Fakultät für Mathematik
Technische Universität München

Technische Universität München
Lehrstuhl für mathematische Physik

Symmetry Methods in Quantum Information Theory

Anna-Lena Karolyn Hashagen

Vollständiger Abdruck der von der Fakultät für Mathematik der Technischen Universität München zur Erlangung des akademischen Grades eines

Doktor der Naturwissenschaften (Dr. rer. nat.)

genehmigten Dissertation.

Vorsitzender:

Prof. Dr. Felix Krahmer

Prüfende der Dissertation:

1. Prof. Dr. Michael M. Wolf
2. Prof. Dr. Otfried Gühne,
   Universität Siegen
3. Prof. Dr. Stephen D. Bartlett,
   The University of Sydney, Australia (schriftliche Beurteilung)

Die Dissertation wurde am 01.10.2018 bei der Technischen Universität München eingereicht und durch die Fakultät für Mathematik am 12.12.2018 angenommen.

Faculty of Mathematics
Technical Univeristy of Munich

Technical University of Munich
Chair of Mathematical Physics

Symmetry Methods in Quantum Information Theory

Anna-Lena Karolyn Hashagen

Full imprint of the dissertation approved by the Faculty of Mathematics of the Technical University of Munich to obtain the academic degree of

Doctor of Natural Sciences (Dr. rer. nat.).

Chairman:

Prof. Dr. Felix Krahmer

Examiners of the dissertation:

1. Prof. Dr. Michael M. Wolf
2. Prof. Dr. Otfried Gühne,
   Universität Siegen
3. Prof. Dr. Stephen D. Bartlett,
   The University of Sydney, Australia

The dissertation was submitted to the Technical University of Munich on 01.10.2018 and was accepted by the Faculty of Mathematics on 12.12.2018.

Abstract

Diese Dissertation behandelt Symmetriemethoden in der Quanteninformationstheorie, insbesondere die Invarianz von Quantenzuständen und Quantenkanälen unter einer gegebenen unitären Darstellung einer endlichen oder kompakten Gruppe. Diese Dissertation analysiert drei unterschiedliche Anwendungsbeispiele, Quantenklonen, Information-Störungs-Austauschbeziehung und randomisierte Benchmarks, in denen Symmetriemethoden genutzt werden, um die Komplexität der Forschungsfrage signifikant zu reduzieren und so eine analytische Lösung zu erreichen.

This dissertation discusses symmetry methods in quantum information theory; in particular, quantum states and quantum channels invariant under a particular unitary representation of a finite or compact group. This thesis analyzes three different examples of application, universal quantum cloning, information-disturbance tradeoffs and randomized benchmarking. Symmetry methods are used to significantly reduce the complexity of the research question, such that it is possible to obtain an analytic solution.

It always seems impossible, until it's done.
– Nelson Mandela

Acknowledgments

This dissertation would not have been possible if it was not for my doctoral adviser Prof. Dr. Michael M. Wolf. I would like to express my sincere gratitude to him for the continuous support of my research, for his patience, the motivation and the immense knowledge. That some want to name the oracle in their next publication "Michael" does not come as a surprise.

I would furthermore like to thank all members of the thesis committee: Prof. Dr. Michael M. Wolf, Prof. Dr. Felix Krahmer, Prof. Dr. Otfried Gühne and Prof. Dr. Stephen D. Bartlett for their time and help.

Besides my adviser and my committee, I would like to thank the whole M5 group. Old, current and new. You created the great atmosphere that turned these Ph.D. years into something very special that I will always love to remember. The group has been a source of friendships as well as good advice and collaboration. I always enjoyed our coffee breaks with stories, cakes and laughter. You also ensured that I learned a lot outside of my research area: I am sure I now know as much about animals as I wished to know, and probably even more. Moreover, I am grateful to our group's administrative assistants Wilma Ghamam and Silvia Schulz who kept us organized and were always ready to help.

I would like to especially thank one person from M5 in particular, who also had the delight of proof-reading this dissertation. My office mate, Daniel Stilck França, with whom I had the greatest time and shared the biggest laughs. Thank you very much, Daniel, for being such an honest and fun soul! You are a true inspiration and surely the best Caipirinha mixer in town. I feel very lucky that I got to share an office with you. My next office mate surely has to step into the biggest footsteps.

Furthermore, I am grateful to all my collaborators: Prof. Dr. Michael M. Wolf, Prof. Dr. David Gross, Prof. Dr. Steven T. Flammia, Prof. Dr. Harald Weinfurter, Dr. Jasmin D. A. Meinecke, Dr. Joel J. Wallmann, Daniel Stilck França, Lukas Knips and Jan Dziewior. It was a great experience to work with and to learn from such distinct researchers. Thank you very much.

I gratefully acknowledge the funding source that made my Ph.D. work possible. I was funded by the Elite Network of Bavaria through the doctorate program "Exploring Quantum Matter". I thank all members of this program and everyone that is involved for such a great work environment and the many stimulating seminars. In particular, I would like to thank my "ExQM Crew", Moritz August, David Leiner and Jakob Wierzbowski, for the trillions coffees that we drank and for the many fun evenings that we spent not discussing research.

A special thank goes to Prof. Dr. Stephen D. Bartlett whose research group at the University of Sydney I visited during my doctoral studies. It was not only great to escape the European winter and enjoy the Australian summer and everything that comes with it, but it was an

amazing experience to be part of your group and to work alongside such splendid researchers. Thank you very much for this opportunity! Thus, I would also like to not miss the chance to thank all the Usyd group members. I had a great time with you all in Australia! You are all like the stereotypical Australian very very kind and you made me feel very welcomed. I would especially like to thank Angela Karanjai, Kamil Korzekwa and Daniel Süß, who I had such a great time with. You are the best!

My time as a doctoral student was made enjoyable by many more great friends, inside and outside of academia. I am grateful for time spent with you all and for all the lovely memories.

Moreover, I would like to thank my parents, Janina Marahrens-Hashagen and Dr. Björn Hashagen, as well as my brother, Jan-Christian Marten Hashagen. Even though they always say that they understand nothing of my research, they were always supportive and always encouraged me to follow my dreams. They fostered my natural interest in mathematics and physics throughout my life and they made it possible for me to have achieved all of what I have today. Especially my older brother, who was always an inspiring example, goad me into pursuing something that I really enjoy and to excel in it. Keeping up with his level of excellence was and still is motivating and inspiring. Thank you for all of this!

Lastly and most of all, I would like to thank Max Büchler (aka Maxi) for all his love and encouragement. He patiently listened to many stories about quantum mechanics and he uncomplainingly read all the manuscripts of my articles as well as the final draft of this dissertation. Your faithful support throughout all stages of this Ph.D. is so appreciated. Thank you.

List of Contributed Articles

Contributed core articles as principal author

1. A. K. Hashagen
   Universal asymmetric quantum cloning revisited
   Quantum Information and Computation, 17(9& 10):0747–0778, August 2017
   (Cf. bibliography entry [1])

2. A. K. Hashagen and M. M. Wolf
   Universality and optimality in the information-disturbance tradeoff
   Accepted in Annales Henri Poincaré, 2018
   (Cf. bibliography entry [2])

4. A. K. Hashagen, S. T. Flammia, D. Gross and J. J. Wallman
   Real randomized benchmarking
   Quantum, 2:85, August 2018
   (Cf. bibliography entry [3])

Contributed further articles as co-author

3. L. Knips, J. Dziewior, A. K. Hashagen, J. D. A. Meinecke, H. Weinfurter and M. M. Wolf
   Measurement-disturbance tradeoff outperforming optimal cloning
   Arxiv e-prints: arXiv:1808.07882 [quant-ph], August 2018
   Submitted to Physical Review Letters
   (Cf. bibliography entry [4])

5. D. S. França and A. K. Hashagen
   Approximate randomized benchmarking for finite groups
   Journal of Physics A: Mathematical and Theoretical, 51(39):395302, August 2018
   (Cf. bibliography entry [5])

I, Anna-Lena Karolyn Hashagen, am principal author of articles 1, 2 and 4. These form the core articles of this dissertation. I am co-author of articles 3 and 5, of which the first one, article 3, is still being reviewed at the moment of the submission of this dissertation.

Contents

List of Figures

1 Introduction

Symmetry methods are widely used in a variety of mathematical fields and physical applications. They draw on a wide range of disciplines and without surprise they can also be found at the heart of quantum information theory. Classifying invariants and reducing complicated objects into a simple canonical form appoint these symmetry methods to a valuable toolbox. These tools are used in this dissertation to study fundamental features of quantum information theory and to answer research questions concerned with information processing tasks.

In particular, this dissertation studies quantum states and quantum channels that are invariant under a particular unitary representation of a compact group. These symmetry considerations then simplify specific fundamental questions in three areas of application. The first area of application is universal asymmetric quantum cloning, in which it can be shown that the optimal quantum cloning channel obeys a specific symmetry. It is then possible to explicitly derive the quantum channel yielding an optimal tradeoff between the individual qualities of the quantum clones. Another area of application is the information-disturbance tradeoff. A two-parameter family of quantum instruments giving the optimal tradeoff between information gain and state disturbance can be derived, again using the symmetry of the underlying optimization problem. The third area of application analyzed in this dissertation is randomized benchmarking. The protocol includes an averaging procedure, which again allows the application of symmetry methods. It is then possible to derive quantitative estimates of the average error of the noise inherit to a physical quantum channel.

I utilize symmetry methods to reduce the complexity of the research question, such that analytic results can be obtained. All articles included in this dissertation fall into these three main areas of application of the symmetry methods,

(a) quantum cloning,

(b) information-disturbance tradeoffs, and

(c) randomized benchmarking.

The research articles included in this dissertation fall into one or two of these main areas. This determines their order of appearance in this dissertation. Article 1 falls into area (a), article 2 is concerned with area (b), article 3 is mainly about area (b), but also treats area (a), article 4 and article 5 solely discuss area (c). Their extensive summary and a statement of the individual contribution is presented in the appendix.

Outline of this dissertation

The following chapters give a short introduction to the mathematical and physical concepts of quantum information theory as well as important symmetry methods used throughout the included articles of this dissertation. Since the articles included are solely concerned with finite-dimensional quantum mechanics, this dissertation only analyzes this special case and thus avoids cumbersome technicalities.

In chapter 2 we introduce the basic formalism of quantum information theory. We start by defining quantum states and measurements performed on systems in specific quantum states in section 2.1 and we give explicit examples commonly used in quantum information theory and used throughout this dissertation. In section 2.2 we define quantum channels and underline their importance to describe physical processes. We discuss different representations of quantum channels that are commonly used in quantum information theory and that manifested themselves as very useful when studying the mathematical and physical properties of quantum channels. These two concepts are then brought together in section 2.3, where we discuss general measurement schemes described by quantum instruments. In chapter 3 we provide an introduction to the symmetry methods used throughout this dissertation. Section 3.1 describes the structure of quantum states under symmetry. We study its general structure, before we give three explicit examples of a symmetry group. In section 3.2 we discuss quantum channels under symmetry. The connection to quantum states under symmetry is established and the same three examples of a symmetry group are discussed. In section 3.3 we define the notion of a group design and discuss an example of a unitary 2-design, the complex Clifford group, and an example of an orthogonal 2-design, the real Clifford group. In chapter 4 we discuss three applications of the symmetry methods presented in the previous chapter. Section 4.1 discusses universal quantum cloning, section 4.2 covers information-disturbance tradeoffs and section 4.3 deals with randomized benchmarking. This chapter shows how we use these symmetry methods to derive analytic results to fundamental questions in quantum information theory.

Throughout the following chapters, references for basic definitions and concepts, commonly used in quantum information theory, may be omitted. In the case of quantum information theory, these can all be found in the famous book by Michael A. Nielsen and Isaac L. Chuang [6] or in the excellent book by Teiko Heinosaari and Mário Ziman [7]. All basic results about finite or compact groups and their representations used throughout the chapter describing symmetry methods in quantum information theory can be found in the marvelous book by Barry Simon [8]. The author does not claim authorship for these results, even if a reference is omitted.

2 Quantum Information Theory

Throughout this dissertation we will only consider finite dimensional Hilbert spaces $\mathbb{C}^d$. We will write $\mathcal{M}_{d_1,d_2}$ for the space of $d_1 \times d_2$ matrices and in the case that $d_1 = d_2 = d$ we use the abbreviation $\mathcal{M}_d$. For any $X \in \mathcal{M}_{d_1,d_2}$, we will denote by $X^t \in \mathcal{M}_{d_2,d_1}$ the usual transposition, by $\bar{X} \in \mathcal{M}_{d_1,d_2}$ the complex conjugation and by $X^* = \bar{X}^t \in \mathcal{M}_{d_2,d_1}$ the conjugate transpose. A matrix $X \in \mathcal{M}_d$ that is equal to its own conjugate transpose, i.e. $X^* = X$, is called hermitian. A matrix $X \in \mathcal{M}_d$ is called positive semidefinite if and only if (iff) it is a hermitian matrix with solely non-negative eigenvalues and we will denote this by $X \geq 0$. The set of all positive semidefinite matrices forms a cone and we will denote it by $\mathcal{M}_d^+ \subset \mathcal{M}_d$. Moreover, for a linear map $T : \mathcal{M}_{d_1} \to \mathcal{M}_{d_2}$ we will denote its adjoint with respect to (w.r.t.) the Hilbert-Schmidt inner product by $T^* : \mathcal{M}_{d_2} \to \mathcal{M}_{d_1}$. Furthermore, let $\mathbb{1}_d \in \mathcal{M}_d$ denote the identity matrix on $\mathcal{M}_d$ and we will use the notation $\mathrm{id}_d : \mathcal{M}_d \to \mathcal{M}_d$ to denote the identity map.

In the following, subscripts denoting the dimension might be omitted if these are clear from context. We use the bra-ket notation, also called Dirac notation, as it is the standard notation used in quantum mechanics [6, 9].

2.1 Quantum states and measurements

Quantum mechanics divides every physical experiment into the preparation of a quantum state and the measurement of an observable associated with a quantum system in that quantum state. A preparation scheme of a quantum system determines the probability distribution of any possible measurement. Therefore, a quantum state may be regarded as the family of preparation schemes that yields identical outcome distributions for all measurements. Similarly, a measurement may be regarded as the family of measurement schemes that yields identical outcome distributions for all quantum states in a statistical experiment.

2.1.1 Quantum states

Every d-dimensional *quantum state*, $d \in \mathbb{N}$, is described by a density matrix $\rho \in \mathcal{M}_d^+$ that is positive semidefinite and has trace 1. In the special cases $d = 2$ and $d = 3$, quantum states are usually called qubits and qutrits, respectively. Analogously, d-dimensional quantum states are sometimes referred to as qudits. We will denote by $\mathcal{D}_d \subseteq \mathcal{M}_d$ the set of all d-dimensional quantum states, i.e.,

$$\mathcal{D}_d := \{\rho \in \mathcal{M}_d \,|\, \rho \geq 0, \mathrm{Tr}[\rho] = 1\}\,. \tag{2.1}$$

This set is convex, i.e., for any $\rho, \sigma \in \mathcal{D}_d$ and $0 \leq \lambda \leq 1$ implies $\lambda\rho + (1-\lambda)\sigma \in \mathcal{D}_d$. The extremal points of this set, i.e., those elements that do not admit a proper convex decomposition, are

called *pure states* and are the one-dimensional projectors $|\psi\rangle\langle\psi|$ for unit vectors $|\psi\rangle \in \mathbb{C}^d$. Any quantum state that is not pure is called *mixed* and we will call $\frac{\mathbb{1}}{d} \in \mathcal{D}_d$ the *maximally mixed quantum state*.

In quantum theory, composite systems are described using tensor products. A multipartite quantum system consisting of $n \in \mathbb{N}$ quantum states with dimensions $d_1, d_2, \ldots, d_n \in \mathbb{N}$ is described by $\mathcal{D}(\mathbb{C}^{d_1} \otimes \ldots \otimes \mathbb{C}^{d_n}) \subset \bigotimes_{i=1}^n \mathcal{M}_{d_i}$. The simplest example is a bipartite system $\rho_{1,2} \in \mathcal{D}(\mathbb{C}^{d_1} \otimes \mathbb{C}^{d_2})$. The *restrictions* of $\rho_{1,2}$ are then obtained using the *partial trace*, where the trace map $\mathrm{Tr} : \mathcal{M}_d \to \mathbb{C}$ is just applied to a subsystem. Thus, if we would like to trace out the second system to obtain the first restricted quantum state we apply the partial trace $\mathrm{Tr}_2 := \mathrm{id} \otimes \mathrm{Tr}$ to $\rho_{1,2}$. The resulting state is then usually denoted by omitting the label of the traced out system, i.e., $\rho_1 = (\mathrm{id} \otimes \mathrm{Tr})(\rho_{1,2}) = \mathrm{Tr}_2(\rho_{1,2})$. We will also call these restrictions *reduced states* or *marginals*.

Multipartite quantum systems may exhibit a characteristic, called *entanglement*, that distinguishes them from their classical counterparts. Informally, we speak of entanglement if the correlation exhibited between the individual systems is not of classical nature. Consider a bipartite quantum state $\rho \in \mathcal{D}(\mathbb{C}^{d_1} \otimes \mathbb{C}^{d_2})$. It is called *separable* or *classically correlated* if there is $m \in \mathbb{N}$ such that it can be written as

$$\rho = \sum_{i=1}^{m} \lambda_i \rho_i^{(1)} \otimes \rho_i^{(2)}, \tag{2.2}$$

with quantum states $\{\rho_i^{(k)}\}_{i=1}^m \subset \mathcal{D}_{d_k}$ and a probability distribution $\{\lambda_i\}_{i=1}^m$ as weights for all $k \in \{1, 2\}$. Otherwise $\rho \in \mathcal{D}(\mathbb{C}^{d_1} \otimes \mathbb{C}^{d_2})$ is called *entangled*.

The most extreme cases regarding entanglement arise if the marginals are maximally mixed. Consider the Hilbert space $\mathbb{C}^d \otimes \mathbb{C}^d$ with the computational basis given by the standard unit vectors $\{|i\rangle\}_{i=1}^d$ on those two spaces. Then we define the maximally entangled state

$$\Omega_d := |\Omega\rangle\langle\Omega| \qquad \text{with} \qquad |\Omega\rangle = \frac{1}{\sqrt{d}} \sum_{i}^{d} |i\rangle \otimes |i\rangle . \tag{2.3}$$

Note that both marginals are maximally mixed, i.e., $\mathrm{Tr}_1[\Omega_d] = \mathrm{Tr}_2[\Omega_d] = \frac{\mathbb{1}}{d}$. It turns out that every maximally entangled state is of the form $(\mathbb{1} \otimes U)\,|\Omega\rangle$, where U is some unitary matrix. However, if we talk about the maximally entangled state, we refer to the one given in the computational basis. The maximally entangled quantum state is closely related to the flip (or swap) operation, defined through $\mathbb{F}\,|i\rangle \otimes |j\rangle = |j\rangle \otimes |i\rangle$ by the correspondence,

$$d\,|\Omega\rangle\langle\Omega|^{t_2} = \mathbb{F}, \tag{2.4}$$

where $\cdot^{t_2}$ denotes the partial transposition on the second system, i.e., for any bipartite $\rho \in \mathcal{D}(\mathbb{C}^d \otimes \mathbb{C}^d)$ the partial transposition is defined as

$$\rho \mapsto \rho^{t_2} := \left(\mathrm{id} \otimes \cdot^t\right)(\rho). \tag{2.5}$$

2.1.2 Measurements

An m-outcome measurement on the space of quantum states is described by a *positive operator-valued measure* (POVM) $E := \{E_i\}_{i=1}^m$, whose elements $E_i \in \mathcal{M}_d^+$, called effect operators, are

positive semidefinite and sum up to the identity $\sum_{i=1}^{m} E_i = \mathbb{1}$. If the measurement is performed on a system in quantum state $\rho \in \mathcal{D}_d$, we get the measurement outcome i with probability $\mathrm{Tr}[\rho E_i]$. This is illustrated in figure 2.1. We denote the set of all such POVMs as $\mathcal{E}_{d,m}$ and we use the shorthand notation $\mathcal{E}_d := \mathcal{E}_{d,d}$. If the effect operators E_i are mutually orthogonal projections, we call the POVM a *von Neumann measurement*.

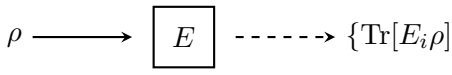


Figure 2.1: Measurement of a system in quantum state $\rho \in \mathcal{D}_d$ represented by a POVM $E \in \mathcal{E}_{d,m}$. The measurement outcome i is observed with probability $\mathrm{Tr}[E_i\rho]$.

We call a POVM *informationally complete*, if the statistics obtained through the measurement allow for a full description of the quantum state. This is maximally efficient, if the measurement's effect operators are rank-one projections. An especially interesting case occurs if these effect operators are symmetric with respect to the Hilbert-Schmidt inner product [10, 11]. This is called a symmetric, informationally complete (SIC) POVM, i.e., a set of d^2 subnormalized rank-one projectors $E = \{P_i/d\}_{i=1}^{d^2} \in \mathcal{E}_{d^2,d}$ with equal pairwise Hilbert-Schmidt inner product, $\mathrm{Tr}[P_iP_j]/d^2 = 1/(d^2(d+1))$ for all $i \neq j$.

We have now defined a quantum measurement of a system in a quantum state. In order to fully describe a quantum measurement scheme, the notion of a quantum channel is needed. We therefore discuss completely positive and trace-preserving linear maps, also called quantum channels, in the next section.

2.2 Quantum channels

Considering the time evolution of our quantum system yields different (but equivalent) pictures related to the splitting of a physical process into preparation and measurement. If the evolution is part of the measurement process, we speak of the *Heisenberg picture*. If, on the other hand, we allow the quantum states to evolve with time, we usually refer to this as the *Schrödinger picture*. In the following, we will mostly adopt the later viewpoint.

Any physical process should be represented by a linear map that maps quantum states to quantum states. It should therefore preserve the positivity as well as the trace characterizing quantum states. A linear map $T : \mathcal{M}_{d_1} \to \mathcal{M}_{d_2}$ is called *positive* iff $T(X) \geq 0$ whenever $X \geq 0$ and *trace-preserving* iff $\mathrm{Tr}[T(X)] = \mathrm{Tr}[X]$ for any $X \in \mathcal{M}_{d_1}$. Positivity alone is, however, not sufficient for a full description of a physical process. Consider a system which is part of a larger system, such as a reduced state in a bipartite setting. If the linear map acts solely onto that subsystem, we still require the overall system to remain positive. We therefore require the stronger notion of complete positivity. A linear map $T : \mathcal{M}_{d_1} \to \mathcal{M}_{d_2}$ is called *completely positive* iff $(\mathrm{id}_n \otimes T) : \mathcal{M}(\mathbb{C}^n \otimes \mathbb{C}^{d_1}) \to \mathcal{M}(\mathbb{C}^n \otimes \mathbb{C}^{d_2})$ is positive for every $n \in \mathbb{N}$. This motivates the following definition.

Definition 2.1 (Quantum channel)**.** A *quantum channel* is a linear map

$$T : \mathcal{M}_{d_1} \to \mathcal{M}_{d_2} \tag{2.6}$$

that is completely positive and trace preserving. The set of all quantum channels of the form (2.6) is denoted by $\mathcal{T}_{d_1,d_2}$, and we will write $\mathcal{T}_{d_1} := \mathcal{T}_{d_1,d_1}$.

This definition is illustrated in figure 2.2. An example of a quantum channel is the identity channel, $\mathrm{id}_d : \mathcal{M}_d \to \mathcal{M}_d$. Intuitively, the identity channel represents the ideal quantum communication channel as the quantum state it acts upon is not altered in the process.

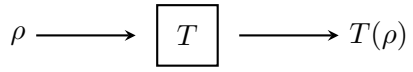


Figure 2.2: A quantum channel $T \in \mathcal{T}_{d_1,d_2}$ acting on a system in a quantum state $\rho \in \mathcal{D}_{d_1}$. The system evolves to the quantum state $T(\rho) \in \mathcal{D}_{d_2}$.

In the Heisenberg picture quantum channels are linear maps

$$T^* : \mathcal{M}_{d_2} \to \mathcal{M}_{d_1}, \tag{2.7}$$

which are completely positive and unital, i.e., $T^*(\mathbb{1}_{d_2}) = \mathbb{1}_{d_1}$. They correspond to the dual map of the quantum channel in the Schrödinger picture with respect to the Hilbert-Schmidt inner product. Performing a measurement described by a POVM $E = \{E_i\}_{i=1}^m \in \mathcal{E}_{d_2,m}$, the outcome probabilities of an evolved state when measured are $\mathrm{Tr}[E_i T(\rho)] = \mathrm{Tr}[T^*(E_i)\rho]$ and thus equal in the two different pictures.

The following section introduces different representations of quantum channels defined above; these are very useful and widely used through quantum information literature.

2.2.1 Representations of quantum channels

The following one-to-one correspondence between linear maps and operators turns out to be very useful. In particular, in the study of quantum channels the theorem allows to infer its properties by the study of the corresponding quantum state.

Theorem 2.2 (Choi-Jamiolkowski isomorphism [12])**.** *The mapping*

$$J : T \mapsto J_T := (T \otimes \mathrm{id}_{d_1})\Omega_{d_1}, \tag{2.8}$$

defines the so-called Choi-Jamiolkowski isomorphism between the vector spaces of linear maps on a d_1-dimensional system $\{T : \mathcal{M}_{d_1} \to \mathcal{M}_{d_2} \mid T \text{ linear}\}$ and linear operators on a $(d_2 \times d_1)$-dimensional Hilbert space $\mathcal{M}(\mathbb{C}^{d_2} \otimes \mathbb{C}^{d_1})$. The matrix $d_1 J_T$ is often called Choi matrix, and if T is a quantum channel, then J_T is usually referred to as the corresponding Jamiolkowski state.

This theorem gives rise to many useful correspondences [12]:

(a) A linear map $T : \mathcal{M}_{d_1} \to \mathcal{M}_{d_2}$ is *completely positive* iff the corresponding Choi matrix is positive semidefinite, i.e., iff $J_T \geq 0$.

(b) A linear map $T : \mathcal{M}_{d_1} \to \mathcal{M}_{d_2}$ is *unital*, i.e., $T(\mathbb{1}) = \mathbb{1}$, iff $\mathrm{Tr}_2[J_T] = \mathbb{1}_{d_2}/d_1$.

(c) A linear map $T : \mathcal{M}_{d_1} \to \mathcal{M}_{d_2}$ is *trace-preserving*, i.e., $T^*(\mathbb{1}) = \mathbb{1}$, iff $\mathrm{Tr}_1[J_T] = \mathbb{1}_{d_1}/d_1$.

(d) A linear map $T : \mathcal{M}_{d_1} \to \mathcal{M}_{d_2}$ is *hermiticity-preserving*, i.e., $T(X) = T(X)^*$ for all $X = X^* \in \mathcal{M}_{d_1}$, iff the corresponding Choi matrix is hermitian, i.e., $J_T = J_T^*$.

Furthermore, this theorem 2.2 gives rise to the following very useful representation of quantum channels. It stems from the convex decomposition of the Jamiolkowski state into rank-one matrices.

Theorem 2.3 (Kraus representation [12, 13])**.** *A liner map $T : \mathcal{M}_{d_1} \to \mathcal{M}_{d_2}$ is completely positive iff there is a $r \in \mathbb{N}$ such that for any $X \in \mathcal{M}_{d_1}$ it can be written in the form*

$$T(X) = \sum_{j=1}^{r} K_j X K_j^*, \tag{2.9}$$

with matrices $\{K_j \in \mathbb{C}^{d_2 \times d_1}\}_{i=1}^r$, which are called the Kraus operators of T. Furthermore, T is trace-preserving iff the Kraus operators satisfy $\sum_{j=1}^r K_j^* K_j = \mathbb{1}$ and T is unital iff $\sum_{j=1}^r K_j K_j^* = \mathbb{1}$. Moreover, the minimum number of Kraus operators is called the Kraus-rank and $r \leq d_1 d_2$.

We have now thoroughly discussed quantum states, quantum measurements as well as quantum channels and are therefore equipped to introduce quantum instruments. These describe a general measurement scheme, a key concept in this dissertation [2, 4], and are defined in the following section.

2.3 Quantum instruments

A general measurement scheme describes the statistics of the measurement results and the evolution of the quantum system on which the measurement is performed. This is described by a *quantum instrument*, illustrated in figure 2.3. Quantum instruments were first introduced by [14] and a detailed description can be found in [15, Chapter 2.3].

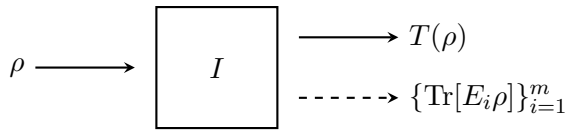


Figure 2.3: A general measurement scheme described by a quantum instrument.

A quantum instrument is a set of completely positive linear maps $I := \{I_i : \mathcal{M}_{d_1} \to \mathcal{M}_{d_2}\}_{i=1}^m$ that fulfill the normalization condition $\sum_{i=1}^m I_i^*(\mathbb{1}) = \mathbb{1}$, where the $I_i^* : \mathcal{M}_{d_2} \to \mathcal{M}_{d_1}$ denote the corresponding dual maps of the elements of the quantum instrument. In this general measurement scheme, if the measurement on the system in state $\rho \in \mathcal{D}_d$ gives the outcome i with probability $p_i := \mathrm{Tr}[I_i(\rho)]$, then the quantum state *after* the measurement is $I_i(\rho)/p_i$.

Therefore, the two marginals of our quantum instrument are the inherent quantum measurement and the inherent quantum channel. We identify the inherent POVM to be $E := \{E_i := I_i^*(\mathbb{1})\}_{i=1}^m \in \mathcal{E}_{d_1,m}$, such that the measurement outcome i is observed with probability

$$p_i = \mathrm{Tr}[I_i(\rho)] = \mathrm{Tr}[I_i^*(\mathbb{1})\rho] = \mathrm{Tr}[E_i\rho]. \tag{2.10}$$

The normalization condition on the quantum instrument ensures that the probabilities add up to one, as expected. On the other hand, if we ignore the measurement outcomes, the sum of the elements of the quantum instrument yields a quantum channel $T : \mathcal{M}_{d_1} \to \mathcal{M}_{d_2}$, i.e.,

$$T(\cdot) := \sum_{i=1}^{m} I_i(\cdot), \tag{2.11}$$

where the normalization condition of the quantum instrument gives the quantum channel its trace-preserving property.

The notion of a quantum instrument naturally describes the fact that a POVM alone does not determine the quantum state after the measurement. Rather the full description of the quantum instrument is needed to describe a general measurement scheme as depicted in figure 2.3. That is, every POVM corresponds to a whole equivalence class of instruments.

The most well known example is the so-called *Lüders instrument* associated to a POVM $E \in \mathcal{E}_{d_1,m}$. It is defined to be the quantum instrument I with elements

$$I_i(\cdot) = \sqrt{E_i} \cdot \sqrt{E_i}, \quad i = 1, \ldots, m, \tag{2.12}$$

where for any $X \in \mathcal{M}_d$ with $X \geq 0$, $\sqrt{X} \in \mathcal{M}_d$ is such that $(\sqrt{X})^* \sqrt{X} = X$. In the special case of a von Neumann measurement, if we measure the outcome i, the system, initially in quantum state ρ, evolves to the post-measurement quantum state $E_i \rho E_i / \operatorname{Tr}[E_i \rho]$. This von Neumann-Lüders measurement is well-known from standard quantum mechanics and arises as a special case in the general quantum instrument description of a measurement [6].

With the help of quantum instruments, it is now possible to formalize the intuitive notion of state disturbance through a measurement. If a measurement is performed on a system in some quantum state in order to gain some information about this system, we necessarily have to introduce some disturbance to the system. This idea is described in the following proposition.

Proposition 2.4 (No information without disturbance [7])**.** *Consider a quantum instrument represented by a set of completely positive linear maps $I := \{I_i : \mathcal{M}_d \to \mathcal{M}_d\}$. If there is no disturbance on average, i.e., $T := \sum_i I_i$ satisfies $T = \mathrm{id}$, then the probability of obtaining an outcome i, which is given by $\operatorname{Tr}[I_i(\rho)]$, is independent of the input quantum state $\rho \in \mathcal{M}_d$. Therefore, no information is gained about this initial quantum state.*

Proof. Using the one-to-one correspondence between a completely positive linear map and its Choi matrix, given in theorem 2.2, the fact that there is no disturbance on average, i.e., $\sum_i I_i = \mathrm{id}$, reads

$$\sum_i J_{I_i} = |\Omega\rangle\langle\Omega| \,. \tag{2.13}$$

The I_i are completely positive, such that $J_{I_i} \geq 0$. Equation (2.13) thus corresponds to a convex decomposition of the maximally entangled state. This is, however, a pure state, such that the decomposition must be trivial, i.e., $J_{I_i} = c_i \, |\Omega\rangle\langle\Omega|$ for some constants $c_i \geq 0$. Therefore, $I_i = c_i \, \mathrm{id}_d$, such that the probability of obtaining outcome i is $\operatorname{Tr}[I_i(\rho)] = c_i \operatorname{Tr}[\rho] = c_i$ and is thus independent of ρ. □

Gaining information about an unknown quantum state without introducing any disturbance to the quantum system would be possible, if the following cloning device was viable:

Theorem 2.5 (No-Cloning Theorem [16])**.** *There is no quantum channel, $T : \mathcal{M}_d \to \mathcal{M}_d \otimes \mathcal{M}_d$ such that for all quantum states $\rho \in \mathcal{D}_d$ the following holds,*

$$T(\rho) = \rho \otimes \rho. \tag{2.14}$$

Proof. The theorem is a consequence of linearity. Let $|\psi\rangle\langle\psi|, |\phi\rangle\langle\phi| \in \mathcal{D}_d$ be two orthogonal pure states. If there was a map T as specified in the theorem, then for $\lambda \in [0, 1]$

$$\lambda T(|\psi\rangle\langle\psi|) + (1-\lambda)T(|\phi\rangle\langle\phi|) = \lambda\left(|\psi\rangle\langle\psi| \otimes |\psi\rangle\langle\psi|\right) + (1-\lambda)\left(|\phi\rangle\langle\phi| \otimes |\phi\rangle\langle\phi|\right), \tag{2.15}$$

which has rank 2, while,

$$T\left(\lambda|\psi\rangle\langle\psi| + (1-\lambda)|\phi\rangle\langle\phi|\right) = \left(\lambda|\psi\rangle\langle\psi| + (1-\lambda)|\phi\rangle\langle\phi|\right) \otimes \left(\lambda|\psi\rangle\langle\psi| + (1-\lambda)|\phi\rangle\langle\phi|\right), \tag{2.16}$$

which has rank 2^2. □

The no-cloning theorem tells us that it is impossible to perfectly clone an unknown quantum state. It is, however, possible to do this in an approximate manner, which has given rise to many applications of the no-cloning theorem in different areas of quantum information theory [17, 18]. Therefore, even though proposition 2.4 and theorem 2.5, as no-go theorems, are negative in nature, they do open up a fruitful field of applications and fundamental questions in quantum information theory [1, 2, 4], which is explored in the articles included in this dissertation.

3 Symmetry Methods

This chapter describes the symmetry methods used throughout the dissertation. For an extensive review of representations of finite and compact groups, which are fundamental for the study of symmetries, the reader is referred to [8].

Consider a finite or compact group $\mathbf{G}$ with elements $g \in \mathbf{G}$ and a unitary representation $U := \{U_g\}_{g\in\mathbf{G}}$ on a finite-dimensional Hilbert space $\mathbb{C}^d$. This unitary representation can be written as a direct sum of irreducible unitary representations (irreps), i.e.,

$$U \simeq \bigoplus_{i=1}^{k} n_i U^i, \tag{3.1}$$

where $n_i > 0$ denotes the degeneracy of the ith irrep. The underlying Hilbert space decomposes as

$$\mathbb{C}^d \simeq \bigoplus_{i=1}^{k} \left(\mathbb{C}^{d_i} \otimes \mathbb{C}^{n_i}\right), \tag{3.2}$$

where the sum runs over all irreps of the group $\mathbf{G}$ and $\mathbb{C}^{d_i}$ carries the ith irrep. Every U_g, $g \in \mathbf{G}$, is then block-diagonal with respect to this decompositions, i.e., it is of the form

$$U_g \simeq \bigoplus_{i=1}^{k} U_g^i \otimes \mathbb{1}_{n_i}\,. \tag{3.3}$$

Let $\mathcal{A}$ be the algebra of operators generated by the unitary representation $\{U_g\}_{g\in\mathbf{G}}$. Then its commutant, denoted by $\mathcal{A}'$, is defined as

$$\mathcal{A}' := \{B \,|\, BA = AB \text{ for all } A \in \mathcal{A}\}\,. \tag{3.4}$$

Considering the decomposition of the underlying Hilbert space, it is clear that

$$\mathcal{A} = \left\{ \bigoplus_{i=1}^{k} A_i \otimes \mathbb{1}_{n_i} \,\middle|\, A_i \in \mathcal{M}_{d_i} \right\}, \tag{3.5}$$

and the commutant has the form,

$$\mathcal{A}' = \left\{ \bigoplus_{i=1}^{k} \mathbb{1}_{d_i} \otimes B_i \,\middle|\, B_i \in \mathcal{M}_{n_i} \right\}. \tag{3.6}$$

If the group $\mathbf{G}$ is compact, it is well-known that there exists a unique probability measure μ that is invariant under left and right group multiplication [8, Chapter VII.3]. This measure is called *Haar measure* and we will denote integrals with respect to this Haar measure by $\int_{\mathbf{G}} \mathrm{d}\mu$.

The following groups are of special relevance in the upcoming discussion about symmetries. Their definitions are given now.

Definition 3.1 (Unitary group)**.** The unitary group $\mathbf{U}(d)$ is the group of $d \times d$ matrices obeying

(a) $U^*U = UU^* = \mathbb{1}$.

Definition 3.2 (Special unitary group)**.** The special unitary group $\mathbf{SU}(d)$ is the group of $d \times d$ matrices obeying

(a) $U^*U = UU^* = \mathbb{1}$ and

(b) $\det(U) = 1$.

Definition 3.3 (Orthogonal group)**.** The orthogonal group $\mathbf{O}(d)$ is the group of $d \times d$ real matrices obeying

(a) $O^tO = OO^t = \mathbb{1}$.

Definition 3.4 (Special orthogonal group)**.** The special orthogonal group $\mathbf{SO}(d)$ is the group of $d \times d$ real matrices obeying

(a) $O^tO = OO^t = \mathbb{1}$ and

(b) $\det(O) = 1$.

We have described the mathematical basics in terms of unitary representations of finite or compact groups and the decomposition of the underlying Hilbert space with respect to the irreducible unitary representations. Furthermore, we have defined the commutant of unitary representations of a finite or compact group. Moreover, we have introduced important groups, that will be used throughout the next section. With all these details at hand, we now study quantum states and quantum channels under symmetry.

3.1 Quantum states under symmetry

We say that a quantum state $\rho \in \mathcal{D}_d$ *obeys a symmetry* of a unitary representation $U = \{U_g\}_{g \in \mathbf{G}}$ of some group $\mathbf{G}$ on $\mathbb{C}^d$, if it is invariant with respect to that symmetry, i.e., if for all $g \in \mathbf{G}$

$$\rho = U_g \rho U_g^* \tag{3.7}$$

holds. Instead of studying the symmetry behavior of specific quantum states, it is of interest to fix a specific symmetry group and then study all quantum states that are invariant under this symmetry group.

We are in particular interested in bipartite quantum states $\rho \in \mathcal{D}(\mathbb{C}_d \otimes \mathbb{C}_d)$ and we focus on closed groups $\mathbf{G}$ of unitaries of the form $U = (U^{(1)} \otimes U^{(2)})$, with unitary $U^{(1)}, U^{(2)} \in \mathbf{U}(d)$. Please note that we simplified our notation, since in the following we always refer to this specific unitary representation. As a closed subgroup of the unitary group, this group $\mathbf{G}$ must be compact and therefore carries a unique Haar measure. This may be used to define the following concept of a twirl.

Definition 3.5 (Twirl of a quantum state)**.** Let $\rho \in \mathcal{D}(\mathbb{C}_d \otimes \mathbb{C}_d)$ be a quantum state, $\mathbf{G}$ a finite or compact group with unitary representation of the form $U = (U^{(1)} \otimes U^{(2)})$, with unitary $U^{(1)}, U^{(2)} \in \mathbf{U}(d)$ and with Haar measure μ. We denote the *twirl* of ρ with respect to $\mathbf{G}$ as $\mathbb{T} : \mathcal{M}_{d^2} \to \mathcal{M}_{d^2}$ and define it to be

$$\mathbb{T}(\rho) := \int_{\mathbf{G}} \left(U^{(1)} \otimes U^{(2)}\right) \rho \left(U^{(1)} \otimes U^{(2)}\right)^* \, \mathrm{d}\mu. \tag{3.8}$$

The twirl maps every matrix $A \in \mathcal{M}_{d^2}$ into the commutant $\mathbf{G}'$ which is the algebra

$$\mathbf{G}' := \left\{ B \in \mathcal{M}_{d^2} \middle| \forall \left(U^{(1)} \otimes U^{(2)}\right) \in \mathbf{G} : \left[B, \left(U^{(1)} \otimes U^{(2)}\right)\right] = 0 \right\}, \tag{3.9}$$

that commutes with all elements of $\mathbf{G}$.[1] This twirl operation is completely positive and doubly stochastic, that is, it maps density matrices to density matrices and the identity to itself. Furthermore, it is a projection. In accordance with equation (3.6), every $B \in \mathbf{G}'$ is thus of the form,

$$B \simeq \bigoplus_{i=1}^{k} \mathbb{1}_{d_i} \otimes B_i, \tag{3.10}$$

with $d_i = \dim \mathbb{C}^{d_i}$ and B_i acting on the n_i-dimensional space appearing in equation (3.3). In the following we will focus on finite or compact groups $\mathbf{G}$ of which all irreducible unitary representations on $\mathbb{C}^d$ are non-degenerate, i.e., for which $n_i = 1$ holds for all $i = \{1, \ldots, k\}$. In this case, the commutant $\mathbf{G}'$ is an abelian algebra [19] that is spanned by a set of orthogonal minimal projections $\{P_i\}_{i=1}^k$, such that equation (3.10) is of the form

$$B = \bigoplus_{i=1}^{k} x_i P_i, \tag{3.11}$$

where $x_i = \mathrm{Tr}[BP_i]/d_i$ and the P_i are the orthogonal projections onto the ith irrep.

In particular, one can show that in the non-degenerate case the twirl is explicitly given by

$$\mathbb{T}(\cdot) = \sum_{i=1}^{k} \frac{1}{d_i} \mathrm{Tr}[\,\cdot\, P_i] P_i, \tag{3.12}$$

through the observation that for all $A \in \mathcal{M}_{d^2}$ we have that

$$\begin{aligned} \mathbb{T}(A) &= \sum_{i=1}^{k} \frac{1}{d_i} \mathrm{Tr}\left[\mathbb{T}(A) P_i\right] P_i \\ &= \sum_{i=1}^{k} \frac{1}{d_i} \mathrm{Tr}\left[A \mathbb{T}(P_i)\right] P_i \\ &= \sum_{i=1}^{k} \frac{1}{d_i} \mathrm{Tr}\left[A P_i\right] P_i, \end{aligned}$$

where we have used the fact that the twirl is self-adjoint with respect to the Hilbert-Schmidt inner product.

[1] Please note again that we use a simplified notation so by $U \in \mathbf{G}$, we strictly speaking refer to the unitary representation U_g, $g \in \mathbf{G}$.

Using the invariance of the Haar measure, the following lemma illustrates the correspondence between the invariance of a quantum state under the twirl and its membership of the commutant.

Lemma 3.6. *For a quantum state $\rho \in \mathcal{D}(\mathbb{C}_d \otimes \mathbb{C}_d)$,*

$$\left[\rho, U^{(1)} \otimes U^{(2)}\right] = 0 \quad \forall\, U^{(1)}, U^{(2)} \in \mathbf{U}(d) \tag{3.13}$$

is equivalent to

$$\int_{\mathbf{G}} \left(U^{(1)} \otimes U^{(2)}\right) \rho \left(U^{(1)} \otimes U^{(2)}\right)^* \mathrm{d}\mu = \rho. \tag{3.14}$$

Proof. If

$$\left[\rho, U^{(1)} \otimes U^{(2)}\right] = 0 \quad \forall\, U^{(1)}, U^{(2)} \in \mathbf{U}(d)$$

then

$$\int_{\mathbf{G}} \left(U^{(1)} \otimes U^{(2)}\right) \rho \left(U^{(1)} \otimes U^{(2)}\right)^* \mathrm{d}\mu = \int_{\mathbf{G}} \rho \left(U^{(1)} \otimes U^{(2)}\right) \left(U^{(1)} \otimes U^{(2)}\right)^* \mathrm{d}\mu = \rho.$$

The other direction follows from

$$\int_{\mathbf{G}} \left(U^{(1)} \otimes U^{(2)}\right) \rho \left(U^{(1)} \otimes U^{(2)}\right)^* \mathrm{d}\mu = \rho.$$

because then for unitary $V^{(1)}, V^{(2)} \in \mathbf{U}(d)$ we have

$$\begin{aligned}
&\rho \left(V^{(1)} \otimes V^{(2)}\right) \\
&= \int_{\mathbf{G}} \left(U^{(1)} \otimes U^{(2)}\right) \rho \left(U^{(1)} \otimes U^{(2)}\right)^* \mathrm{d}\mu \left(V^{(1)} \otimes V^{(2)}\right) \\
&= \left(V^{(1)} \otimes V^{(2)}\right) \left(V^{(1)} \otimes V^{(2)}\right)^* \int_{\mathbf{G}} \left(U^{(1)} \otimes U^{(2)}\right) \rho \left(U^{(1)} \otimes U^{(2)}\right)^* \mathrm{d}\mu \left(V^{(1)} \otimes V^{(2)}\right) \\
&= \left(V^{(1)} \otimes V^{(2)}\right) \int_{\mathbf{G}} \left(U^{(1)} \otimes U^{(2)}\right) \rho \left(U^{(1)} \otimes U^{(2)}\right)^* \mathrm{d}\mu \\
&= \left(V^{(1)} \otimes V^{(2)}\right) \rho,
\end{aligned}$$

where we have used the invariance property. □

Therefore, in order to study the symmetry behavior, we can focus on analyzing the commutant and its structure. In the following, we will look at three specific examples of a symmetry group, which are of particular interest [20]. We study the UU-invariant quantum states, the $U\bar{U}$-invariant quantum states and the OO-invariant quantum states.

3.1.1 UU-invariant states – Werner states

The first and most prominent example is the *Werner state* [21, 22]. These are all quantum states $\rho \in \mathcal{D}(\mathbb{C}_d \otimes \mathbb{C}_d)$ that are invariant under unitaries of the form $(U \otimes U)$, with unitary $U \in \mathbf{U}(d)$, forming a representation of a group $\mathbf{G}$ on $\mathbb{C}^d \otimes \mathbb{C}^d$. The notation reflects that

this is just the unitary group and we therefore write the projection of any quantum state $\rho \in \mathcal{D}(\mathbb{C}^d \otimes \mathbb{C}^d)$ onto the Werner states, the *UU-twirl*, as

$$\mathbb{T}^{(UU)}(\rho) = \int_{\mathbf{U}(d)} (U \otimes U)\, \rho\, (U \otimes U)^* \, \mathrm{d}\mu. \tag{3.15}$$

The structure of all Werner states can be analyzed using a well-known result from group theory, given in the following theorem [1].

Theorem 3.7 (Weyl [23, Chapter IV])**.** *If an operator ρ acting on the n-fold tensor product $\left(\mathbb{C}^d\right)^{\otimes n}$ obeys $[\rho, U^{\otimes n}] = 0$ for all unitaries $U \in \mathbf{U}(d)$, then it is a linear combination of operators V_π representing the permutation group on $\left(\mathbb{C}^d\right)^{\otimes n}$, i.e., it is of the form*

$$\rho = \sum_{\pi \in \mathbf{S}_n} a_\pi V_\pi, \tag{3.16}$$

where $\mathbf{S}_n$ is the symmetric group on n elements, π are all possible permutations of n elements and $a_\pi \in \mathbb{C}$. The permutation operators V_π are defined via

$$V_\pi\, (v_1 \otimes \ldots \otimes v_n) = v_{\pi^{-1}(1)} \otimes \ldots \otimes v_{\pi^{-1}(n)}. \tag{3.17}$$

Proof. The theorem immediately follows from [8, Theorem IX.11.5]. Denote by $\mathbf{SU}(d)$ the special unitary group of finite degree d and by $\mathbf{S}_n$ the symmetric group on n elements. Let $\mathcal{A}$ be the group algebra of $\mathbf{SU}(d)$ and $\mathcal{B}$ be the group algebra of $\mathbf{S}_n$ generated by their respective unitary representation on $\left(\mathbb{C}^d\right)^{\otimes n}$. Since $\mathbf{SU}(d)$ and $\mathbf{S}_n$ act dually on $\left(\mathbb{C}^d\right)^{\otimes n}$, we get $\mathcal{A}' = \mathcal{B}$. The commutant is therefore exactly the algebra generated by the permutation operators V_π. Thus, if an operator commutes with all unitaries of the form $U^{\otimes n}$, it must be an element of this commutant, i.e. it must be a linear combination of permutation operators. □

Therefore, the most simple version of this theorem applies to our case with $n = 2$. We, therefore, only need to consider two permutation operators, namely the identity $\mathbb{1} \in \mathcal{M}_{d^2}$ and the flip (or swap) operator $\mathbb{F} \in \mathcal{M}_{d^2}$, which was already introduced in section 2.1 in equation (2.4). Hence, the current analysis yields that a quantum state $\rho \in \mathcal{M}_{d^2}$ with a UU-symmetry must be of the form

$$\rho = a\frac{\mathbb{1}}{d^2} + b\mathbb{F}, \tag{3.18}$$

with appropriate $a, b \in \mathbb{C}$. The normalization condition, which says that $\mathrm{Tr}[\rho] = 1$, gives $b = (1-a)/d$. The positivity requirement, $\rho \geq 0$, then yields

$$\rho = a\frac{\mathbb{1}}{d^2} + (1-a)\frac{\mathbb{F}}{d}, \quad a \in \left[\frac{d}{d+1}, \frac{d}{d-1}\right], \tag{3.19}$$

where we have used the fact that the flip $\mathbb{F}$ has eigenvalues ± 1 (with multiplicities $d(d \pm 1)/2$) corresponding to symmetric and antisymmetric eigenvectors respectively.

Another very common representation of the Werner states is in terms of these eigenprojections of the flip $\mathbb{F}$ [20]. Let us denote these by $P_\pm$, i.e., $\mathbb{F}P_\pm = \pm P_\pm$, given by $P_\pm = (\mathbb{1} \pm \mathbb{F})/2$. As

already mentioned, these have multiplicities $d_\pm = d(d \pm 1)/2$ corresponding to the symmetric and antisymmetric subspace. Every Werner state can then be written in the following form

$$\rho = a\frac{P_+}{d_+} + (1-a)\frac{P_-}{d_-}, \quad a \in [0,1]\,. \tag{3.20}$$

The next example is closely related to the Werner state and the next section is therefore going to build on the analysis just presented.

3.1.2 $U\bar{U}$-invariant states – Isotropic states

A second very prominent example can be derived by considering the partial transposition of the second system of a Werner state. These are all states that are invariant under the compact group $\mathbf{G}$ represented by unitaries of the form $\left(U \otimes \bar{U}\right)$, with unitary $U \in \mathbf{U}(d)$, on $\mathbb{C}^d \otimes \mathbb{C}^d$. The $\bar{\cdot}$ again denotes complex conjugation. This is easily seen by considering lemma 3.6 and observing that if $\rho \in \mathcal{D}(\mathbb{C}^d \otimes \mathbb{C}^d)$ commutes with unitaries $U \otimes U$, then $\rho^{t_2} \in \mathcal{D}(\mathbb{C}^d \otimes \mathbb{C}^d)$ must commute with $U \otimes \bar{U}$ by using that $U^* = \bar{U}^t$ holds. A state obeying this symmetry is called an *isotropic state* [24]. The $U\bar{U}$-twirl projects every quantum state $\rho \in \mathcal{D}(\mathbb{C}^d \otimes \mathbb{C}^d)$ onto the isotropic states,

$$\mathbb{T}^{(U\bar{U})}(\rho) = \int_{\mathbf{U}(d)} \left(U \otimes \bar{U}\right) \rho \left(U \otimes \bar{U}\right)^* \mathrm{d}\mu. \tag{3.21}$$

The structure of these states can be determined by using the previous analysis regarding Werner states and our observation regarding the partial transposition [19]. First we should note how the partial transposition alters the action of the unitaries under consideration. We find for any $A_1, A_2 \in \mathcal{M}_d$ and every $U^{(1)}, U^{(2)} \in \mathbf{U}(d)$,

$$\begin{aligned}
&\left(\left(U^{(1)} \otimes U^{(2)}\right)(A_1 \otimes A_2)\left(U^{(1)} \otimes U^{(2)}\right)^*\right)^{t_2} \\
&= \left(U^{(1)} A_1 U^{(1)*}\right) \otimes \left(U^{(2)} A_2 U^{(2)*}\right)^t \\
&= \left(U^{(1)} A_1 U^{(1)*}\right) \otimes \left(\left(U^{(2)*}\right)^t A_2^t \left(U^{(2)}\right)^t\right) \\
&= \left(U^{(1)} \otimes \bar{U}^{(2)}\right)(A_1 \otimes A_2)^{t_2}\left(U^{(1)} \otimes \bar{U}^{(2)}\right)^*.
\end{aligned} \tag{3.22}$$

By linearity this holds for any $A \in \mathcal{M}_{d^2}$, too. Let $\mathbf{G}$ be a compact group with unitary representation of the form $\left(U^{(1)} \otimes U^{(2)}\right)$ and let $\mathbf{K}$ be a compact group with unitary representation of the form $\left(U^{(1)} \otimes \bar{U}^{(2)}\right)$ with unitaries $U^{(1)}, U^{(2)} \in \mathbf{U}(d)$. Furthermore, denote their respective twirls by $\mathbb{T}^{(U_1 U_2)}$ and $\mathbb{T}^{(U_1 \bar{U}_2)}$. Using the computation from equation (3.22), it is then immediately possible to see that the twirls are related through the partial transposition

$$\left(\mathbb{T}^{(U_1 U_2)}(\cdot)\right)^{t_2} = \mathbb{T}^{(U_1 \bar{U}_2)}\left((\cdot)^{t_2}\right). \tag{3.23}$$

As a consequence, we find that the commutants, as ranges of the twirls, obey the fundamental relation

$$\left(\mathbf{G}'\right)^{t_2} = \mathbf{K}'. \tag{3.24}$$

Therefore, the quantum states invariant under $\mathbf{K}$ are precisely the partial transposes of the quantum states invariant under $\mathbf{G}$.

Every isotropic state $\rho \in \mathcal{M}_{d^2}$ is thus of the form

$$\rho = a\frac{\mathbb{1}}{d^2} + b\,|\Omega\rangle\langle\Omega|\,, \tag{3.25}$$

for appropriate $a, b \in \mathbb{C}$, where $|\Omega\rangle\langle\Omega|$ denotes again the maximally entangled quantum state, which is the partial transposition of the flip, i.e.,

$$|\Omega\rangle\langle\Omega| = d\mathbb{F}^{t_2}. \tag{3.26}$$

Normalization yields that $b = 1 - a$ and together with positivity we get that every isotropic state is of the form

$$\rho = a\frac{\mathbb{1}}{d^2} + (1-a)\,|\Omega\rangle\langle\Omega|\,, \quad a \in \left[0, \frac{d^2}{d^2-1}\right]. \tag{3.27}$$

Combining the last two examples leads to a third prominent example of symmetry group, which is discussed in the next section.

3.1.3 OO-invariant states

The third example we would like to analyze is all quantum states $\rho \in \mathcal{D}\left(\mathbb{C}^d \otimes \mathbb{C}^d\right)$ that are invariant under the compact group $\mathbf{G}$ represented by real orthogonal matrices of the form $(O \otimes O)$, with orthogonal $O \in \mathbf{O}(d)$, on $\mathbb{C}^d \otimes \mathbb{C}^d$. These states are usually just referred to as OO-invariant quantum states [19]. The OO-twirl of a quantum state $\rho \in \mathcal{D}(\mathbb{C}^d \otimes \mathbb{C}^d)$ is given by

$$\mathbb{T}^{(OO)}(\rho) = \int_{\mathbf{O}(d)} (O \otimes O)\,\rho\,(O \otimes O)^*\ \mathrm{d}\mu, \tag{3.28}$$

where μ now denotes the Haar measure on $\mathbf{O}(d)$. The OO-twirl projects every quantum state onto the OO-invariant quantum states, i.e., into the commutant of $\mathbf{G}$.

The two previous examples fall into this class, since both UU-invariant quantum states as well as $U\bar{U}$-invariant quantum states are OO-invariant. Therefore, we know that the commutant $\mathbf{G}'$ is at least the algebra generated by the commutant of the group represented by unitaries of the form $(U \otimes U)$ and the group represented by unitaries of the form $\left(U \otimes \bar{U}\right)$, with $U \in \mathbf{U}(d)$. Considering the well-known theory of representation of the orthogonal group $\mathbf{O}(d)$, we know that there is no additional element to its commutant [8], i.e., every OO-invariant quantum state is of the form

$$\rho = a\,\mathbb{1} + b\mathbb{F} + c\,|\Omega\rangle\langle\Omega|\,, \tag{3.29}$$

with appropriate $a, b, c \in \mathbb{C}$. The corresponding minimal orthogonal projections in the flavor of equation (3.10) that span the commutant algebra are

$$P_0 = |\Omega\rangle\langle\Omega|\,, \qquad P_1 = \frac{1}{2}\left(\mathbb{1} - \mathbb{F}\right), \text{ and} \qquad P_2 = \frac{1}{2}\left(\mathbb{1} + \mathbb{F}\right) - |\Omega\rangle\langle\Omega|\,. \tag{3.30}$$

Therefore, every OO-invariant state can be written as a linear combination of these projections, i.e.,

$$\rho = (1 - a_1 - a_2)\,P_0 + a_1\frac{P_1}{\mathrm{Tr}[P_1]} + a_2\frac{P_2}{\mathrm{Tr}[P_2]}, \quad a_1, a_2 \geq 0, a_1 + a_2 \leq 1. \tag{3.31}$$

In this section, we discussed quantum states under symmetries and the decomposition of the underlying Hilbert space. We presented three examples that are very prominent in the literature, and we analyzed the structure of quantum states that are invariant with respect to these three symmetries.

The next section is going to analyze quantum channels under symmetries. We will start with a general discussion about covariant quantum channels and then discuss quantum channels under the symmetries particularly emphasized in this section.

3.2 Quantum channels under symmetry

This section uses previously presented results regarding quantum states under symmetry to discuss quantum channels under symmetry. We will, in particular, study the case of the orthogonal group related to the third example discussed in the last section. Before we do this, we start by defining covariance property of a quantum channel.

Definition 3.8 (Covariant quantum channel)**.** A quantum channel $T : \mathcal{M}_d \to \mathcal{M}_d$ is said to be *covariant* with respect to unitary representations $U^{(1)}, U^{(2)}$ of a finite or compact group $\mathbf{G}$, if for all $X \in \mathcal{M}_d$ and for all $g \in \mathbf{G}$,

$$T\left(U_g^{(2)^*} X U_g^{(2)}\right) = U_g^{(1)^*} T(X) U_g^{(1)}. \tag{3.32}$$

The twirl of a quantum channel is a very important concept and since it is related to covariance we define it in the following.

Definition 3.9 (Twirl of a quantum channel)**.** Let $T : \mathcal{M}_d \to \mathcal{M}_d$ be a quantum channel, $\mathbf{G}$ a finite or compact group with Haar measure μ and $U^{(1)}, U^{(2)}$ unitary representations of $\mathbf{G}$. We denote the *twirl* of T with respect to $\mathbf{G}$ as $\mathbb{T}_T : \mathcal{M}_d \to \mathcal{M}_d$ and define it to be

$$\mathbb{T}_T(\cdot) := \int_{\mathbf{G}} \mathcal{U}_g^{(1)} \circ T \circ \mathcal{U}_g^{(2)^*}(\cdot)\, \mathrm{d}\mu, \tag{3.33}$$

where $\mathcal{U}$ denotes the adjoint representation of $\mathbf{G}$ defined through its action on any $X \in \mathcal{M}_d$ by conjugation, i.e., for all $g \in \mathbf{G}$

$$\mathcal{U}_g(X) = U_g X U_g^*. \tag{3.34}$$

The following lemma shows that the notion of covariance and invariance under the twirl are closely linked. This is in the same spirit as lemma 3.6, as we will show later on.

Lemma 3.10. *Let $\mathbf{G}$ be a compact or finite group with unitary representations $U^{(1)}$ and $U^{(2)}$ and Haar measure μ. For a quantum channel $T : \mathcal{M}_d \to \mathcal{M}_d$*

$$U_g^{(1)} T\left(U_g^{(2)^*} \cdot U_g^{(2)}\right) U_g^{(1)^*} = T(\cdot) \quad \forall g \in \mathbf{G} \tag{3.35}$$

is equivalent to

$$\int_{\mathbf{G}} U_g^{(1)} T\left(U_g^{(2)^*} \cdot U_g^{(2)}\right) U_g^{(1)^*}\, \mathrm{d}\mu = T(\cdot). \tag{3.36}$$

Proof. If

$$U_g^{(1)} T \left(U_g^{(2)^*} \cdot U_g^{(2)} \right) U_g^{(1)^*} = T(\cdot)$$

holds for all $g \in \mathbf{G}$, then it immediately follows that

$$\int_{\mathbf{G}} U_g^{(1)} T \left(U_g^{(2)^*} \cdot U_g^{(2)} \right) U_g^{(1)^*} \, \mathrm{d}\mu = T(\cdot)$$

by the properties of the Haar measure. If, on the other hand,

$$\int_{\mathbf{G}} U_g^{(1)} T \left(U_g^{(2)^*} \cdot U_g^{(2)} \right) U_g^{(1)^*} \, \mathrm{d}\mu = T(\cdot),$$

then for fixed h,

$$\begin{aligned}
T \left(U_h^{(2)^*} \cdot U_h^{(2)} \right) &= \int_{\mathbf{G}} U_g^{(1)} T \left(U_g^{(2)^*} U_h^{(2)^*} \cdot U_h^{(2)} U_g^{(2)} \right) U_g^{(1)^*} \, \mathrm{d}\mu \\
&= \int_{\mathbf{G}} U_h^{(1)^*} U_h^{(1)} U_g^{(1)} T \left(U_g^{(2)^*} U_h^{(2)^*} \cdot U_h^{(2)} U_g^{(2)} \right) U_g^{(1)^*} U_h^{(1)^*} U_h^{(1)} \, \mathrm{d}\mu \\
&= U_h^{(1)^*} \left(\int_{\mathbf{G}} U_h^{(1)} U_g^{(1)} T \left(U_g^{(2)^*} U_h^{(2)^*} \cdot U_h^{(2)} U_g^{(2)} \right) U_g^{(1)^*} U_h^{(1)^*} \, \mathrm{d}\mu \right) U_h^{(1)} \\
&= U_h^{(1)^*} \left(\int_{\mathbf{G}} U_{hg}^{(1)} T \left(U_{hg}^{(2)^*} \cdot U_{hg}^{(2)} \right) U_{hg}^{(1)^*} \, \mathrm{d}\mu \right) U_h^{(1)} \\
&= U_h^{(1)^*} \left(\int_{\mathbf{G}} U_g^{(1)} T \left(U_g^{(2)^*} \cdot U_g^{(2)} \right) U_g^{(1)^*} \, \mathrm{d}\mu \right) U_h^{(1)} \\
&= U_h^{(1)^*} T(\cdot) U_h^{(1)},
\end{aligned}$$

where we have used the properties of the Haar measure and the fact that for h fixed, $g \mapsto hg$ is a bijection, so as g runs through $\mathbf{G}$, so does hg. □

We are in particular interested in bipartite quantum states $\rho \in \mathcal{D}\left(\mathbb{C}^d \otimes \mathbb{C}^d\right)$ and we again focus on closed groups $\mathbf{G}$ of unitaries of the form $U = \left(U^{(1)} \otimes U^{(2)}\right)$ with unitary $U^{(1)}, U^{(2)} \in \mathbf{U}(d)$. Similarly to the previous section, where we discussed quantum states under symmetry, we will also use a simplified notation in the rest of this section.

The Choi-Jamiolkowski state $J_T \in \mathcal{M}_{d^2}$ of a quantum channel $T : \mathcal{M}_d \to \mathcal{M}_d$ is invariant under twirling if it is a covariant quantum channel. This connection is formalized in the following lemma.

Lemma 3.11. *For a quantum channel $T : \mathcal{M}_d \to \mathcal{M}_d$*

$$U^{(1)} T \left(U^{(2)^*} \ \cdot \ U^{(2)} \right) U^{(1)^*} = T(\cdot) \quad \forall \, U^{(1)}, U^{(2)} \in \mathbf{U}(d) \tag{3.37}$$

is equivalent to

$$\int_{\mathbf{G}} \left(U^{(1)} \otimes \bar{U}^{(2)} \right) J_T \left(U^{(1)} \otimes \bar{U}^{(2)} \right)^* \mathrm{d}\mu = J_T, \tag{3.38}$$

where the group $\mathbf{G}$ is of the form $U^{(1)} \otimes \bar{U}^{(2)}$ with $U^{(1)}, U^{(2)} \in \mathbf{U}(d)$ and μ denotes the group's Haar measure.

Proof. If

$$U^{(1)}T\left(U^{(2)^*}\ \cdot\ U^{(2)}\right)U^{(1)^*} = T(\cdot)\ \ \forall\, U^{(1)}, U^{(2)} \in \mathbf{U}\,(d)$$

then

$$\begin{aligned}
&\int_{\mathbf{G}} \left(U^{(1)} \otimes \bar{U}^{(2)}\right) J_T \left(U^{(1)} \otimes \bar{U}^{(2)}\right)^* \mathrm{d}\mu \\
=&\int_{\mathbf{G}} \left(U^{(1)} \otimes \bar{U}^{(2)}\right) (T \otimes \mathrm{id}) \left|\Omega\right\rangle\!\left\langle\Omega\right| \left(U^{(1)} \otimes \bar{U}^{(2)}\right)^* \mathrm{d}\mu \\
=&\int_{\mathbf{G}} \left(U^{(1)} \otimes \bar{U}^{(2)}\right) (T \otimes \mathrm{id}) \left(U^{(2)} \otimes \bar{U}^{(2)}\right)^* \left|\Omega\right\rangle\!\left\langle\Omega\right| \left(U^{(2)} \otimes \bar{U}^{(2)}\right) \left(U^{(1)} \otimes \bar{U}^{(2)}\right)^* \mathrm{d}\mu \\
=&\int_{\mathbf{G}} \left(U^{(1)} T\left(U^{(2)^*} \cdot U^{(2)}\right) U^{(1)^*} \otimes \mathrm{id}\right) \left|\Omega\right\rangle\!\left\langle\Omega\right| \mathrm{d}\mu \\
=&\int_{\mathbf{G}} (T \otimes \mathrm{id}) \left|\Omega\right\rangle\!\left\langle\Omega\right| \mathrm{d}\mu \\
=&J_T,
\end{aligned}$$

where we have again used the properties of the Haar measure and the fact that $(U \otimes \bar{U})\left|\Omega\right\rangle = \left|\Omega\right\rangle$ for any unitary $U \in \mathbf{U}(d)$. The other direction follows from

$$\int_{\mathbf{G}} \left(U^{(1)} \otimes U^{(2)}\right) J_T \left(U^{(1)} \otimes U^{(2)}\right)^* \mathrm{d}\mu = J_T,$$

and application of lemma 3.10. □

This lemma 3.11 allows us to use the previous analysis from section 3.1 to analyze the structure of quantum channels under the three different symmetries studied before (see also [20]).

3.2.1 UU-covariant quantum channels

The first example that we discuss is the UU-covariant quantum channel $T : \mathcal{M}_d \to \mathcal{M}_d$ satisfying for any $X \in \mathcal{M}_d$

$$T\left(UXU^*\right) = UT(X)U^* \ \ \forall U \in \mathbf{U}(d). \tag{3.39}$$

By lemma 3.11 its corresponding Jamiolkowski state is an isotropic state. Therefore, consider the following linear maps $S_1, S_2 : \mathcal{M}_d \to \mathcal{M}_d$ that act on $X \in \mathcal{M}_d$ as

$$S_1(X) := d\,\mathrm{Tr}[X]\,\mathbb{1} \qquad \text{and} \qquad S_2(X) := X. \tag{3.40}$$

Their Choi matrices are the identity $\mathbb{1} \in \mathcal{M}_{d^2}$ and the maximally entangled state $\left|\Omega\right\rangle\!\left\langle\Omega\right| \in \mathcal{M}_{d^2}$

$$J_{S_1} = (S_1 \otimes \mathrm{id})\left|\Omega\right\rangle\!\left\langle\Omega\right| = \mathbb{1} \quad \text{and} \tag{3.41}$$

$$J_{S_2} = (S_2 \otimes \mathrm{id})\left|\Omega\right\rangle\!\left\langle\Omega\right| = \left|\Omega\right\rangle\!\left\langle\Omega\right|. \tag{3.42}$$

Using equation (3.27) therefore yields the well-known result that UU-covariant quantum channels are depolarizing, i.e., the UU-covariant quantum channel $T : \mathcal{M}_d \to \mathcal{M}_d$ has the structure

$$T(\cdot) = a\,\mathrm{Tr}[\cdot]\frac{\mathbb{1}}{d} + (1-a)\,\mathrm{id}, \quad a \in \left[0, \frac{d^2}{d^2-1}\right]. \tag{3.43}$$

This is a prominently used result throughout quantum information theory [20] and was derived here again using the invariance of its Jamiolkowski state.

3.2.2 $U\bar{U}$-covariant quantum channels

The second example is that of a quantum channel $T : \mathcal{M}_d \to \mathcal{M}_d$ that is $U\bar{U}$-covariant, i.e., that satisfies for every $X \in \mathcal{M}_d$

$$T\left(\bar{U}X\bar{U}^*\right) = UT(X)U^* \quad \forall U \in \mathbf{U}(d). \tag{3.44}$$

By lemma 3.11 its corresponding Jamiolkowski state is a Werner state. Therefore, in order to analyze its structure consider the following linear maps $S_1, S_3 : \mathcal{M}_d \to \mathcal{M}_d$ that act on any $X \in \mathcal{M}_d$ as

$$S_1(X) := d\,\mathrm{Tr}[X]\,\mathbb{1} \qquad \text{and} \qquad S_3(X) := dX^t. \tag{3.45}$$

Please note that these linear maps are not quantum channels, but they do have the correct invariance properties. Their Choi states are exactly the identity $\mathbb{1} \in \mathcal{M}_{d^2}$ and the flip $\mathbb{F} \in \mathcal{M}_{d^2}$,

$$J_{S_1} = (S_1 \otimes \mathrm{id})\,|\Omega\rangle\langle\Omega| = \mathbb{1} \quad \text{and} \tag{3.46}$$

$$J_{S_3} = (S_3 \otimes \mathrm{id})\,|\Omega\rangle\langle\Omega| = \mathbb{F}. \tag{3.47}$$

Using lemma 3.11 together with equation (3.20) shows that a $U\bar{U}$-covariant quantum channel $T : \mathcal{M}_d \to \mathcal{M}_d$ is of the form

$$T(\cdot) = \frac{a}{d+1}\left(\mathrm{Tr}[\cdot]\,\mathbb{1} + (\cdot)^t\right) + \frac{1-a}{d-1}\left(\mathrm{Tr}[\cdot]\,\mathbb{1} - (\cdot)^t\right), \quad a \in [0,1]. \tag{3.48}$$

This concludes the second example.

3.2.3 OO-covariant quantum channels

The third example combines the first two. Consider an OO-covariant quantum channel $T : \mathcal{M}_d \to \mathcal{M}_d$ satisfying for any $X \in \mathcal{M}_d$

$$T\left(OXO^*\right) = OT(X)O^* \quad \forall O \in \mathbf{O}(d). \tag{3.49}$$

By lemma 3.11 its corresponding Jamiolkowski state is an OO-invariant quantum state. In order to analyze its structure consider the following linear maps defined in the two previous examples, $S_1, S_2, S_3 : \mathcal{M}_d \to \mathcal{M}_d$ that act on any $X \in \mathcal{M}_d$ as

$$S_1(X) := d\,\mathrm{Tr}[X]\,\mathbb{1}, \qquad S_2(X) := X \qquad \text{and} \qquad S_2(X) := dX^t. \tag{3.50}$$

As discussed previously, their Choi matrices are exactly the identity $\mathbb{1} \in \mathcal{M}_{d^2}$, the maximally entangled state $|\Omega\rangle\langle\Omega| \in \mathcal{M}_{d^2}$ and the flip $\mathbb{F} \in \mathcal{M}_{d^2}$,

$$J_{S_1} = (S_1 \otimes \mathrm{id})\,|\Omega\rangle\langle\Omega| = \mathbb{1}, \tag{3.51}$$

$$J_{S_2} = (S_2 \otimes \mathrm{id})\,|\Omega\rangle\langle\Omega| = |\Omega\rangle\langle\Omega| \quad \text{and} \tag{3.52}$$

$$J_{S_3} = (S_3 \otimes \mathrm{id})\,|\Omega\rangle\langle\Omega| = \mathbb{F}. \tag{3.53}$$

Using equation (3.31) then shows that an OO-covariant quantum channel $T : \mathcal{M}_d \to \mathcal{M}_d$ has the structure

$$T(\cdot) = (1 - a_1 - a_2)\,\mathrm{id} + \frac{a_1}{d-1}\left(\mathrm{Tr}[\cdot]\,\mathbb{1} - (\cdot)^t\right) + \frac{2a_2}{d(d+1)-2}\left(\frac{d}{2}\left(\mathrm{Tr}[\cdot]\,\mathbb{1} - (\cdot)^t\right) - \mathrm{id}\right), \tag{3.54}$$

$a_1, a_2 \geq 0, a_1 + a_2 \leq 1$.

Therefore, the two previous examples are special cases of this quantum channel.

In this section we have discussed quantum states and quantum channels under symmetry. We analyzed three important symmetry examples and derived the explicit structure of the commutant. In the next section, we focus on group designs, in particular on unitary and orthogonal 2-designs. These are sets of unitary (or real orthogonal) matrices that are "nicely distributed" in the sense that the average of any 2nd degree polynomial over these sets equals the average over the entire unitary (or orthogonal) group [25, 26].

3.3 Group designs

Many algorithms and protocols in quantum information theory are based on random quantum states generated according to the Haar measure [27–32]. Generating these according to the Haar measure is, however, inefficient in practice, since the number of gates required grows exponentially with the number of qubits [33].Therefore, it is very useful to identify subsets of the unitary group that can adequately simulate the Haar measure for a class of operational tasks and to furthermore determine efficient gate decompositions for these subsets. This leads to the notion of a group design and more specifically unitary design, which will be the focus in this section. Unitary t-designs were, to our knowledge, first introduced to quantum information theory by Dankert in his thesis [34] and in a proceeding paper [33]. In the following we will focus on the special case of unitary 2-designs, where $t = 2$. All definitions can, however, be easily generalized.

Definition 3.12 (Unitary 2-design)**.** A set $\mathbf{D} = \{U_k \in \mathbf{U}(d)\}_{k=1,\dots,K}$ of unitary matrices on $\mathbb{C}^d$ is a *unitary* 2*-design* if it fulfills the following equivalent conditions:

(a) *(Averages)* Let $p(U)$ be a polynomial of degree at most 2 in the matrix elements of U and at most 2 in the complex conjugates of those matrix elements. Then the following relation should be fulfilled for any polynomial of this type,

$$\frac{1}{K}\sum_{k=1}^{K} p(U_k) = \int_{\mathbf{U}(d)} p(U)\,\mathrm{d}\mu. \tag{3.55}$$

(b) *(Twirling of quantum states)* For all bipartite quantum states $\rho \in \mathcal{D}(\mathbb{C}^d \otimes \mathbb{C}^d)$ the following holds

$$\frac{1}{K}\sum_{U_k \in \mathbf{D}} (U_k \otimes U_k)\rho(U_k \otimes U_k)^* = \int_{\mathbf{U}(d)} (U \otimes U)\rho(U \otimes U)^*\,\mathrm{d}\mu. \tag{3.56}$$

(c) *(Twirling of quantum channels)* For any quantum channel $T : \mathcal{M}_d \to \mathcal{M}_d$ and for any quantum state $\rho \in \mathcal{D}_d$ the following holds

$$\frac{1}{K}\sum_{U_k \in \mathbf{D}} U_k^* T\left(U_k \rho U_k^*\right) U_k = \int_{\mathbf{U}(d)} U^* T\left(U \rho U^*\right) U\,\mathrm{d}\mu. \tag{3.57}$$

In all these equivalent statements, the integral is again taken with respect to the Haar measure of $\mathbf{U}(d)$, denoted by μ.

The second equivalent formulation of a unitary 2-design, (b), reveals the link to the previous sections: We recover the notion of a Werner state. That is, for any unitary 2-design $\mathbf{D}$, we may use the analysis of the previous sections to specify the structure of its twirl or any quantum channel or quantum state under its twirl. In the following, we will focus on designs that form a group, since, to our best knowledge, all relevant examples fall into this class [35]. We can then immediately give an equivalent definition of a unitary t-design through the commutant:

Definition 3.13 (Unitary t-design)**.** Let $\mathbf{G}$ be a finite or compact group acting on $\mathbb{C}^d$ for some dimension d. Then $\mathbf{G}$ is a *unitary t-design* if we have the equality of commutants, i.e.,

$$\left\{U^{\otimes t} \,\middle|\, U \in \mathbf{G}\right\}' = \left\{U^{\otimes t} \,\middle|\, U \in \mathbf{U}(d)\right\}'. \tag{3.58}$$

This definition naturally lends itself to a generalization to arbitrary reference groups, beyond the well-studied case of $\mathbf{U}(d)$. In particular, the following will be of interest to us:

Definition 3.14 (Orthogonal t-design)**.** Let $\mathbf{G}$ be a finite or compact group acting on $\mathbb{C}^d$ for some dimension d. Then $\mathbf{G}$ is an *orthogonal t-design* if we have the equality of commutants, i.e.,

$$\left\{O^{\otimes t} \,\middle|\, O \in \mathbf{G}\right\}' = \left\{O^{\otimes t} \,\middle|\, O \in \mathbf{O}(d)\right\}', \tag{3.59}$$

where $\mathbf{O}(d)$ is the real orthogonal group acting on $\mathbb{C}^d$.

In the next sections, we will give two examples of designs. The first one is the *complex Clifford group*, which is a prominent example of a unitary 2-design in the quantum information theory literature, examples including randomized benchmarking [5, 36–38], quantum state tomography [39–42], quantum process tomography [43, 44], quantum cryptography [45] and quantum error correction [46]. The second example is the *real Clifford group*, which is an orthogonal 2-design [3]. It is not as popular as its complex counterpart in quantum information theory, but does find applications in randomized benchmarking [3] or quantum computing on rebits [47, 48].

3.3.1 Complex Clifford group

The complex Clifford group is a unitary 2-design, a unitary 3-design, but fails to be a unitary 4-design [49–54]. Therefore, let us start by defining the complex Clifford group.

The full Pauli group on one qubit $\mathbf{P}(1)$ is defined as the group generated by

$$\mathbf{P}(1) := \langle X, Z, iI \rangle, \tag{3.60}$$

where I, X, Y, Z are the standard Pauli matrices defined as

$$I := \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \quad X := \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad Y := \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} \text{ and } \quad Z := \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}. \tag{3.61}$$

The full Pauli group on n qubits is defined to be the n-fold tensor power

$$\mathbf{P}(n) := \mathbf{P}(1)^{\otimes n}. \tag{3.62}$$

Definition 3.15 (Complex Clifford group)**.** The complex Clifford group $\mathbf{X}(n)$ is the group-theoretic normalizer of the full Pauli group $\mathbf{P}(n)$ in the unitary group $\mathbf{U}(2^n)$, i.e., it is defined as

$$\mathbf{X}(n) := \{U \in \mathbf{U}(2^n) \,|\, U\mathbf{P}(n) = \mathbf{P}(n)U\}\,. \tag{3.63}$$

In the simplest case, for one qubit $n = 1$, the complex Clifford group is just given as

$$\mathbf{X}(1) = \langle H, P\rangle\,, \tag{3.64}$$

where H is the Hadamard gate defined as

$$H := \frac{1}{\sqrt{2}}\begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}, \tag{3.65}$$

and P is the $\pi/4$-phase gate defined as

$$P := \begin{pmatrix} 1 & 0 \\ 0 & i \end{pmatrix}. \tag{3.66}$$

In the simple two qubit case, $n = 2$, the complex Clifford group is just

$$\mathbf{X}(2) = \langle \mathbf{X}(1) \otimes \mathbf{X}(1), CZ\rangle\,, \tag{3.67}$$

where CZ is the controlled Z-gate defined as

$$CZ := \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix}. \tag{3.68}$$

The complex Clifford group is the most prominent non-trivial example of a unitary 2-design. A short proof of this fact is stated in the following. Please note that studying the commutant structure significantly simplifies the proof (cf. the proof in [35]).

Theorem 3.16. *The representation $\{U_g \otimes \bar{U}_g\}$ of the complex Clifford group $\mathbf{X}(n)$ decomposes into two non-degenerate irreducible unitary representations.*

Proof. See [55, theorem 6.8.1.] and proofs therein. □

Theorem 3.17. *The complex Clifford group $\mathbf{X}(n)$ is a unitary 2-design.*

Proof. The complex Clifford group $\mathbf{X}(n)$ is a subgroup of the unitary group $\mathbf{U}(2^n)$. Therefore, its commutant must contain the commutant of the unitary group. By theorem 3.16 both commutants have the same dimensions, and are thus equal. Application of definition 3.13 of a unitary 2-design proves the claim. □

Another example of interest is the real Clifford group. This is a subgroup of the complex Clifford group and is discussed in the next section.

3.3.2 Real Clifford group

In a similar fashion to the complex Clifford group, we can define a group called the *real Clifford group*. To do so, we first need to define the *real Pauli group* $\mathbf{E}(n)$ on n qubits as

$$\mathbf{E}(n) := \left\langle \mathbf{E}(1)^{\otimes n} \right\rangle, \tag{3.69}$$

where $\mathbf{E}(1)$ is just the real Pauli group on 1 qubit defined as

$$\mathbf{E}(1) := \left\langle X := \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, Z := \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \right\rangle. \tag{3.70}$$

$\mathbf{E}(n)$ is thus generated by tensor products of the Pauli matrices X and Z with 2×2 identity matrices $\mathbb{1}_2$. Therefore, as the matrices are real, it is called the real Pauli group.

Definition 3.18 (Real Clifford group)**.** The real Clifford group $\mathbf{C}(n)$ is the group-theoretic normalizer of the real Pauli group $\mathbf{E}(n)$ in the real orthogonal group $\mathbf{O}\left(2^n\right)$, i.e., it is defined as

$$\mathbf{C}(n) := \{O \in \mathbf{O}(2^n) \,|\, O\mathbf{E}(n) = \mathbf{E}(n)O\}. \tag{3.71}$$

In the simplest case when $n = 1$ the real Clifford group is generated by the Pauli Z-gate and the Hadamard gate H,

$$\mathbf{C}(1) = \left\langle Z := \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, H := \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix} \right\rangle, \tag{3.72}$$

and in the two qubit case when $n = 2$ the real Clifford group is

$$\mathbf{C}(2) = \langle \mathbf{C}(1) \otimes \mathbf{C}(1), CZ \rangle, \tag{3.73}$$

where CZ is again the controlled Z-gate, defined as

$$CZ := \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix}. \tag{3.74}$$

A thorough discussion of the real Clifford group can be found in [55].

The real Clifford group is therefore a subgroup of the complex Clifford group, $\mathbf{C}(n) \subset \mathbf{X}(n)$, and we necessarily have that for any $t \in \mathbb{N}$,

$$\{O^{\otimes t} | O \in \mathbf{C}(n)\}' \supset \{U^{\otimes t} | U \in \mathbf{X}(n)\}'.$$

Similarly, the orthogonal group is a subgroup of the unitary group, $\mathbf{O}(2^n) \subset \mathbf{U}(2^n)$, and we thus have that

$$\{O^{\otimes t} | O \in \mathbf{O}(2^n)\}' \supset \{U^{\otimes t} | U \in \mathbf{U}(2^n)\}'.$$

In the special case $t = 2$, we get the following correspondence:

$$\begin{array}{ccc} \{O^{\otimes 2} | O \in \mathbf{C}(n)\}' & \supset & \{U^{\otimes 2} | U \in \mathbf{X}(n)\}' \\ \| & & \| \\ \{O^{\otimes 2} | O \in \mathbf{O}(2^n)\}' & \supset & \{U^{\otimes 2} | U \in \mathbf{U}(2^n)\}' \end{array}$$

This is proven in the following.

Theorem 3.19. *The representation* $\{O_g \otimes O_g\}$ *of the real Clifford group* $\mathbf{C}(n)$ *decomposes into three non-degenerate irreducible unitary representations.*

Proof. See [55, theorem 6.8.1.] and proofs therein. □

Theorem 3.20. *The real Clifford group* $\mathbf{C}(n)$ *is an orthogonal* 2*-design.*

Proof. The real Clifford group $\mathbf{C}(n)$ is a subgroup of the real orthogonal group $\mathbf{O}(2^n)$. Therefore, the commutant of the real orthogonal group must be contained in the commutant of the real Clifford group. By theorem 3.19 both commutants have the same dimensions, and are thus equal. Application of definition 3.14 of an orthogonal 2-design proves the claim. □

We have discussed group designs and in particular unitary and orthogonal 2-designs. We have given one example each: the complex Clifford group is a unitary 2-design and one of its subgroups, the real Clifford group, is an orthogonal 2-design. These groups are of particular importance, since it is often the case that in practice it suffices to draw random matrices from the unitary (or orthogonal) design, instead of the full unitary (or orthogonal) group with respect to the Haar measure. We have studied the structure of quantum states and quantum channels with respect to unitary or orthogonal symmetries thoroughly. These findings then also apply to the groups design. These symmetry considerations have been applied throughout all articles included in this dissertation. We are in particular interested in the application of symmetries in the area of universal quantum cloning, information-disturbance tradeoffs as well as randomized benchmarking.

4 Applications

This chapter discusses three different applications of the symmetry methods presented in the previous chapter. The first application of interest is universal quantum cloning [1, 4], which is discussed in the next section. Another application of interest is within the area of information-disturbance tradeoffs [2, 4]. The third application of the symmetry methods is within randomized benchmarking [3, 5], which is presented in the third and final section of this chapter.

4.1 Universal quantum cloning

Quantum cloning is the process of perfectly cloning an arbitrary unknown quantum state and obtaining an exact copy of that quantum state without altering it, i.e., for any quantum state $\rho \in \mathcal{D}_d$ a quantum cloning device performs

$$\rho \mapsto \rho \otimes \rho. \tag{4.1}$$

Such a process is forbidden by the laws of quantum mechanics as was shown in the famous *No-Cloning Theorem* by Wootters and Zurek [16] and explained in theorem 2.5. Even though it is simple to prove by just invoking the linearity of quantum mechanics, this no-go theorem has fundamental consequences in the field of quantum information and yet again underlines the striking difference between quantum information and its classical counterpart. On a first look, it seems like the no-cloning theorem only tells us what cannot be done, it gives an intrinsic impossibility; however, it only tells us what cannot be done in an exact manner. On the other hand, most of the tasks cannot be performed in an exact manner using real devices. It is, therefore, more intriguing to study the bounds inherit to such a device. This finds applications to many tasks in quantum information theory, such as quantum state estimation [56–58] and eavesdropping in quantum cryptography [59–61]. The idea is that, in general, there is no way how an eavesdropper could perfectly copy a quantum state encoding some information. Consequently, as long as this quantum state is received unperturbed, the receiver knows that it has not been copied by any adversary. This is just one of the many useful application of the no-cloning theorem in information theory.

Historically, the no-cloning theorem was the result of many discussions about a proposal made by Nick Herbert [62]. He proposed a *FLASH communication scheme* (First Light Amplification Superluminal Hookup) to use quantum correlations to communicate faster than light. His proposal, however, needed that copying an unknown quantum state is possible. This triggered many discussions and responses [16, 63–65] and lead to the no-cloning theorem.

Even though perfect quantum cloning is not possible, it is possible to do so in an approximate manner. Quantum cloning, therefore, refers to the process of finding a quantum channel that

clones a quantum state as best as possible. If all quantum clones have the same quality, the cloning procedure is called symmetric. On the other hand, if the quantum clones are allowed to have different qualities, the cloning procedure is called asymmetric quantum cloning. Therefore, symmetric quantum cloning is a special case of asymmetric quantum cloning. In the case of universal quantum cloning, the cloning process is independent of the input quantum state, i.e., all input quantum states are copied equally well. Non-universal quantum cloning is thus mostly referred to as state-dependent cloning. In the following we will restrict to universal asymmetric quantum cloning, in which the clones may have different qualities, but this is independent of the input quantum state.

The first authors who went beyond the no-cloning theorem were Bužek and Hillery [66], who studied the optimal universal $1 \to 2$ qubit quantum cloning machine [67]. This gave rise to a vast amount of articles generalizing their universal symmetric quantum cloning machine to either qudits or the $N \to M$ cloning case [57, 68–73].

At the same time, the thorough exploration of universal asymmetric quantum cloning began [1, 71, 74–86], where the produced quantum clones are allowed to have different qualities. For more details, the reader is referred to the two excellent review articles [17, 18]. This special case is of particular interest to us, and it will thus be discussed in the following.

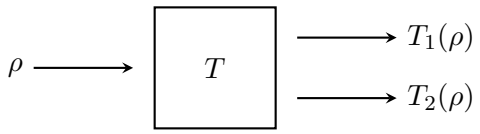


Figure 4.1: The main setup of universal asymmetric $1 \to 2$ quantum cloning. The quantum cloning channel $T : \mathcal{M}_d \to \mathcal{M}_d \otimes \mathcal{M}_d$ takes any initial state $\rho \in \mathcal{D}_d$ and gives two approximate clones $T_1(\rho) \in \mathcal{D}_d$ and $T_2(\rho) \in \mathcal{D}_d$ with possibly different qualities.

Let us consider the case of universal asymmetric $1 \to 2$ quantum cloning. Its main setup is depicted in figure 4.1. We are interested in finding the optimal cloning channel $T : \mathcal{M}_d \to \mathcal{M}_d \otimes \mathcal{M}_d$ that yields the best possible quantum clones, given through its marginals by $T_1(\rho) \in \mathcal{D}_d$ and $T_2(\rho) \in \mathcal{D}_d$. Intuitively, the better the quality of one of the clones, the worst the quality of the other must be. In order to assess the quality of the individual clones, we consider the distance of the marginal channel T_i to the identity channel. The goal is then to fully specify the set of all attainable single quantum clone qualities,

$$C = \left\{ z \in \mathbb{R}^2 \,\middle|\, z = \begin{pmatrix} d(T_1) \\ d(T_2) \end{pmatrix} \right\}, \tag{4.2}$$

where the distance measures $d : \mathcal{T}_d \to [0, \infty]$ must fulfill specific assumptions.

Assumption 4.1 (on the distance measure to the identity channel [1])**.** For $d : \mathcal{T}_d \to [0, \infty]$ we assume that

(a) d is concave,

(b) d is unitarily invariant, i.e., for every unitary $U \in \mathcal{M}_d$ and every quantum channel $T \in \mathcal{T}_d$, we have that

$$d\left(UT(U^* \cdot U)U^*\right) = d(T), \tag{4.3}$$

and

(c) the origin is attainable, i.e., $\{0\} \in C$.

It is then possible to show that without loss of generality (w.l.o.g.) the optimal quantum cloning channel is UU-covariant [1, 72, 73], such that the marginal maps take the form of a depolarizing channel, i.e., for $i = 1, 2$,

$$T_i(\cdot) = a_i^2 \operatorname{Tr}[\cdot]\frac{\mathbb{1}}{d} + (1 - a_i^2)\,\mathrm{id}, \quad a_i \in \left[0, \frac{d^2}{d^2-1}\right], \tag{4.4}$$

in accordance with section 3.2. Casting the universal asymmetric quantum cloning problem as an optimization problem, then allows to achieve analytic results for the optimal universal quantum cloning channel.

Theorem 4.2 (Optimal universal $1 \to 2$ asymmetric quantum cloning channel [1, theorem 3])**.** *The optimal universal* $1 \to 2$ *asymmetric quantum cloning channel* $T : \mathcal{M}_d \to \mathcal{M}_d \otimes \mathcal{M}_d$ *for any quantum state* $\rho \in \mathcal{D}_d$ *is given by*

$$T(\rho) = (a_2\,\mathbb{1} + a_1\mathbb{F})\left(\rho \otimes \frac{\mathbb{1}}{d}\right)(a_2\,\mathbb{1} + a_1\mathbb{F}), \tag{4.5}$$

with $(a_1)^2 + (a_2)^2 + \frac{2a_1a_2}{d} = 1$, $a_1, a_2 \in \mathbb{R}$, *and where* $\mathbb{F}$ *is the flip or swap operator.*

This application shows that symmetry considerations are very useful and it is possible to truncate the problem in question extensively. The symmetry considerably reduces the complexity of the problem while still preserving the interesting features of the full structure. This can also be observed in the next application that we will discuss: information-disturbance tradeoffs.

4.2 Information-disturbance tradeoff

The information-disturbance tradeoff quantifies how much information can be extracted through a quantum measurement from a quantum system and how much noise has to be necessarily introduced to that system. The goal is to mathematically describe the accessible region and find the optimal tradeoff. For a given amount of information, the disturbance is minimized, or, for a fixed amount of tolerable disturbance, the amount of information extracted from the system is maximized.

Heisenberg's uncertainty principle, mostly stated using the famous inequality for position and momentum,

$$\Delta q\,\Delta p \geq \hbar, \tag{4.6}$$

certainly comprises the information-disturbance tradeoff. Its mathematical development to include a generalized measurement scheme [14, 87], lead to numerous works on its assertion and scope. The inherent negative formulation regarding limitations of quantum preparations and measurement has lead to many underestimations of its fundamental relevance. The interest in a thorough mathematical analysis of the information-disturbance tradeoff has, however, seen

a significant increase, especially since the emergence of quantum information theory. This is especially the case, since it is relevant for some practical applications, such as quantum cryptography [59, 88–90].

This interest is represented by the vast amount of papers quantifying the tradeoff. One way, in which these papers differ, is whether a reference measurement is used to quantify the measurement error and whether a second reference system is used to quantify the disturbance. In [91–103], for example, both, the measurement error as well as the disturbance, are quantified with respect to a reference measurement. Contrarily, in [104–111] no reference measurement is used at all. We are particularly interested in an intermediate approach. The information gained is quantified through a quantum measurement, which is assumed to approximate a given target measurement, which serves as our reference. On the other hand, the disturbance is quantified without specifying any reference measurement. This differentiation is illustrated in figure B.1. This approach seems to be more natural, and also more applicable to quantum communication setups within the laboratory [4].

Another way, in which the papers quantifying the information-disturbance tradeoff differ, is the distance measures used. For example, [91, 100, 101, 105, 106, 109] use various entropic distance measures, [96, 103, 107] use norm-based distance measures, [104, 105, 109, 110] use fidelities, [95, 111] use Fisher information, and [97, 102] use transport-cost functions. These all stand on an equal footing and their respective advantages lie within their applications.

In the following we will consider the setup illustrated in figure 4.2. We fix a target measurement $E \in \mathcal{E}_{d,m}$ and we quantify the tradeoff between the quality of a measurement $E' \in \mathcal{E}_{d,m}$, that approximates the target measurement, and the disturbance that this measurement introduces to the system. The disturbance is described by a quantum channel $T \in \mathcal{T}_d$ and its distance to the identity channel $\mathrm{id} \in \mathcal{T}_d$. Every quantum instrument then describes one pair (E', T) allowed by quantum mechanics.

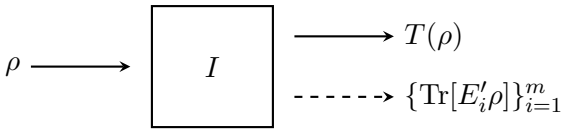


Figure 4.2: The main setup of the information-disturbance tradeoff. The quantum instrument $I = \{I_i : \mathcal{M}_d \to \mathcal{M}_d\}_{i=1}^m$ describes a measurement scheme. It gives the statistics of the measurement results corresponding to the POVM $E' \in \mathcal{E}_{d,m}$ (classical output) and the evolution of the initial quantum system $\rho \in \mathcal{D}_d$ to its final state after the measurement $T(\rho) \in \mathcal{D}_d$ (quantum output). The information-disturbance tradeoff quantifies the minimal disturbance to the system necessary to obtain a specific amount of information.

The disturbance is quantified by a functional $\Delta : \mathcal{T}_d \to [0, \infty]$ that determines how much $T \in \mathcal{T}_d$ differs from the identity channel $\mathrm{id} \in \mathcal{T}_d$. We will need the following assumption.

Assumption 4.3 (on the distance measure to the identity channel [2, assumption 1])**.**
For $\Delta : \mathcal{T}_d \to [0, \infty]$ we assume that

(a) $\Delta(\mathrm{id}) = 0$,

(b) Δ is convex,

(c) Δ is basis-independent, i.e., for every unitary $U \in \mathcal{M}_d$ and every quantum channel $T \in \mathcal{T}_d$, we have that

$$\Delta\left(UT(U^* \cdot U)U^*\right) = \Delta(T). \tag{4.7}$$

The measurement error is quantified by a functional $\delta : \mathcal{E}_{d,m} \to [0, \infty]$ that determines how much E' differs from the target measurement E. In the case of a non-degenerate von Neumann target measurement, which is the case we will focus on, we require the following assumptions.

Assumption 4.4 (on the distance measure to the target measurement [2, assumption 2])**.** For $\delta : \mathcal{E}_d \to [0, \infty]$ we assume that

(a) $\delta\big((|i\rangle\langle i|)_{i=1}^d\big) = 0$,

(b) δ is convex,

(c) δ is permutation-invariant, i.e., for every permutation $\pi \in S_d$ and any measurement $M \in \mathcal{E}_d$

$$M_i' = U_\pi^* M_{\pi(i)} U_\pi \ \forall i \ \Rightarrow \ \delta(M') = \delta(M), \tag{4.8}$$

where U_π is the permutation matrix that acts as $U_\pi|i\rangle = |\pi(i)\rangle$, and

(d) that for every diagonal unitary $D \in \mathcal{M}_d$ and any measurement $M \in \mathcal{E}_d$

$$M_i' = D^* M_i D \ \forall i \ \Rightarrow \ \delta(M') = \delta(M). \tag{4.9}$$

Using these assumptions on the two functionals, allows us to exploit the symmetry properties of the optimization problem. Every quantum instrument can be regarded as a quantum channel with a classical output and a quantum output. If we adopt this viewpoint, it is possible to show the following proposition.

Proposition 4.5 (Reduction to symmetric channels [2, proposition 1])**.** *Let* $\mathbf{G}$ *be the group generated by all diagonal unitaries and permutation matrices in* $\mathcal{M}_d$*. If* Δ *and* δ *satisfy assumptions 4.3 and 4.4, respectively, the optimal tradeoff between them can be attained within the set of channels* $\tilde{T} : \mathcal{M}_d \to \mathcal{M}_d \otimes \mathcal{M}_d$ *for which*

$$(U \otimes U)\tilde{T}\big(U^* \rho U\big)(U \otimes U)^* = \tilde{T}(\rho) \quad \forall U \in \mathbf{G}, \rho \in \mathcal{D}_d. \tag{4.10}$$

The optimization is thus invariant under the group of all diagonal unitaries and permutation matrices. The results derived in section 3.2.1 can therefore only be used partly, because we do not consider the full unitary group here. Given that the group $\mathbf{G}$, generated by all diagonal unitaries and permutation matrices, is a subgroup of the full unitary group, we expect its commutant to be larger. The structure of the marginals has an additional term, which is given in the next lemma.

Lemma 4.6 (Structure of marginals of symmetric channels [2, lemma 1])**.** *Let* $\mathbf{G}$ *be the group generated by all diagonal unitaries and permutation matrices in* $\mathcal{M}_d$ *and* $T : \mathcal{M}_d \to \mathcal{M}_d$ *a quantum channel. Then the following are equivalent:*

(a) $T(\rho) = UT\left(U^*\rho U\right)U^* \quad \forall U \in \mathbf{G}, \rho \in \mathcal{D}_d$.

(b) There are $a, b, c \in \mathbb{R}$ *with* $a + b + c = 1$ *so that*

$$T(\cdot) = a\,\mathrm{Tr}[\cdot]\frac{\mathbb{1}}{d} + b\,\mathrm{id} + c\sum_{i=1}^{d} |i\rangle\langle i|\,\langle i|\cdot|i\rangle\,. \tag{4.11}$$

These symmetry considerations, therefore, once again significantly reduce the complexity of the problem. It is therefore possible to obtain analytic results regarding the information-disturbance tradeoff as is shown in [2]. We can show that if the target measurement is a non-degenerate von Neumann measurement, then the optimal information-disturbance tradeoff can always be achieved within a two-parameter family of quantum instruments, which is independent of the chosen distance measure.

Theorem 4.7 ((Almost universal) optimal instruments [2, theorem 1])**.** *Let* Δ *and* δ *be distance-measures for quantifying disturbance and measurement-error that satisfy assumptions 4.3 and 4.4, respectively. Then the optimal* $\Delta - \delta$*-tradeoff w.r.t. a target measurement that is given by an orthonormal basis* $\{|i\rangle \in \mathbb{C}^d\}_{i=1}^d$ *is attained within the two-parameter family of instruments defined by*

$$I_i(\rho) := z\,\langle i|\,\rho\,|i\rangle\,\frac{\mathbb{1}_d - |i\rangle\langle i|}{d-1} + (1-z)K_i\rho K_i, \quad K_i := \mu\,\mathbb{1}_d + \nu\,|i\rangle\langle i|\,, \tag{4.12}$$

where $z \in [0,1]$ *and* $\mu, \nu \in \mathbb{R}$ *satisfy* $d\mu^2 + \nu^2 + 2\mu\nu = 1$ *(which makes* $\sum_i I_i$ *trace preserving).*

The next section discusses one more application of these symmetry methods, namely randomized benchmarking.

4.3 Randomized benchmarking

Randomized benchmarking is a method for obtaining quantitative estimates of the average error rate of a physical quantum channel. This is a crucial ingredient to many applications, in which only a sufficiently low error guarantees a successful implementation, such as quantum computing. Full characterization of the error of the physical quantum channel is possible through quantum process tomography [7]. This method is, however, not feasible in practice. It requires a number of experimental configurations that grows exponentially with the system size. Furthermore, it unrealistically assumes that the measurements and the state preparation admit lower errors than the process itself. Randomized benchmarking overcomes these difficulties and allows us to quantitatively determine the noise in a system efficiently, even though only partial information is obtained. Randomized benchmarking has become a popular tool to assess the quality of quantum processes [36, 112–120].

The standard approach to randomized benchmarking is to consider quantum gates taken from the complex Clifford group [31, 33, 37, 121–128]. However, randomized benchmarking is possible using any other finite or compact group $\mathbf{G}$ as shown in [3, 5, 123, 124, 129]. Sequences of quantum gates are then generated, such that the net sequence, if realized without errors, is

the identity operation. An average sequence fidelity can then be obtained by averaging over many random realizations of quantum gate sequences of the same length. Repeating these steps many times for different sequence lengths, we obtain an estimate of the expected value of the sequence fidelity to which experimental data may be calibrated. This yields information about the average error rate of the noise of the system.

More concretely, consider a sequence of $m+1$ quantum gates $\mathcal{C}_j : \mathcal{M}_{2^n} \to \mathcal{M}_{2^n}$, $j = 1, \ldots, m$, taken from a finite or compact group $\mathbf{G}$ that act on an initial quantum state $\rho \in \mathcal{D}_{2^n}$, followed by a measurement represented by a POVM $E \in \mathcal{E}_{2^n}$. This is visualized in figure 4.3. The physical quantum channels are

$$\tilde{\mathcal{C}}_j = \mathcal{C}_j \circ \mathcal{E}, \tag{4.13}$$

where $\mathcal{C}_j(\cdot) = C_j \cdot C_j$, $C_j \in \mathbf{G}$, is a quantum gate taken from the representation of the group $\mathbf{G}$ and $\mathcal{E} : \mathcal{M}_{2^n} \to \mathcal{M}_{2^n}$ is the associated error quantum channel. It is common to assume that the error quantum channel is both gate and time independent. A generalization is straightforward, but introduces unnecessary complicated notation. Therefore, we will also assume gate and time independence of the error quantum channel here.

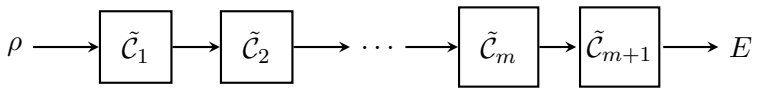


Figure 4.3: The main setup of randomized benchmarking. For a fixed $m \in \mathbb{N}$, a sequence of $m+1$ quantum gates is applied to an initial quantum state $\rho \in \mathcal{D}_{2^n}$. A measurement $E \in \mathcal{E}_{2^n}$ is then performed.

The sequence is generated, such that in the case of its ideal implementation, it gives the identity operation, i.e.,

$$\mathcal{C}_{m+1} \circ \mathcal{C}_m \circ \ldots \circ \mathcal{C}_2 \circ \mathcal{C}_1 = \mathrm{id}\,. \tag{4.14}$$

A subsequent measurement is performed given by an effect operator of a POVM, $E \in \mathcal{E}_{2^n}$, to measure the survival probability. Averaging over $M \in \mathbb{N}$ random realizations of sequences of length m gives the average sequence fidelity. This averaging procedure gives a twirl over that respective group. In the case of the complex Clifford group $\mathbf{X}(n)$, which we know forms a unitary 2-design by theorem 3.17, we can use the symmetry considerations from section 3.2.1 to deduce that the average over the complex Clifford group gives a depolarizing error quantum channel, i.e.,

$$\mathcal{E}(\cdot) = a\,\mathrm{Tr}[\cdot]\frac{\mathbb{1}}{d} + (1-a)\,\mathrm{id}, \quad a \in \left[0, \frac{d^2}{d^2-1}\right]. \tag{4.15}$$

In the case of the real Clifford group $\mathbf{C}(n)$, which is an orthogonal 2-design by theorem 3.20, we know from section 3.2.3 that the average over the real Clifford group yields an error quantum channel of the form

$$\mathcal{E}(\cdot) = (1-a_1-a_2)\,\mathrm{id} + \frac{a_1}{d-1}\left(\mathrm{Tr}[\cdot]\,\mathbb{1} - (\cdot)^t\right) + \frac{2a_2}{d(d+1)-2}\left(\frac{d}{2}\left(\mathrm{Tr}[\cdot]\,\mathbb{1} - (\cdot)^t\right) - \mathrm{id}\right), \tag{4.16}$$

$a_1, a_2 \geq 0, a_1 + a_2 \leq 1$.

The symmetry considerations again significantly simplify the problem under consideration. It is then possible to derive the exponential decay of the expected fidelity. The parameters to which experimental data is then calibrated are exactly those parameters appearing in equations (4.15) and (4.16) or the alike parameters if another symmetry group is considered. These then assess the quality of the quantum process.

The three applications discussed in the three previous sections clearly show that symmetry methods provide a powerful toolbox to study fundamental features of quantum mechanics and they demonstrate their exploitation for information processing tasks. Symmetry methods within quantum information theory allow to significantly reduce the complexity of the problem under investigation, so that it is possible to obtain analytic results.

Bibliography

[1] A. K. Hashagen. Universal asymmetric quantum cloning revisited. *Quant. Inf. Comp.*, 17(9& 10):0747–0778, August 2017.

[2] A. K. Hashagen and M. M. Wolf. Universality and optimality in the information-disturbance tradeoff. *Accepted in Ann. Henri Poincaré*, 2018.

[3] A. K. Hashagen, S. T. Flammia, D. Gross, and J. J. Wallman. Real randomized benchmarking. *Quantum*, 2:85, August 2018.

[4] L. Knips, J. Dziewior, A. K. Hashagen, J. D. A. Meinecke, H. Weinfurter, and M. M. Wolf. Measurement-disturbance tradeoff outperforming optimal cloning. *ArXiv e-prints: arXiv:1808.07882 [quant-ph]*, August 2018.

[5] D. S. França and A. K. Hashagen. Approximate randomized benchmarking for finite groups. *J. Phys. A: Math. Theor.*, 51(39):395302, August 2018.

[6] M. A. Nielsen and I. L. Chuang. *Quantum Computation and Quantum Information*. Cambridge Series on Information and the Natural Sciences. Cambridge University Press, 1st edition, January 2004.

[7] T. Heinosaari and M. Ziman. *The Mathematical Language of Quantum Theory: From Uncertainty to Entanglement*. Cambridge University Press, 2012.

[8] B. Simon. *Representations of finite and compact groups*, volume 10 of *Graduate studies in mathematics*. American Mathematical Society, 1996.

[9] P. A. M. Dirac. A new notation for quantum mechanics. *Math. Proc. Cambridge Philos. Soc.*, 35(3):416–418, 1939.

[10] J. M. Renes, R. Blume-Kohout, A. J. Scott, and C. M. Caves. Symmetric informationally complete quantum measurements. *J. Math. Phys.*, 45(6):2171–2180, 2004.

[11] C. A. Fuchs, M. C. Hoang, and B. C. Stacey. The SIC question: History and state of play. *Axioms*, 6(3):21, 2017.

[12] M.-D. Choi. Completely positive linear maps on complex matrices. *Linear Algebra Appl.*, 10(3):285–290, 1975.

[13] K. Kraus. *States, Effects, and Operations*, volume 190 of *Lecture Notes in Physics*. Springer-Verlag Berlin Heidelberg, 1st edition, 1983.

[14] E. B. Davies and J. T. Lewis. An operational approach to quantum probability. *Comm. Math. Phys.*, 17(3):239–260, 1970.

[15] J. Watrous. *The Theory of Quantum Information*. Cambridge University Press, 2018.

[16] W. K. Wootters and W. H. Zurek. A single quantum cannot be cloned. *Nature*, 299:802–803, October 1982.

[17] N. J. Cerf and J. Fiurášek. Optical quantum cloning. In E. Wolf, editor, *Progress in Optics*, volume 49, pages 455–545. Elsevier, 2006.

[18] V. Scarani, S. Iblisdir, N. Gisin, and A. Acín. Quantum cloning. *Rev. Mod. Phys.*, 77:1225–1256, November 2005.

[19] K. G. H. Vollbrecht and R. F. Werner. Entanglement measures under symmetry. *Phys. Rev. A*, 64:062307, November 2001.

[20] M. Keyl. Fundamentals of quantum information theory. *Phys. Rep.*, 369(5):431–548, 2002.

[21] R. F. Werner. Quantum states with Einstein-Podolsky-Rosen correlations admitting a hidden-variable model. *Phys. Rev. A*, 40:4277–4281, October 1989.

[22] S. Popescu. Bell's inequalities versus teleportation: What is nonlocality? *Phys. Rev. Lett.*, 72:797–799, February 1994.

[23] H. Weyl. *The Classical Groups: Their Invariants and Representations*. Princeton landmarks in mathematics and physics. Princeton University Press, 1997.

[24] E. M. Rains. Bound on distillable entanglement. *Phys. Rev. A*, 60:179–184, July 1999.

[25] E. Bannai. On some spherical t-designs. *J. Combin. Theory Ser. A*, 26(2):157–161, 1979.

[26] E. Bannai and E. Bannai. A survey on spherical designs and algebraic combinatorics on spheres. *European J. Combin.*, 30(6):1392–1425, 2009. Association Schemes: Ideas and Perspectives.

[27] J. Radhakrishnan, M. Rötteler, and P. Sen. Random measurement bases, quantum state distinction and applications to the hidden subgroup problem. *Algorithmica*, 55(3):490–516, 2005.

[28] P. Hayden, D. Leung, P. W. Shor, and A. Winter. Randomizing quantum states: Constructions and applications. *Comm. Math. Phys.*, 250(2):371–391, September 2004.

[29] C. H. Bennett, P. Hayden, D. W. Leung, P. W. Shor, and A. Winter. Remote preparation of quantum states. *IEEE Trans. Inf. Theory*, 51(1):56–74, January 2005.

[30] A. Harrow, P. Hayden, and D. Leung. Superdense coding of quantum states. *Phys. Rev. Lett.*, 92:187901, May 2004.

[31] J. Emerson, R. Alicki, and K. Zyczkowski. Scalable noise estimation with random unitary operators. *J. Opt. B*, 7(10):S347, 2005.

[32] A. Abeyesinghe, I. Devetak, P. Hayden, and A. Winter. The mother of all protocols: restructuring quantum information's family tree. *Proc. R. Soc. A*, 465:2537–2563, June 2009.

[33] C. Dankert, R. Cleve, J. Emerson, and E. Livine. Exact and approximate unitary 2-designs and their application to fidelity estimation. *Phys. Rev. A*, 80:012304, July 2009.

[34] C. Dankert. Efficient simulation of random quantum states and operators. Master's thesis, University of Waterloo, 2005.

[35] D. Gross, K. Audenaert, and J. Eisert. Evenly distributed unitaries: On the structure of unitary designs. *J. Math. Phys.*, 48(5):052104, 2007.

[36] E. Knill, D. Leibfried, R. Reichle, J. Britton, R. B. Blakestad, J. D. Jost, C. Langer, R. Ozeri, S. Seidelin, and D. J. Wineland. Randomized benchmarking of quantum gates. *Phys. Rev. A*, 77:012307, January 2008.

[37] E. Magesan, J. M. Gambetta, and J. Emerson. Robust randomized benchmarking of quantum processes. *Phys. Rev. Lett.*, 106:180504, 2011.

[38] J. J. Wallman and S. T. Flammia. Randomized benchmarking with confidence. *New J. Phys.*, 16(10):103032, 2014.

[39] A. Hayashi, T. Hashimoto, and M. Horibe. Reexamination of optimal quantum state estimation of pure states. *Phys. Rev. A*, 72:032325, September 2005.

[40] A. J. Scott. Tight informationally complete quantum measurements. *J. Phys. A*, 39(43):13507, 2006.

[41] H. Zhu. Quantum state estimation with informationally overcomplete measurements. *Phys. Rev. A*, 90:012115, July 2014.

[42] H. Zhu and B. Englert. Quantum state tomography with fully symmetric measurements and product measurements. *Phys. Rev. A*, 84:022327, August 2011.

[43] Y.-K. Liu and S. Kimmel. Quantum compressed sensing using 2-designs. In *APS Meeting Abstracts*, page B44.006, 2016.

[44] A. J. Scott. Optimizing quantum process tomography with unitary 2-designs. *J. Phys. A*, 41(5):055308, 2008.

[45] A. Ambainis, J. Bouda, and A. Winter. Nonmalleable encryption of quantum information. *J. Math. Phys.*, 50(4):042106, 2009.

[46] D. Gottesman. *Stabilizer codes and quantum error correction.* PhD thesis, California Institute of Technology, 1997.

[47] T. Rudolph and L. Grover. A 2 rebit gate universal for quantum computing. *ArXiv e-prints: arXiv:quant-ph/0210187*, October 2002.

[48] A. R. Calderbank, E. M. Rains, P. W. Shor, and N. J. A. Sloane. Quantum error correction and orthogonal geometry. *Phys. Rev. Lett.*, 78:405–408, January 1997.

[49] Z. Webb. The Clifford group forms a unitary 3-design. *Quant. Inf. Comp.*, 16:1379–1400, 2016.

[50] H. Zhu. Multiqubit Clifford groups are unitary 3-designs. *Phys. Rev. A*, 96(6):062336, 2017.

[51] R. Kueng and D. Gross. Qubit stabilizer states are complex projective 3-designs. *ArXiv e-prints: arXiv:1510.02767 [quant-ph]*, 2015.

[52] J. Helsen, J. J. Wallman, and S. Wehner. Representations of the multi-qubit Clifford group. *J. Math. Phys.*, 59(7):072201, 2018.

[53] H. Zhu, R. Kueng, M. Grassl, and D. Gross. The Clifford group fails gracefully to be a unitary 4-design. *ArXiv e-prints: arXiv:1609.08172 [quant-ph]*, September 2016.

[54] D. Gross, S. Nezami, and M. Walter. Schur-Weyl duality for the Clifford group with applications: Property testing, a robust Hudson theorem, and de Finetti representations. *arXiv e-print: arXiv:1712.08628 [quant-ph]*, 2017.

[55] G. Nebe, E. M. Rains, and N. J. A. Sloane. *Self-Dual Codes and Invariant Theory.* Algorithms and Computation in Mathematics. Springer Berlin Heidelberg, 2006.

[56] G. Chiribella. On quantum estimation, quantum cloning and finite quantum de Finetti theorems. *Lect. Notes Comput. Sci.*, 6519:9, 2011.

[57] D. Bruß, A. Ekert, and C. Macchiavello. Optimal universal quantum cloning and state estimation. *Phys. Rev. Lett.*, 81:2598–2601, September 1998.

[58] J. Bae and A. Acín. Asymptotic quantum cloning is state estimation. *Phys. Rev. Lett.*, 97(3):030402, July 2006.

[59] C. H. Bennet and G. Brassard. Quantum cryptography: Public key distribution and coin tossing. In *Proceedings of IEEE International Conference on Computers, Systems and Signal Processing*, pages 175–179, December 1984.

[60] N. J. Cerf, M. Bourennane, A. Karlsson, and N. Gisin. Security of quantum key distribution using d-level systems. *Phys. Rev. Lett.*, 88:127902, March 2002.

[61] L. Gyongyosi and S. Imre. Fidelity analysis of quantum cloning based attacks in quantum cryptography. In *2009 10th International Conference on Telecommunications*, pages 221–227, June 2009.

[62] N. Herbert. FLASH—a superluminal communicator based upon a new kind of quantum measurement. *Found. Phys.*, 12(12):1171–1179, Dec 1982.

[63] D. Dieks. Communication by EPR devices. *Phys. Lett. A*, 92(6):271–272, 1982.

[64] P. W. Milonni and M. L. Hardies. Photons cannot always be replicated. *Phys. Lett. A*, 92(7):321–322, 1982.

[65] L. Mandel. Is a photon amplifier always polarization dependent? *Nature*, 304:188, July 1983.

[66] V. Bužek and M. Hillery. Quantum copying: Beyond the no-cloning theorem. *Phys. Rev. A*, 54(3):1844, 1996.

[67] D. Bruß, D. P. DiVincenzo, A. Ekert, C. A. Fuchs, C. Macchiavello, and J. A. Smolin. Optimal universal and state-dependent quantum cloning. *Phys. Rev. A*, 57:2368–2378, April 1998.

[68] N. Gisin and S. Massar. Optimal quantum cloning machines. *Phys. Rev. Lett.*, 79:2153–2156, September 1997.

[69] M. Hillery and V. Bužek. Quantum copying: Fundamental inequalities. *Phys. Rev. A*, 56:1212–1216, August 1997.

[70] V. Bužek and M. Hillery. Universal optimal cloning of arbitrary quantum states: From qubits to quantum registers. *Phys. Rev. Lett.*, 81:5003–5006, November 1998.

[71] N. J. Cerf. Asymmetric quantum cloning machines. *Acta Phys. Slovaca*, 48(3):115–132, June 1998.

[72] R. F. Werner. Optimal cloning of pure states. *Phys. Rev. A*, 58:1827–1832, 1998.

[73] M. Keyl and R. F. Werner. Optimal cloning of pure states, testing single clones. *J. Math. Phys.*, 40(7):3283–3299, July 1999.

[74] V. Bužek, M. Hillery, and R. Bednik. Controlling the flow of information in quantum cloners: Asymmetric cloning. *Acta Phys. Slovaca*, 48:177, 1998.

[75] C. Niu and R. B. Griffiths. Optimal copying of one quantum bit. *Phys. Rev. A*, 58:4377–4393, December 1998.

[76] N. J. Cerf. Pauli cloning of a quantum bit. *Phys. Rev. Lett.*, 84(19):4497–4500, 2000.

[77] S. L. Braunstein, V. Bužek, and M. Hillery. Quantum-information distributors: Quantum network for symmetric and asymmetric cloning in arbitrary dimension and continuous limit. *Phys. Rev. A*, 63:052313, April 2001.

[78] N. J. Cerf. Asymmetric quantum cloning in any dimension. *J. Mod. Opt.*, 47(2–3):187–209, 2000.

[79] S. Iblisdir, A. Acín, N. J. Cerf, R. Filip, J. Fiurášek, and N. Gisin. Multipartite asymmetric quantum cloning. *Phys. Rev. A*, 72:042328, October 2005.

[80] S. Iblisdir, A. Acín, and N. Gisin. Generalised asymmetric quantum cloning machines. *Quant. Inform. Comp.*, 6(4):410–435, July 2006.

[81] J. Fiurášek, R. Filip, and N. J. Cerf. Highly asymmetric quantum cloning in arbitrary dimension. *Quant. Inf. Comp.*, 5:583–592, 2005.

[82] P. Ćwikliński, M. Horodecki, and M. Studziński. Region of fidelities for a $1 \to N$ universal qubit quantum cloner. *Phys. Lett. A*, 376(32):2178–2187, 2012.

[83] M. Studziński, P. Ćwikliński, M. Horodecki, and M. Mozrzymas. Group-representation approach to $1 \to N$ universal quantum cloning machines. *Phys. Rev. A*, 89:052322, May 2014.

[84] A. Kay, D. Kaszlikowski, and R. Ramanathan. Optimal cloning and singlet monogamy. *Phys. Rev. Lett.*, 103:050501, July 2009.

[85] A. Kay, R. Ramanathan, and D. Kaszlikowshi. Optimal asymmetric quantum cloning for quantum information and computation. *Quant. Inf. Comp.*, 13(9-10):880–900, September 2013.

[86] A. Kay. Optimal universal quantum cloning: Asymmetries and fidelity measures. *Quant. Inf. Comp.*, 16:991, 2016.

[87] G. Lüders. Über die Zustandsänderung durch den Meßprozeß. *Ann. Phys.*, 443(5-8):322–328, 1950.

[88] A. K. Ekert. Quantum cryptography based on Bell's theorem. *Phys. Rev. Lett.*, 67:661–663, August 1991.

[89] C. A. Fuchs and A. Peres. Quantum-state disturbance versus information gain: Uncertainty relations for quantum information. *Phys. Rev. A*, 53:2038–2045, April 1996.

[90] C. A. Fuchs. *Information Gain vs. State Disturbance in Quantum Theory*, chapter 13, pages 229–259. Wiley-VCH Verlag GmbH & Co. KGaA, January 2005.

[91] H. Martens and W. M. de Muynck. Disturbance, conservation laws and the uncertainty principle. *J. Phys. A*, 25(18):4887, 1992.

[92] M. Ozawa. Universally valid reformulation of the Heisenberg uncertainty principle on noise and disturbance in measurement. *Phys. Rev. A*, 67(4):042105, April 2003.

[93] M. Ozawa. Uncertainty relations for noise and disturbance in generalized quantum measurements. *Ann. Phys.*, 311(2):350–416, 2004.

[94] T. Heinosaari and M. M. Wolf. Nondisturbing quantum measurements. *J. Math. Phys.*, 51(9):092201, 2010.

[95] Y. Watanabe and M. Ueda. Quantum estimation theory of error and disturbance in quantum measurement. *ArXiv e-prints: arXiv:1106.2526 [quant-ph]*, June 2011.

[96] A. C. Ipsen. Error-disturbance relations for finite dimensional systems. *ArXiv e-prints: arXiv:1311.0259 [quant-ph]*, November 2013.

[97] P. Busch, P. Lahti, and R. F. Werner. Proof of Heisenberg's error-disturbance relation. *Phys. Rev. Lett.*, 111(16):160405, October 2013.

[98] P. Busch, P. Lahti, and R. F. Werner. Colloquium: Quantum root-mean-square error and measurement uncertainty relations. *Rev. Mod. Phys.*, 86:1261–1281, October 2014.

[99] C. Branciard. How well can one jointly measure two incompatible observables on a given quantum state? *Proc. Natl. Acad. Sci. USA*, 110(17):6742–6747, April 2013.

[100] F. Buscemi, M. J. W. Hall, M. Ozawa, and M. M. Wilde. Noise and disturbance in quantum measurements: An information-theoretic approach. *Phys. Rev. Lett.*, 112(5):050401, February 2014.

[101] P. J. Coles and F. Furrer. State-dependent approach to entropic measurement-disturbance relations. *Phys. Lett. A*, 379:105–112, January 2015.

[102] R. Schwonnek, D. Reeb, and R. F. Werner. Measurement uncertainty for finite quantum observables. *Mathematics*, 4(2):38, June 2016.

[103] J. M. Renes, V. B. Scholz, and S. Huber. Uncertainty relations: An operational approach to the error-disturbance tradeoff. *Quantum*, 1:20, July 2017.

[104] K. Banaszek. Fidelity balance in quantum operations. *Phys. Rev. Lett.*, 86:1366–1369, February 2001.

[105] H. Barnum. Information-disturbance tradeoff in quantum measurement on the uniform ensemble. In *Proceedings of IEEE International Symposium on Information Theory*, page 277, 2001.

[106] L. Maccone. Entropic information-disturbance tradeoff. *Europhys. Lett.*, 77(4):40002, 2007.

[107] D. Kretschmann, D. Schlingemann, and R. F. Werner. The information-disturbance tradeoff and the continuity of Stinespring's representation. *IEEE Trans. Inf. Theory*, 54(4):1708–1717, April 2008.

[108] F. Buscemi, M. Hayashi, and M. Horodecki. Global information balance in quantum measurements. *Phys. Rev. Lett.*, 100:210504, May 2008.

[109] F. Buscemi and M. Horodecki. Towards a unified approach to information-disturbance tradeoffs in quantum measurements. *Open Syst. Inf. Dyn.*, 16(01):29–48, 2009.

[110] A. Bisio, G. Chiribella, G. M. D'Ariano, and P. Perinotti. Information-disturbance tradeoff in estimating a unitary transformation. *Phys. Rev. A*, 82:062305, December 2010.

[111] T. Shitara, Y. Kuramochi, and M. Ueda. Trade-off relation between information and disturbance in quantum measurement. *Phys. Rev. A*, 93:032134, March 2016.

[112] J. M. Chow, J. M. Gambetta, L. Tornberg, J. Koch, L. S. Bishop, A. A. Houck, B. R. Johnson, L. Frunzio, S. M. Girvin, and R. J. Schoelkopf. Randomized benchmarking and process tomography for gate errors in a solid-state qubit. *Phys. Rev. Lett.*, 102:090502, March 2009.

[113] C. A. Ryan, M. Laforest, and R. Laflamme. Randomized benchmarking of single- and multi-qubit control in liquid-state NMR quantum information processing. *New J. Phys.*, 11(1):013034, 2009.

[114] S. Olmschenk, R. Chicireanu, K. D. Nelson, and J. V. Porto. Randomized benchmarking of atomic qubits in an optical lattice. *New J. Phys.*, 12(11):113007, 2010.

[115] K. R. Brown, A. C. Wilson, Y. Colombe, C. Ospelkaus, A. M. Meier, E. Knill, D. Leibfried, and D. J. Wineland. Single-qubit-gate error below $\mathbf{10^{-4}}$ in a trapped ion. *Phys. Rev. A*, 84:030303, September 2011.

[116] J. P. Gaebler, A. M. Meier, T. R. Tan, R. Bowler, Y. Lin, D. Hanneke, J. D. Jost, J. P. Home, E. Knill, D. Leibfried, and D. J. Wineland. Randomized benchmarking of multiqubit gates. *Phys. Rev. Lett.*, 108:260503, June 2012.

[117] R. Barends, J. Kelly, A. Megrant, A. Veitia, D. Sank, E. Jeffrey, T. C. White, J. Mutus, A. G. Fowler, B. Campbell, Y. Chen, Z. Chen, B. Chiaro, A. Dunsworth, C. Neill, P. O'Malley, P. Roushan, A. Vainsencher, J. Wenner, A. N. Korotkov, A. N. Cleland, and J. M. Martinis. Superconducting quantum circuits at the surface code threshold for fault tolerance. *Nature*, 508:500–503, April 2014.

[118] T. Xia, M. Lichtman, K. Maller, A. W. Carr, M. J. Piotrowicz, L. Isenhower, and M. Saffman. Randomized benchmarking of single-qubit gates in a 2D array of neutral-atom qubits. *Phys. Rev. Lett.*, 114:100503, March 2015.

[119] J. T. Muhonen, A. Laucht, S. Simmons, J. P. Dehollain, R. Kalra, F. E. Hudson, S. Freer, K. M. Itoh, D. N. Jamieson, J. C. McCallum, A. S. Dzurak, and A. Morello. Quantifying the quantum gate fidelity of single-atom spin qubits in silicon by randomized benchmarking. *J. Phys. Condens. Matter*, 27(15):154205, 2015.

[120] S. Asaad, C. Dickel, N. K. Langford, S. Poletto, A. Bruno, M. A. Rol, D. Deurloo, and L. DiCarlo. Independent, extensible control of same-frequency superconducting qubits by selective broadcasting. *npj Quantum Inf.*, 2:16029, August 2016.

[121] J. J. Wallman. Randomized benchmarking with gate-dependent noise. *Quantum*, 2:47, January 2018.

[122] J. Helsen, J. J. Wallman, S. T. Flammia, and S. Wehner. Multi-qubit randomized benchmarking using few samples. *ArXiv e-prints: arXiv:1701.04299 [quant-ph]*, January 2017.

[123] A. W. Cross, E. Magesan, L. S. Bishop, J. A. Smolin, and J. M. Gambetta. Scalable randomized benchmarking of non-Clifford gates. *npj Quantum Inf.*, 2:16012, April 2016.

[124] A. Carignan-Dugas, J. J. Wallman, and J. Emerson. Characterizing universal gate sets via dihedral benchmarking. *Phys. Rev. A*, 92:060302, December 2015.

[125] J. M. Gambetta, A. D. Córcoles, S. T. Merkel, B. R. Johnson, J. A. Smolin, J. M. Chow, C. A. Ryan, C. Rigetti, S. Poletto, T. A. Ohki, M. B. Ketchen, and M. Steffen. Characterization of addressability by simultaneous randomized benchmarking. *Phys. Rev. Lett.*, 109:240504, December 2012.

[126] E. Magesan, J. M. Gambetta, and J. Emerson. Characterizing quantum gates via randomized benchmarking. *Phys. Rev. A*, 85:042311, April 2012.

[127] J. Emerson, M. Silva, O. Moussa, C. Ryan, M. Laforest, J. Baugh, D. G. Cory, and R. Laflamme. Symmetrized characterization of noisy quantum processes. *Science*, 317(5846):1893–1896, 2007.

[128] B. Lévi, C. C. López, J. Emerson, and D. G. Cory. Efficient error characterization in quantum information processing. *Phys. Rev. A*, 75:022314, February 2007.

[129] W. G. Brown and B. Eastin. Randomized benchmarking with restricted gate sets. *Phys. Rev. A*, 97(6):062323, 2018.

[130] C. Scheiderer. Semidefinite representation for convex hulls of real algebraic curves. *SIAM J. Appl. Algebra Geometry*, 2(1):1–25, 2018.

[131] C. Scheiderer. Semidefinitely representable convex sets. *ArXiv e-prints: arXiv:1612.07048 [math.OC]*, December 2016.

[132] C. Scheiderer. Spectrahedral shadows. *SIAM J. Appl. Algebra Geometry*, 2(1):26–44, 2018.

A Contributed core article: Article 1

A. K. Hashagen
Universal asymmetric quantum cloning revisited
Quantum Information and Computation, 17(9& 10):0747–0778, August 2017

Summary of article 1: Universal Asymmetric Quantum Cloning Revisited [1]

In this article, we investigate the universal asymmetric $1 \to 2$ quantum cloning problem. Quantum cloning is only possible in an approximate manner. Its exact statement is excluded by linearity of quantum mechanics (cf. theorem 2.5) and it highlights the striking difference between classical and quantum information theory. It thus forms a fundamental no-go theorem in quantum information theory and its bounds give rise to many applications (cf. section 4.1) [56–61]. In this work we look at universal quantum cloning, in which the figure of merit assessing the quality of the quantum clones is state-independent, i.e., it is independent of the input quantum state. Moreover, we study the more general case of asymmetric universal quantum cloning, in which the qualities of the clones are allowed to differ. We focus on the $1 \to 2$ quantum cloning problem; one quantum state is transformed to two approximate copies, which we call its quantum clones. Intuitively, the better the quality of one of these clones, the worse the quality of the second must be. In this article, we quantitatively describe this intuitive behavior. We derive the quantum cloning channel and we fully specify the set of all attainable single quantum clone qualities.

The universal asymmetric quantum cloning problem can be cast as an optimization problem. To derive the optimal quantum cloning channel, we maximize the quality of one of the clones, while keeping the quality of the second clone fixed. The qualities of the single clones are assessed using a distance measure that fulfills the property of joint concavity and unitary invariance. For technical reasons we furthermore assume that the origin is attainable. An example of such distance measure is the entanglement fidelity. We then identify the symmetry properties of this optimization problem. This reduces the complexity of the underlying problem and allows for an analytic solution. In theorem [1, theorem 3] we give the optimal universal $1 \to 2$ asymmetric quantum cloning channel, which is independent of the choice of distance measure:

Theorem A.1 (Optimal universal $1 \to 2$ asymmetric quantum cloning channel [1, theorem 3])**.** *The optimal universal* $1 \to 2$ *asymmetric quantum cloning channel* $T : \mathcal{M}_d \to \mathcal{M}_d \otimes \mathcal{M}_d$

for any quantum state $\rho \in \mathcal{D}_d$ is given by

$$T(\rho) = (\alpha_2 \mathbb{1} + \alpha_1 \mathbb{F}) \left(\rho \otimes \frac{\mathbb{1}}{d} \right) (\alpha_2 \mathbb{1} + \alpha_1 \mathbb{F}), \tag{A.1}$$

with $(\alpha_1)^2 + (\alpha_2)^2 + \frac{2\alpha_1\alpha_2}{d} = 1$, $\alpha_1, \alpha_2 \in \mathbb{R}$, and where $\mathbb{F}$ is the flip or swap operator.

Furthermore, we show that it is possible to rewrite the optimization problem as a semidefinite program (SDP). This might be useful when considering applications of universal asymmetric quantum cloning.

Moreover, we analytically derive the set of all achievable single quantum clone qualities for different figures of merit. The optimization problem describing the universal asymmetric quantum cloning problem can be rewritten using the notion of a one-sided polar [1, definition 3]. It is then possible to use the bipolar theorem [1, theorem 4] to fully specify the set of all attainable single quantum clone qualities. In its application it is important that the distance measure used is such that the origin is attainable. We then obtain the following results:

Theorem A.2 (Set of all attainable single clone fidelities within universal $1 \to 2$ asymmetric quantum cloning [1, theorem 5])**.** *The set of all attainable clone qualities in terms of single clone fidelities $\langle\Omega| J_{T_i} |\Omega\rangle$ with $i = 1, 2$, where $J_{T_i} := \mathrm{id} \otimes T_i (|\Omega\rangle\langle\Omega|)$ is the Choi-Jamiolkowski state of the marginals of the optimal quantum cloning channel, given by equation* (A.1)*, with $|\Omega\rangle\langle\Omega|$ being the maximally entangled state, is given by*

$$C = \left\{ x \in \mathbb{R}^2 \;\middle|\; \sup_{v \in \mathbb{R}^2} \frac{\langle x, (v, 1-v)\rangle}{\lambda_{max}\left(H_{(v,1-v)}\right)} \leq 1 \right. \\ \left. \wedge \sup_{v \in \mathbb{R}^2} \frac{\langle x, (-v, v-1)\rangle}{\lambda_{max}\left(H_{(-v,v-1)}\right)} \leq 1 \wedge \sup_{v \in \mathbb{R}^2} \frac{\langle x, (\pm v, \mp v)\rangle}{\lambda_{max}\left(H_{(\pm v,\mp v)}\right)} \leq 1 \right\}, \tag{A.2}$$

where $\lambda_{max}(H_z)$ denotes the maximum eigenvalue of

$$H_z = z_1 |\Omega\rangle\langle\Omega|_{01} \otimes \mathbb{1}_2 + z_2 |\Omega\rangle\langle\Omega|_{02} \otimes \mathbb{1}_1,$$

given by

$$\lambda_{max}\left(H_{\binom{v}{1-v}}\right) = \frac{1}{2d}\left(d + \sqrt{d^2 + 4(d^2-1)(v-1)v}\right),$$

$$\lambda_{max}\left(H_{\binom{-v}{v-1}}\right) = \begin{cases} 0 & \text{if } 0 \geq v \geq 1, \\ \frac{1}{2d}\left(-d + \sqrt{d^2 + 4(d^2-1)(v-1)v}\right) & \text{otherwise,} \end{cases}$$

$$\lambda_{max}\left(H_{\binom{\pm v}{\mp v}}\right) = v\sqrt{\frac{d^2-1}{d^2}}.$$

The upper boundary of this set is described by

$$\frac{1}{d+1}\left(\sqrt{\langle\Omega| J_{T_1} |\Omega\rangle} + \sqrt{\langle\Omega| J_{T_2} |\Omega\rangle}\right)^2 + \frac{1}{d-1}\left(\sqrt{\langle\Omega| J_{T_1} |\Omega\rangle} - \sqrt{\langle\Omega| J_{T_2} |\Omega\rangle}\right)^2 = \frac{2}{d}. \tag{A.3}$$

We extended this result to other examples of distance measures. Figures illustrating these results can be found in [1, appendix].

Statement of individual contribution

I, Anna-Lena Karolyn Hashagen, am the principal author of this article and was thus responsible for writing this article. The idea for this project was the result of many discussions with my doctoral supervisor Prof. Dr. Michael M. Wolf. I am the single author of this article and was thus solely involved in all parts of it.

Journal permission and article

Subject **Re: Permission to include article (DOI: https://doi.org/10.26421 /QIC17.9-10) in my dissertation**

From QIC Editorial <qic@rintonpress.com>
Sender QIC Editorial <qic@rintonpress.com>
To Anna-Lena Hashagen <hashagen@ma.tum.de>
Date 2018-03-26 15:31

Dear Anna-Lena,

Yes, you can your work any way you wish. Good luck in preparing your dissertation.

Regards, Wei

On 3/26/2018 9:24 AM, Anna-Lena Hashagen wrote:

> Dear Sir or Madam,
>
> I am currently preparing a cumulative dissertation. I am the author of the article
>
> Anna-Lena Hashagen
> Universal Asymmetric Quantum Cloning Revisited
> Quantum Information and Computation, Vol. 17, No. 9&10 (2017) 0747-0778
>
> and I would like to ask for permission to include this article in my dissertation.
>
> Kind regards,
> Anna-Lena Hashagen

Guidelines for Authors

Manuscripts:

1) QIC publishes papers in all areas of quantum information processing, including quantum algorithms, quantum information theory, quantum complexity theory, quantum cryptology, quantum communication and measurements, proposals and experiments on the implementation of quantum computation, communications, and entanglement in all areas of science such as ion traps, cavity QED, photons, nuclear magnetic resonance, and solid-state proposals.
2) Original articles, survey articles, reviews, tutorials, perspectives, and correspondences are all welcome. Computer science, physics and mathematics are covered. Both theories and experiments are suitable.
3) The default upper limit for each article type is: 2 (journal) pages for a correspondence, perspective, or book review; 6 pages for a letter; 40 pages for an original research article; 40 pages for a survey, review, or tutorial article.

Submissions:

1) Submission of a paper implies that it has not been published and is not being considered for publication in another journal. All submitted papers would be acknowledged and peer-reviewed.
2) Author(s), when submitting a paper, are encouraged to provide us with a) the names of two or three "guardian" editors from the list of editors and b) a list of several (preferably 3 to 5) names for potential referees. Authors are also encouraged to specify, in their cover letter, the type of the submitted article as an original article, survey article, review, tutorial, perspective, or correspondence.
3) Author(s) may request to submit their paper for the express category, which offers fast-track review. To do so, in submission letters, author(s) should a) explicitly mention that they are submitting their paper for the express category and b) provide compelling reasons why their paper is of significant interest on timely research topics that deserve express treatment.
4) A paper can be submitted by sending the manuscript in pdf format, along with a cover letter, to

QIC@rintonpress.com

Publication:

1) Once a paper is accepted for publication in QIC, the copyright is transferred to the publisher. The author retains the rights to use all or part of the paper in his (her) future publications, in which the source of the original publication is acknowledged.
2) The authors of accepted papers are requested, if necessary, to reformat their papers according to QIC layout. The instruction and latex template for preparing manuscripts for QIC can be downloaded at http://www.rintonpress.com/style

Quantum Information and Computation, Vol. 17, No. 9&10 (2017) 0747–0778

UNIVERSAL ASYMMETRIC QUANTUM CLONING REVISITED

ANNA-LENA HASHAGEN
Department of Mathematics, Technische Universität München
Boltzmannstraße 3, 85748 Garching bei München, Germany

Received July 21, 2016
Revised May 31, 2017

This paper revisits the universal asymmetric $1 \to 2$ quantum cloning problem. We identify the symmetry properties of this optimization problem, giving us access to the optimal quantum cloning map. Furthermore, we use the bipolar theorem, a famous method from convex analysis, to completely characterize the set of achievable single quantum clone qualities using the fidelity as our figure of merit; from this it is easier to give the optimal cloning map and to quantify the quality tradeoff in universal asymmetric quantum cloning. Additionally, it allows us to analytically specify the set of achievable single quantum clone qualities using a range of different figures of merit.

Keywords: Quantum Information Theory, Quantum Cloning
Communicated by: I Cirac & A Harrow

1 Introduction

One of the most fundamental, but nevertheless intriguing, feature of quantum mechanics is the impossibility to perfectly clone an arbitrary quantum state. The intrinsic linearity of quantum mechanics facilitates this remarkable difference between classical information and its quantum counterpart, quantum information, which cannot be copied in a perfect manner. This observation is known as the "No-Cloning Theorem 1" [1]. It is deeply intertwined with the impossibility of superluminal communication, the impossibility of classical teleportation as well as the impossibility of fully determining an unknown quantum state [2]. Even though the No-Cloning Theorem 1 gives rise to a lot of impossibilities, it also allows advantageous use within for example quantum cryptography.

The possibilities that open up with the study of the No-Cloning Theorem 1 gave rise to a vast research area. This research was fuelled even further by experimental advancements; in these experiments approximate quantum cloning was realized using different techniques [3, 4, 5, 6]. One question, which turns out to be especially interesting is the question of approximate quantum cloning and its inherent boundaries. An intensive review is given by Cerf and Fiurášek [7] and by Scarani, Iblisdir and Gisin [8]. Even though perfect quantum cloning is impossible, it can be done in an approximate manner. This means that we are looking for a quantum channel, which clones any input state as good as possible. This is called universal quantum cloning, because this setting is independent of the input state, i.e. the figure of merit assessing the quality of the clones is state-independent. In the case in which all clones have the same quality, the cloning procedure is named universal symmetric quantum cloning. If the clones may have different qualities, the cloning procedure is named universal asymmetric quantum cloning. The symmetric quantum cloning is thus a special case of the asymmetric quantum cloning.

Quantum cloning has been studied immensely after the universal symmetric $1 \to 2$ qubit quantum cloning machine was discovered by Bužek and Hillery in 1996 [9]. Their machine was shown to be optimal by Bruß, DiVincenzo, Ekert, Fuchs, Macchiavello and Smolin two years later [11]. At about the same time, Gisin and Massar presented their work on universal symmetric quantum cloning machines that transform N identical qubits into M identical clones and gave a numerical suggestion for optimality [12]. Full optimality was then provided through an analytical proof by Bruß, Ekert and Macchiavello [13]. Naturally, all these universal symmetric quantum cloning machines were extended to the qudit case. This was independently done by Bužek and Hillery [14] and Cerf [15], who analyzed the $1 \to 2$ quantum cloning case for qudits, as well as Werner, who constructed the unique optimal symmetric $N \to M$ qudits quantum cloning machine and together with Keyl shows full optimality using group theoretical methods [16, 17].

The thorough exploration of the asymmetric case began with papers by Bužek, Hillery and Bednik [18], Niu and Griffiths [19] and Cerf [15, 20], in which they independently analyze and derive the universal asymmetric $1 \to 2$ qubit quantum cloning machine. Furthermore, they generalized their results from qubits to qudits [21, 22]. The more general universal asymmetric $N \to M$ qubit cloning machines were introduced by Iblisdir, Acín, Gisin, Fiurážek, Filip and Cerf [23] and, using a technique from group theory, by Iblisdir, Acín and Gisin [24]. The extensions to qudits was then presented by Fiurážek, Filip and Cerf in an additional paper [25]. The inherent tradeoff among various output fidelities was further clarified and visualized

by Jiang and Yu [26]. Moreover, Ćwikliński, Horodecki and Studziński further discussed the asymmetric quantum cloning case in their paper [27], in which they provide a general result on an admissible region of fidelities for universal $1 \to N$ qubit quantum cloning machines. This result was extended to qudits by Studziński, Ćwikliński, Horodecki and Mozrzymas using a general group representation approach [28]. Simultaneously, Kay together with Ramanathan and Kaszlikowshi analyzed special cases of the universal $N \to M$ qudit quantum cloning machines, such as the universal asymmetric $1 \to N$ qudit cloning problem or the universal asymmetric $N-1 \to N$ qubit cloning problem [29, 30, 31].

This paper is concerned with the universal asymmetric $1 \to 2$ quantum cloning. We are thus interested in a quantum channel, also called optimal cloning map, that produces two good approximate clones from one input state, such that the qualities of these two clones must not be equal and are independent of the input state. In other words, if we fix the quality of one of the clones, the optimal cloning map maximizes the quality of the other clone independent of the input state. There exists a natural tradeoff between the qualities of the two clones: if the quality of one clone increases, intuitively it is clear that the quality of the other clone must decrease, complying to the No-Cloning Theorem 1. The goal of this paper is to quantify this intuitive behavior regarding the quality of the two clones and to rediscover this optimal cloning map corresponding to universal asymmetric $1 \to 2$ quantum cloning. The arising asymmetric quantum cloning map agrees with previous results; we derive it, however, using methods from Eggeling and Werner [32] and Vollbrecht and Werner [33] originally used in order to study separability properties and entanglement measures under symmetry respectively. Furthermore, in this paper we analytically derive the set of achievable single quantum clone qualities using different figures of merit by means of convex analysis techniques. This powerful but simple method is what sets it apart from previous results.

The paper is organized as follows: In the next chapter, we give a brief overview of the setting under consideration in this paper. Chapter 3 discusses figures of merit with which the quality of the clones are assessed. In order to quantify the asymmetric tradeoff in the quality of the clones, single clone figures of merit are investigated solely. In Chapter 4 we observe that the optimal quantum cloning channel is a quantum channel featuring specific symmetry properties. These symmetry properties determine the Choi-Jamiolkowski state. The Choi-Jamiolkowski channel state duality establishes that all properties of the quantum channel are encoded in the corresponding state. Reformulating the asymmetric quantum cloning problem using this Choi-Jamiolkowski state yields the optimal quantum cloning channel, given in Theorem 3 in Chapter 5. Furthermore, in this chapter, we draw the connection to semidefinite programming, which may also be used to solve the convex quantum cloning optimization problem. In Chapter 6 we use the bipolar theorem, a technique known from convex analysis, to fully characterize the set of all attainable single quantum clone fidelities. Theorem 5 summarizes this main result. Additionally, the set of all achievable single quantum clone qualities using a range of different figures of merit are given in Corollary 2. These sets are depicted in figures found in the appendix.

2 The Setting of Universal Asymmetric Quantum Cloning

We consider systems on a finite dimensional Hilbert space $\mathcal{H} = \mathbb{C}^d$. Denote as $\mathcal{M}_d$ the set of all complex-valued $d \times d$-matrices. Every quantum state is described by a density matrix $\rho \in \mathcal{M}_d$

with normalization $\mathrm{Tr}\,[\rho] = 1$ and positivity property $\rho \geq 0$. The set of all d-dimensional density matrices or quantum states is denoted as $\mathcal{D}_d := \{\rho \in \mathcal{M}_d | \rho \geq 0, \mathrm{Tr}\,[\rho] = 1\}$. A transformation of a quantum state is described by a quantum channel, which is a completely positive trace preserving linear map $T : \mathcal{M}_d \to \mathcal{M}_{d'}$. Furthermore, we denote by $\mathcal{U}(d) := \{U \in \mathcal{M}_d | UU^* = U^*U = \mathbb{1}\}$ the unitary group acting on our Hilbert space $\mathcal{H} = \mathbb{C}^d$. Moreover, $\mathbb{1}$ is the identity matrix in $\mathcal{M}_d$.

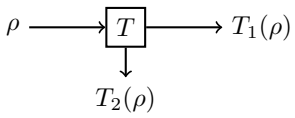


Fig. 1. The main setup of universal asymmetric $1 \to 2$ quantum cloning.

We are considering the universal asymmetric $1 \to 2$ quantum cloning case. The main setup is illustrated in Figure 1. The quantum cloning channel, $T \in \mathcal{B}(\mathcal{M}_d)$, $T : \mathcal{M}_d \to \mathcal{M}_d \otimes \mathcal{M}_d$ is a trace preserving completely positive linear map with marginal maps $T_i : \mathcal{M}_d \to \mathcal{M}_d$ for $i = 1, 2$, defined as $T_i(\rho) := \mathrm{Tr}_{\bar{i}}\,[T(\rho)]$, where the involution $i \mapsto \bar{i}$ corresponds to the permutation $\{2, 1\}$ of $\{1, 2\}$.[a] Subscripts usually denote the underlying system. The corresponding Choi-matrix defined as $\mathcal{M}_{d^3} \ni \tau_{012} := (\mathrm{id} \otimes T)(|\Omega\rangle\langle\Omega|)$, where $|\Omega\rangle\langle\Omega|$ denotes the maximally entangled state, therefore has three subscripts, 1 and 2 corresponding to the two marginals of the quantum channel T and a third subscript 0 corresponding to the identity channel to which T is tensored.

Intuition lets us postulate, that the closer $T_1(\rho)$ is to ρ, the further away is $T_2(\rho)$ to ρ. Otherwise the No-Cloning Theorem 1 is violated. In order to analyze this intuition and to quantitatively describe it, some further definitions are needed.

Theorem 1 (No-Cloning Theorem [1])**.** *Consider quantum systems on a finite dimensional Hilbert space $\mathcal{H} = \mathbb{C}^d$. There is no completely positive trace preserving linear map, called a quantum channel, $T : \mathcal{M}_d \to \mathcal{M}_d \otimes \mathcal{M}_d$ such that for all quantum states $\rho \in \mathcal{D}_d$ the following holds,*

$$T(\rho) = \rho \otimes \rho. \tag{1}$$

Proof. See [1], or for the convenience of the reader we give a proof in the following. The theorem is a consequence of linearity. Let $\{|\psi_i\rangle\langle\psi_i|\}_{i=1}^n$ be a set of orthogonal pure states and $\{\lambda_i\}_{i=1}^n$ be a set of probabilities such that $\lambda_i \neq 0$ for all i. If there was a map T as specified in the theorem, then

$$\sum_i \lambda_i T(|\psi_i\rangle\langle\psi_i|) = \sum_i \lambda_i |\psi_i\rangle\langle\psi_i| \otimes |\psi_i\rangle\langle\psi_i|, \tag{2}$$

which has rank n, while,

$$\sum_i \lambda_i T(|\psi_i\rangle\langle\psi_i|) = T(\sum_i \lambda_i |\psi_i\rangle\langle\psi_i|) = \sum_{ij} \lambda_i\lambda_j |\psi_i\rangle\langle\psi_i| \otimes |\psi_j\rangle\langle\psi_j|, \tag{3}$$

[a] If $i = 1$ then $\bar{i} = 2$ and if $i = 2$ then $\bar{i} = 1$.

which has rank n^2 $\square$.

3 The Figure of Merit assessing Single Clone Qualities

In order to assess the quality of the clones, we will consider a distance measure d on the space of linear operators. This distance measure $d(\cdot,\cdot) : \mathcal{B}(\mathcal{M}_d) \times \mathcal{B}(\mathcal{M}_d) \to \mathbb{R}_+$ quantifies the quality of a clone. Since we are interested in the asymmetric tradeoff within the quality of the clones, the figure of merit is used to quantify the quality of a single clone. We will thus compare each marginal T_i to the identity map; that is, we are going to consider $d(T_i, \mathrm{id})$ for $i = 1, 2$. Our goal is to fully specify the set of all attainable single quantum clone qualities,

$$\mathcal{C} = \left\{ z \in \mathbb{R}^2 \middle| z = \begin{pmatrix} d(T_1, \mathrm{id}) \\ d(T_2, \mathrm{id}) \end{pmatrix} \right\},$$

using this figure of merit.

Required properties of our figure of merit: Let $L, S : \mathcal{M}_d \to \mathcal{M}_d$ be quantum channels. We require the figure of merit to have the following properties due to technical reasons.

(i) Joint concavity:

$$d(L, S) \geq \lambda d\left(L^{(1)}, S^{(1)}\right) + (1-\lambda) d\left(L^{(2)}, S^{(2)}\right)$$

for all L and S, where $L = \lambda L^{(1)} + (1-\lambda) L^{(2)}$ and $S = \lambda S^{(1)} + (1-\lambda) S^{(2)}$, with $\lambda \in [0, 1]$.

(ii) Unitary invariance:

$$d(\mathbf{U} \circ L \circ \mathbf{U}^*, \mathbf{U} \circ S \circ \mathbf{U}^*) = d(L, S)$$

for all ideal channels $\mathbf{U}$ defined by $\mathbf{U}(\rho) = U \rho U^*$ with unitary $U \in \mathcal{U}(d)$ and where $\cdot^*$ denotes the adjoint or conjugate transpose.

(iii) Furthermore, for reasons that will become clear later, we require that the origin is attainable, i.e. that $\{0\} \in \mathcal{C}$. This requirement means that we are not necessarily considering a metric as a distance measure.

An example of a valid distance measure that fulfills these properties is given by a variant of the induced trace norm distance

$$d^1(T_i, \mathrm{id}) = 1 - \frac{1}{2} \sup_{\rho \in \mathcal{D}_d} \left\| T_i(\rho) - \rho \right\|_1 . \tag{4}$$

It characterizes the maximum probability of not distinguishing the outputs of the two channels T_i, $i = 1, 2$, and the identity channel id over all pure state inputs. This distance measure is specifically chosen in this way to always contain the origin,

$$\{0\} \in \mathcal{C}^1 = \left\{ z \in \mathbb{R}^2 \middle| z = \begin{pmatrix} d^1(T_1, \mathrm{id}) \\ d^1(T_2, \mathrm{id}) \end{pmatrix} \right\},$$

as illustrated in Figure 2, for reasons that will become clear later. Furthermore, we can notice that if $1 - \frac{1}{2} \sup_\rho \left\| T_i(\rho) - \rho \right\|_1 = 1$, then the marginal must be given by $T_i(\rho) = \rho$, with $i = 1, 2$. If we now let the other marginal be given by $T_{\bar{i}}(\rho) = \sigma$, with $\bar{i} = 2, 1$, for some fixed

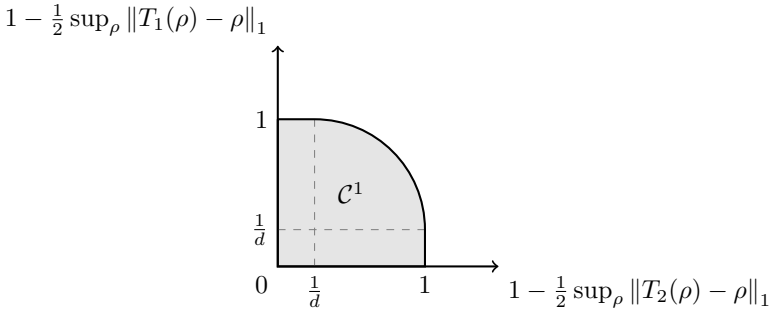


Fig. 2. Set of all attainable single quantum clone qualities $\mathcal{C}^1$. The figure of merit is $d^1(T_i,\mathrm{id}) = 1-\frac{1}{2}\sup_\rho \|T_i(\rho)-\rho\|_1$ for $i=1,2$.

quantum state $\sigma \in \mathcal{D}_d$, then $1-\frac{1}{2}\sup_\rho \|T_{\bar{i}}(\rho)-\rho\|_1 = \lambda_{\min}(\sigma)$, where $\lambda_{\min}(\sigma) \in \left[0,\frac{1}{d}\right]$ is the smallest eigenvalue of σ. These boundary points are visualized in Figure 2.

Another example, which is going to be of interest to us later on, is the fidelity

$$d^F(T_i,\mathrm{id}) = \langle\Omega|\,\tau_{0i}\,|\Omega\rangle\,, \tag{5}$$

where $\tau_{0i} := \mathrm{id}\otimes T_i\left(|\Omega\rangle\langle\Omega|\right)$ is the Choi-Jamiolkowski state of the marginal map and $|\Omega\rangle\langle\Omega|$ is the maximally entangled state with $|\Omega\rangle = \frac{1}{\sqrt{d}}\sum_{i=1}^d |ii\rangle$. It measures the overlap of the output with the maximally entangled state, if T_i acts on half a maximally entangled state.

Other examples that might be of interest are a variant of the induced Frobenius norm distance

$$d^2(T_i,\mathrm{id}) = 1-\sup_{\rho\in\mathcal{D}_d}\|T_i(\rho)-\rho\|_2\,, \tag{6}$$

a variant of the induced operator norm distance

$$d^\infty(T_i,\mathrm{id}) = 1-\sup_{\rho\in\mathcal{D}_d}\|T_i(\rho)-\rho\|_\infty\,, \tag{7}$$

and a variant of the diamond norm distance, which is a stabilized version of the induced trace norm distance,

$$d^\diamond(T_i,\mathrm{id}) = 1-\frac{1}{2}\|T_i-\mathrm{id}\|_\diamond = 1-\frac{1}{2}\sup_{\rho\in\mathcal{D}_{d^2}}\|(T_i\otimes\mathrm{id}_d)(\rho)-\rho\|_1\,, \tag{8}$$

with $i=1,2$. Note that all these distance measures are adjusted by $d^k_{\max}$, $k=1,2,\infty,\diamond$, the maximum value that the norm may take such that the origin is always contained in the corresponding set of all attainable single quantum clone qualities. This is the case, because we always look at the worst case scenario over all quantum states.

The goal is to characterize all possible quantum clone qualities; it is thus of interest to us to consider the following optimization problem

$$\sup_T\left[z_1 d^k(T_1,\mathrm{id}) + z_2 d^k(T_2,\mathrm{id})\right], \tag{9}$$

with $z_1, z_2 \in \mathbb{R}$ for different figures of merit, $k = F, 1, 2, \infty, \diamond$. This optimization problem gives the upper boundary of the set of all attainable single quantum clone qualities. The set of admissible quantum channels T is compact, since the set is bounded and closed in a finite dimensional vector space. The supremum in Eq. (9) is therefore attained for some optimal quantum channel T_{optimal}, which is the optimal quantum cloning channel, as it gives the best tradeoff possible within the qualities of the quantum clones.

4 The Symmetry Properties of the Optimal Quantum Cloning Channel

In order to find the optimal quantum cloning channel, it is of interest to us to identify special symmetry properties of this quantum channel. Let us define what we mean by a symmetrized quantum channel, because it turns out that the optimal quantum cloning channel is exactly of this type.

Definition 1 (Symmetrized quantum channel)**.** *A symmetrized quantum channel $\tilde{T} : \mathcal{M}_d \to \mathcal{M}_d$ is defined via the map*

$$\begin{aligned} T(\cdot) \mapsto \tilde{T}(\cdot) &= \int_{\mathcal{U}(d)} \mathbf{U} \circ T \circ \mathbf{U}^*(\cdot)\, \mathrm{d}\mathbf{U} \\ &= \int_{\mathcal{U}(d)} UT(U^* \cdot U)U^*\, \mathrm{d}U \quad \forall\, U \in \mathcal{U}(d), \end{aligned} \tag{10}$$

where $\mathrm{d}U$ denotes the normalized Haar measure on the unitary group $\mathcal{U}(d)$ and $\mathbf{U}(\cdot) = U \cdot U^*$ is the ideal quantum channel. Note that we will always use a tilde to denote a symmetrized quantum channel.

This symmetrization (also called twirling and in quantum information first introduced in [32, 33]) can be considered as averaging over the unitary group $\mathcal{U}(d)$ on our Hilbert space $\mathcal{H}$. Let us consider a symmetrized quantum cloning channel $\tilde{T} : \mathcal{M}_d \to \mathcal{M}_d \otimes \mathcal{M}_d$ with

$$T(\rho) \mapsto \tilde{T}(\rho) = \int_{\mathcal{U}(d)} (U \otimes U)T(U^*\rho U)(U \otimes U)^*\, \mathrm{d}U$$

for every quantum state $\rho \in \mathcal{D}_d$. It turns out that this symmetry property also applies to its marginals; see the following Lemma 1.

Lemma 1. *The marginal maps $T_i : \mathcal{M}_d \to \mathcal{M}_d$, $i = 1, 2$, of a symmetrized quantum channel $\tilde{T} : \mathcal{M}_d \to \mathcal{M}_d \otimes \mathcal{M}_d$ are symmetrized,*

$$\tilde{T}_i(\rho) = \int_{\mathcal{U}(d)} UT_i(U^*\rho U)U^*\, \mathrm{d}U,$$

for $i = 1, 2$, for every quantum state $\rho \in \mathcal{D}_d$.

Proof. The marginal maps of a symmetrized quantum channel are given by

$$\begin{aligned}\tilde{T}_i(\rho) &= \mathrm{Tr}_{\bar{i}}\left[\tilde{T}(\rho)\right] \\ &= \mathrm{Tr}_{\bar{i}}\left[\int_{\mathcal{U}(d)} (U\otimes U)\, T\left(U^*\rho U\right)(U\otimes U)^*\ \mathrm{d}U\right] \\ &= \int_{\mathcal{U}(d)} \mathrm{Tr}_{\bar{i}}\left[(U\otimes U)\, T\left(U^*\rho U\right)(U\otimes U)^*\right]\ \mathrm{d}U \\ &= \int_{\mathcal{U}(d)} U\,\mathrm{Tr}_{\bar{i}}\left[T\left(U^*\rho U\right)\right]U^*\,\mathrm{d}U \\ &= \int_{\mathcal{U}(d)} UT_i\left(U^*\rho U\right)U^*\,\mathrm{d}U,\end{aligned}$$

for both marginals $i = 1, 2$ □.

These symmetrized marginal maps satisfy the so-called covariance property $\tilde{T}_i(V\rho V^*) = V\tilde{T}_i(\rho)V^*$ with $i = 1, 2$ for unitary $V \in \mathcal{U}(d)$, as stated in the Lemma 2 below [32, 33].

Definition 2 (Covariant)**.** *A quantum channel $T : \mathcal{M}_d \to \mathcal{M}_d$ is called covariant with respect to V if*

$$T\left(V\cdot V^*\right) = VT\left(\cdot\right)V^*$$

holds for all $V \in \mathcal{U}(d)$.

Lemma 2. *A symmetrized quantum channel $\tilde{T} : \mathcal{M}_d \to \mathcal{M}_d$ is covariant.*

Proof. Using the definition of a symmetrized quantum channel yields

$$\begin{aligned}&\tilde{T}\left(V\cdot V^*\right) \\ =&\int_{\mathcal{U}(d)} UT\left(U^*V\cdot V^*U\right)U^*\,\mathrm{d}U \\ =&V\int_{\mathcal{U}(d)} V^*UT\left(U^*V\cdot V^*U\right)U^*V\,\mathrm{d}UV^* \\ =&V\int_{\mathcal{U}(d)} WT\left(W^*\cdot W\right)W^*\,\mathrm{d}(VW)V^* \\ =&V\int_{\mathcal{U}(d)} WT\left(W^*\cdot W\right)W^*\,\mathrm{d}(W)V^* \\ =&V\tilde{T}(\cdot)V^*,\end{aligned}$$

where we have defined $W := V^*U$ with unitaries $U, V \in \mathcal{U}(d)$ and used the properties of the Haar measure □.

The figure of merit that assesses the single clone qualities is influenced by this covariance

property in the following way, namely

$$\begin{aligned} d^k(\tilde{T}_i, \mathrm{id}) &= d^k\left(\int_{\mathcal{U}(d)} \mathbf{U} \circ T_i \circ \mathbf{U}^* \, \mathrm{d}\mathbf{U}, \mathrm{id}\right) \\ &\geq \int_{\mathcal{U}(d)} d^k\left(\mathbf{U} \circ T_i \circ \mathbf{U}^*, \mathrm{id}\right) \, \mathrm{d}\mathbf{U} \\ &= \int_{\mathcal{U}(d)} d^k\left(T_i, \mathbf{U} \circ \mathrm{id} \circ \mathbf{U}^*\right) \, \mathrm{d}\mathbf{U} \\ &= d^k(T_i, \mathrm{id}), \end{aligned}$$

for $k = F, 1, 2, \infty, \diamond$, where we have used the joint concavity property (i) of the figure of merit. The optimization problem given by Eq. (9) therefore simplifies, because the supremum is attained for a symmetrized quantum channel $\tilde{T}$, i.e. we have

$$\sup_{T} \left[z_1 d^k(T_1, \mathrm{id}) + z_2 d^k(T_2, \mathrm{id})\right] = \sup_{\tilde{T}} \left[z_1 d^k(\tilde{T}_1, \mathrm{id}) + z_2 d^k(\tilde{T}_2, \mathrm{id})\right], \tag{11}$$

with $z_1, z_2 \in \mathbb{R}$, for $k = F, 1, 2, \infty, \diamond$.

5 The Optimal Quantum Cloning Channel

The last chapter has shown that the optimal quantum cloning channel is of a symmetrized form. This gives rise to a specific structure of its Choi-Jamiolkowski state, which we are going to exploit to solve the optimization problem in Eq. (11). This chapter therefore discusses the implication of the symmetrized optimal quantum cloning channel on its Choi-Jamiolkowski state and uses this additional structure to derive the optimal quantum cloning channel.

The Choi-Jamiolkowski state of a quantum channel T is defined as

$$\tau := \mathrm{id} \otimes T\left(|\Omega\rangle\langle\Omega|\right), \tag{12}$$

where $|\Omega\rangle = \frac{1}{\sqrt{d}} \sum_{i=1}^{d} |ii\rangle$. We would like to simplify our optimization problem given by Eq. (11) even further using this Choi-Jamiolkowski state. For this purpose, we would like to show that $\left[\tau, \bar{U} \otimes U \otimes U\right] = 0$ for unitary $U \in \mathcal{U}(d)$, in the case of a symmetrized quantum channel, where $\bar{\cdot}$ denotes the complex conjugate.

Lemma 3. *For a Choi-Jamiolkowski state $\tau \in \mathcal{M}_{d^3}$,*

$$\left[\tau, \bar{U} \otimes U \otimes U\right] = 0 \quad \forall \; U \in \mathcal{U}(d)$$

is equivalent to

$$\int_{\mathcal{U}(d)} \left(\bar{U} \otimes U \otimes U\right) \tau \left(\bar{U} \otimes U \otimes U\right)^* \, \mathrm{d}U = \tau.$$

The proof of this Lemma 3 can be found in the Appendix A.1. We may now use this Lemma 3 to prove the following Corollary 1.

Corollary 1. *Let* $\tau := \mathrm{id} \otimes \tilde{T}\left(|\Omega\rangle\langle\Omega|\right)$ *be the Choi-Jamiolkowski state of some symmetrized quantum channel* $\tilde{T}$*, as in Definition 1, then the following holds*

$$\left[\tau, \bar{U} \otimes U \otimes U\right] = 0 \quad \forall\, U \in \mathcal{U}(d)\,.$$

Proof. Remembering that $U \otimes \mathbb{1}\,|\Omega\rangle = \mathbb{1} \otimes U^T\,|\Omega\rangle$, such that $U \otimes \bar{U}\,|\Omega\rangle = |\Omega\rangle$, we get

$$\begin{aligned}
&\int_{\mathcal{U}(d)} \left(\bar{U} \otimes U \otimes U\right) \tau \left(\bar{U} \otimes U \otimes U\right)^* \mathrm{d}U \\
=&\int_{\mathcal{U}(d)} \left(\bar{U} \otimes U \otimes U\right) \left(\mathrm{id} \otimes \tilde{T}\,|\Omega\rangle\langle\Omega|\right) \left(\bar{U} \otimes U \otimes U\right)^* \mathrm{d}U \\
=&\int_{\mathcal{U}(d)} (\mathbb{1} \otimes U \otimes U) \left(\mathrm{id} \otimes \tilde{T}\right) \left((\mathbb{1} \otimes U^*)\,|\Omega\rangle\langle\Omega|\,(\mathbb{1} \otimes U)\right) (\mathbb{1} \otimes U \otimes U)^* \mathrm{d}U \\
=&\tau,
\end{aligned}$$

due to the covariance property of $\tilde{T}$. Application of Lemma 3 finishes the proof □.

Proposition 1. *If for a Choi-Jamiolkowski state* $\tau \in \mathcal{M}_{d^3}$ *we have that*

$$\left[\tau, \bar{U} \otimes U \otimes U\right] = 0 \quad \forall\, U \in \mathcal{U}(d)\,,$$

then

$$\left[\tau^{t_0}, U \otimes U \otimes U\right] = 0,$$

where t_0 *denotes the partial transpose on the first system.*

The proof of this Proposition 1 can be found in the Appendix A.2.

The optimization problem given by Eq. (11) was reduced to a supremum over all symmetrized quantum channels, because we found that the optimal quantum cloning channel must be of this form. Using Corollary 1, we may, without loss of generality, restrict to quantum cloning channels whose Choi-Jamiolkowski matrix τ commutes with $\left\{\bar{U} \otimes U \otimes U : U \in \mathcal{U}(d)\right\}$ and we would like to reformulate our problem by means of τ.

Theorem 2 (Weyl [34, Chapter IV])**.** *Let* $\mathcal{H}$ *be a finite-dimensional Hilbert space. If an operator* τ *acting on* $\mathcal{H}^{\otimes n}$ *fulfills* $[\tau, U^{\otimes n}] = 0$ *for all unitaries* $U \in \mathcal{U}(d)$*, then it is a linear combination of operators* V_π *representing the permutation group on* $\mathcal{H}^{\otimes n}$*,*

$$\tau = \sum_{\pi \in \mathrm{S}_n} a_\pi V_\pi,$$

where S_n *is the symmetric group on* n *elements,* π *are all possible permutations of* n *elements and* $a_\pi \in \mathbb{C}$*. The permutation operators* V_π *are defined via*

$$V_\pi\left(v_1 \otimes \ldots \otimes v_n\right) = v_{\pi^{-1}(1)} \otimes \ldots \otimes v_{\pi^{-1}(n)}.$$

Proof. The theorem immediately follows from [35, Theorem IX.11.5]. Denote by $\mathrm{SU}(d)$ the special unitary group of finite degree d and by S_n the symmetric group on n elements.

Let $\mathcal{A}$ be the group algebra of $\mathrm{SU}(d)$ and $\mathcal{B}$ be the group algebra of S_n generated by their unitary representation on $\mathcal{H}$. Since $\mathrm{SU}(d)$ and S_n act dually on $\left(\mathbb{C}^d\right)^{\otimes n}$, we have $\mathcal{A}' = \mathcal{B}$. The commutant is thus exactly the algebra generated by the permutation operators V_π. If an operator commutes with all unitaries of the form $U^{\otimes n}$, it must therefore be an element of this algebra, i.e. a linear combination of permutation operators □.

Considering Lemma 1, we know that τ^{t_0} commutes with all unitaries of the form $U \otimes U \otimes U$. Furthermore, by Theorem 2, τ^{t_0} must be given by a linear combination of permutation operators, in our case acting on three elements,

$$\tau_{012}^{t_0} = \sum_{\pi \in \mathrm{S}_3} a_\pi V_\pi = a_1 \mathbb{1} + a_2 V_{(01)} + a_3 V_{(02)} + a_4 V_{(12)} + a_5 V_{(012)} + a_6 V_{(210)}, \tag{13}$$

with $a_\pi \in \mathbb{C}$, where $V_{(01)}$ denotes the permutation operator of the first two factors (similarly for $V_{(02)}$ and $V_{(12)}$), $V_{(012)}$ denotes the cyclic permutation and similarly $V_{(210)}$ denotes the anticyclic permutation [32, 33].

The marginal maps are thus given as

$$\tau_{01}^{t_0} = \mathrm{Tr}_2[\tau_{012}^{t_0}] = (a_1 d + a_3 + a_4)\mathbb{1} + (a_2 d + a_5 + a_6)\mathbb{F}, \tag{14a}$$
$$\tau_{01} = (a_1 d + a_3 + a_4)\mathbb{1} + (a_2 d + a_5 + a_6) d \left|\Omega\right\rangle\!\left\langle\Omega\right|, \tag{14b}$$
$$\tilde{T}_1(\rho) = (a_1 d + a_3 + a_4) d \mathbb{1}\, \mathrm{Tr}[\rho] + (a_2 d + a_5 + a_6) d\rho, \tag{14c}$$

and

$$\tau_{02}^{t_0} = \mathrm{Tr}_1[\tau_{012}^{t_0}] = (a_1 d + a_2 + a_4)\mathbb{1} + (a_3 d + a_5 + a_6)\mathbb{F}, \tag{15a}$$
$$\tau_{02} = (a_1 d + a_2 + a_4)\mathbb{1} + (a_3 d + a_5 + a_6) d \left|\Omega\right\rangle\!\left\langle\Omega\right|, \tag{15b}$$
$$\tilde{T}_2(\rho) = (a_1 d + a_2 + a_4) d \mathbb{1}\, \mathrm{Tr}[\rho] + (a_3 d + a_5 + a_6) d\rho, \tag{15c}$$

with $a_1, \ldots, a_6 \in \mathbb{C}$, where we again denote by $\left|\Omega\right\rangle\!\left\langle\Omega\right| := \frac{1}{d}\sum_{i,j=1}^{d} \left|ii\right\rangle\!\left\langle jj\right|$ the maximally entangled state and by $\mathbb{F} := \sum_{i,j=1}^{d} \left|ji\right\rangle\!\left\langle ij\right|$ the flip (or swap) operator.

As a quantum channel, $\tilde{T}$ is a completely positive trace preserving linear map and it must thus fulfill specific properties. Due to the Choi-Jamiolkowski state-channel duality, the operator τ encodes all of its properties [36, Chapter 4.4.3]. Denote by $\tilde{T}^*$ the dual of the quantum channel $\tilde{T}$ corresponding to the Heisenberg picture.

Properties:

(i) Hermiticity: $\tau = \tau^*$, i.e.

$$a_1, \ldots, a_4 \in \mathbb{R} \text{ and } a_5 = \bar{a}_6 \in \mathbb{C}.$$

(ii) Normalization: $\mathrm{Tr}[\tau] = \frac{1}{d}\,\mathrm{Tr}\left[\tilde{T}^*(\mathbb{1})\right]$, i.e.

$$a_1 d^3 + (a_2 + a_3 + a_4) d^2 + (a_5 + a_6) d = 1.$$

(iii) Preservation of trace: $\tilde{T}^*(\mathbb{1}) = \mathbb{1}$ if and only if $\mathrm{Tr}_{12}[\tau] = \frac{\mathbb{1}}{d}$.

(iv) Complete positivity: $\tilde{T}$ is completely positive if and only if $\tau \geq 0$.

Note that if $\tilde{T}(\rho)$ is a completely positive trace preserving linear map, then so are its marginal maps $\tilde{T}_i(\rho) := \mathrm{Tr}_{\bar{i}}\left[\tilde{T}(\rho)\right]$, $i = 1, 2$.

In order to simplify notation and to visualize agreement to previously known results [21, 22], let

$$\begin{aligned} (\alpha_1)^2 &:= a_1 d^3 + a_3 d^2 + a_4 d^2, & (\beta_1)^2 &:= a_2 d^2 + a_5 d + a_6 d, \\ (\alpha_2)^2 &:= a_1 d^3 + a_2 d^2 + a_4 d^2, & (\beta_2)^2 &:= a_3 d^2 + a_5 d + a_6 d. \end{aligned}$$

Then the Choi-Jamiolkowski states τ_{0i}, $i = 1, 2$, of the marginal maps $\tilde{T}_i$ are

$$\tau_{0i} = \alpha_i^2 \frac{\mathbb{1}}{d^2} + \beta_i^2 \left|\Omega\right\rangle\!\left\langle\Omega\right|. \tag{16}$$

The preservation of trace, property (iii), namely $\mathrm{Tr}_i\left[\tau_{0i}\right] = \mathbb{1}/d$, gives a condition on β_i, namely that

$$\begin{aligned} \mathrm{Tr}_i\left[\tau_{0i}\right] = \mathrm{Tr}_i\left[\alpha_i^2 \frac{\mathbb{1}}{d^2} + \beta_i^2 \left|\Omega\right\rangle\!\left\langle\Omega\right|\right] = (\alpha_i^2 + \beta_i^2)\frac{\mathbb{1}}{d} = \frac{\mathbb{1}}{d} \\ \Leftrightarrow \beta_i^2 = 1 - \alpha_i^2. \end{aligned}$$

Another property that the marginals must fulfill is complete positivity, property (iv), namely $\tau_{0i} \geq 0$. This yields

$$\begin{aligned} \tau_{0i} = \alpha_i^2 \frac{\mathbb{1}}{d^2} + \beta_i^2 \left|\Omega\right\rangle\!\left\langle\Omega\right| \geq 0 \\ \Leftrightarrow \alpha_i^2 \geq 0 \text{ and } \beta_i^2 \geq -\frac{\alpha_i^2}{d^2}. \end{aligned}$$

Therefore, the marginal maps and their corresponding Choi-Jamiolkowski states are given as

$$\tau_{0i} = \alpha_i^2 \frac{\mathbb{1}}{d^2} + (1 - \alpha_i^2) \left|\Omega\right\rangle\!\left\langle\Omega\right|, \tag{17a}$$

$$\tilde{T}(\rho) = \alpha_i^2 \frac{\mathbb{1}}{d} \mathrm{Tr}[\rho] + (1 - \alpha_i^2)\rho, \tag{17b}$$

with $\alpha_i^2 \in \left[0, \frac{d^2}{d^2-1}\right]$.

Since these properties must not only hold for the marginals, but also for the full quantum channel, consider

$$\tau_{012} = \sum_{\pi \in \mathrm{S}_3} a_\pi V_\pi^{t_0}. \tag{18}$$

The preservation of trace property (iii), namely that $\mathrm{Tr}_{12}\left[\tau_{012}\right] = \mathbb{1}/d$, yields

$$\begin{aligned} \mathrm{Tr}_{12}\left[\tau_{012}\right] = \mathrm{Tr}_{12}\left[\sum_{\pi \in \mathrm{S}_3} a_\pi V_\pi^{t_0}\right] = (a_1 d^2 + a_2 d + a_3 d + a_4 d + a_5 + a_6)\mathbb{1} = \frac{\mathbb{1}}{d} \\ \Leftrightarrow a_1 d^2 + a_2 d + a_3 d + a_4 d + a_5 + a_6 = \frac{1}{d}. \end{aligned}$$

Deciding complete positivity, property (iv), is a bit more tricky and we thus follow the idea of the following papers [32, 33]. One should first of all notice that

$$\mathcal{A}^{t_0} := \left\{ \sum_{\pi \in \mathrm{S}_3} a_\pi V_\pi^{t_0} \middle| a_\pi \in \mathbb{C} \right\}$$

is a six-dimensional non-commutative unital C^*-algebra. In general, if a von Neumann algebra $\mathcal{B} \subseteq \mathcal{A} \simeq \mathcal{M}_d(\mathbb{C})$ is a subalgebra of a finite-dimensional matrix algebra, then there exists a unitary U such that

$$\mathcal{B} = U \left(0 \oplus \bigoplus_{k=1}^{K} \mathcal{M}_{d_k} \otimes \mathbb{1}_{m_k} \right) U^*,$$

for a decomposition of the Hilbert space $\mathbb{C}^d = \mathbb{C}^{d_0} \oplus \bigoplus_{k=1}^{K} \mathbb{C}^{d_k} \otimes \mathbb{C}^{m_k}$, where each factor k is isomorphic to a full matrix algebra of dimension d_k^2 which appears with multiplicity m_k [37, Chapter 3.6]. Since von Neumann algebras and C^*-algebras coincide in finite dimensions, $\mathcal{A}^{t_0}$ is isomorphic to a sum of two one dimensional and a two dimensional matrix algebra, i.e. $6 = \sum_{k=1}^{K} d_k^2 = 2^2 + 1^2 + 1^2$ (Note that it cannot be a sum of six one dimensional matrix algebras, due to the non-commutativity). Using the same notation as in [32], namely,

$$\begin{aligned} X &= V_{(01)}^{t_0} \quad \text{and} \\ V &= V_{(12)}^{t_0} = V_{(12)}, \end{aligned}$$

with $X^* = X$ and $V^* = V$, such that

$$\begin{aligned} \mathbb{1}^{t_0} &= \mathbb{1}, \\ V_{(02)}^{t_0} &= VXV, \\ V_{(012)}^{t_0} &= XV, \\ V_{(210)}^{t_0} &= VX, \end{aligned}$$

we get that

$$\begin{aligned} X^2 &= dX, \\ V^2 &= \mathbb{1} \quad \text{and} \\ XVX &= X. \end{aligned}$$

A convenient basis is then given by [32]

$$\begin{aligned}
S_+ &= \frac{\mathbb{1}+V}{2}\left(\mathbb{1}-\frac{2X}{d+1}\right)\frac{\mathbb{1}+V}{2},\\
S_- &= \frac{\mathbb{1}-V}{2}\left(\mathbb{1}-\frac{2X}{d-1}\right)\frac{\mathbb{1}-V}{2},\\
S_0 &= \frac{1}{d^2-1}\left(d(X+VXV)-(XV+VX)\right),\\
S_1 &= \frac{1}{d^2-1}\left(d(XV+VX)-(X+VXV)\right),\\
S_2 &= \frac{1}{\sqrt{d^2-1}}\left(X-VXV\right),\\
S_3 &= \frac{i}{\sqrt{d^2-1}}\left(XV-VX\right).
\end{aligned}$$

Denoting by $s_k(\rho) := \mathrm{Tr}[\rho S_k]$ for $k \in \{+,-,0,1,2,3\}$ and using the results of Eggeling and Werner [32], we get the following criteria for complete positivity,

$$s_+, s_-, s_0 \geq 0, \qquad s_0^2 \geq s_1^2+s_2^2+s_3^2, \qquad s_+ + s_- + s_0 = 1.$$

Translating this result back into our original notation (see Appendix B) reduces the optimization problem given in Eq. (11) to the following convex optimization.
Find

$$\sup_{\tilde{T}}\left[z_1 d^k\left(\tilde{T}_1,\mathrm{id}\right)+z_2 d^k\left(\tilde{T}_2,\mathrm{id}\right)\right], \tag{19a}$$

for $k = F, 1, 2, \infty, \diamond$ with $z_1, z_2 \in \mathbb{R}$, where the supremum is taken over all quantum channels of the form

$$\begin{aligned}
\tilde{T}(\rho) = a_1 d\mathbb{1}\,\mathrm{Tr}[\rho] + a_2 d^2\rho\otimes\frac{\mathbb{1}}{d} + a_3 d^2\frac{\mathbb{1}}{d}\otimes\rho + a_4 d\mathbb{F}\,\mathrm{Tr}[\rho]\\
+ a_5 d^2\left(\rho\otimes\frac{\mathbb{1}}{d}\right)\mathbb{F} + a_6 d^2\mathbb{F}\left(\rho\otimes\frac{\mathbb{1}}{d}\right),
\end{aligned} \tag{19b}$$

with the corresponding Choi-Jamiolkowski state given by

$$\begin{aligned}
\tau_{012} = a_1\mathbb{1}_{012} + a_2 d\,|\Omega\rangle\langle\Omega|_{01}\otimes\mathbb{1}_2 + a_3 d\,|\Omega\rangle\langle\Omega|_{02}\otimes\mathbb{1}_1\\
+ a_4\mathbb{1}_0\otimes\mathbb{F}_{12} + a_5\sum_{ijk}|jjk\rangle\langle iki| + a_6\sum_{ijk}|kjk\rangle\langle iij|,
\end{aligned} \tag{19c}$$

such that

$$0 \leq a_1 + a_4, \tag{19d}$$

$$0 \leq (a_1 - a_4)\frac{1}{2}d(d-2)(d+1), \tag{19e}$$

$$0 \leq 2a_1 + (a_2+a_3)d + a_5 + a_6, \tag{19f}$$

$$1 = a_1 d^3 + (a_2+a_3+a_4)d^2 + (a_5+a_6)d, \tag{19g}$$

$$-(a_2+a_4)(a_3+a_4)+(a_1+a_5)(a_1+a_6) \\ +(a_1(a_2+a_3)-a_4(a_5+a_6))d+(a_2a_3-a_5a_6)d^2\geq 0. \quad (19h)$$

In case $d>2$, it is then clear that $a_1=a_4=0$ and without loss of generality $a_5=a_6\in\mathbb{R}$. Then the optimal cloning map and its Choi-Jamiolkowski state is given by

$$\tilde{T}(\rho)=(\alpha_2\mathbb{1}+\alpha_1\mathbb{F})\left(\rho\otimes\frac{\mathbb{1}}{d}\right)(\alpha_2\mathbb{1}+\alpha_1\mathbb{F}), \quad (20a)$$

$$\tau_{012}=(\alpha_2)^2\left(|\Omega\rangle\langle\Omega|_{01}\otimes\frac{\mathbb{1}_2}{d}\right)+(\alpha_1)^2\left(|\Omega\rangle\langle\Omega|_{02}\otimes\frac{\mathbb{1}_1}{d}\right) \\ +\frac{\alpha_1\alpha_2}{d^2}\sum_{ijk}|jjk\rangle\langle iki|+\frac{\alpha_1\alpha_2}{d^2}\sum_{ijk}|kjk\rangle\langle iij|, \quad (20b)$$

with

$$(\alpha_1)^2+(\alpha_2)^2+\frac{2\alpha_1\alpha_2}{d}=1. \quad (20c)$$

In the case $d=2$, however, the second inequality given by Eq. (19e) vanishes. The optimization therefore does not necessarily yield the result $a_1=a_4=0$ anymore, since these might now take negative values. It turns out that this is a freedom in the parametrization, however, still yielding the same universal optimal quantum cloning channel. We may therefore state the following Theorem 3 in full agreement with [21, 22, 23], in which the optimal universal $1\to 2$ asymmetric quantum cloning channel has been derived too. We have, however, mostly used the symmetry idea of Eggeling and Werner as well as Vollbrecht and Werner [32, 33] that exploit a similar symmetry property of the quantum states to study separability properties and entanglement measures.

Theorem 3 (Optimal universal $1\to 2$ asymmetric quantum cloning channel)**.** *The optimal universal $1\to 2$ asymmetric quantum cloning channel $\tilde{T}_{optimal}:\mathcal{M}_d\to\mathcal{M}_d\otimes\mathcal{M}_d$ for any quantum state $\rho\in\mathcal{D}_d$ is given by*

$$\tilde{T}_{optimal}(\rho)=(\alpha_2\mathbb{1}+\alpha_1\mathbb{F})\left(\rho\otimes\frac{\mathbb{1}}{d}\right)(\alpha_2\mathbb{1}+\alpha_1\mathbb{F}), \quad (21)$$

with $(\alpha_1)^2+(\alpha_2)^2+\frac{2\alpha_1\alpha_2}{d}=1$, $\alpha_1,\alpha_2\in\mathbb{R}$, and where $\mathbb{F}:=\sum_{i,j=1}^{d}|ji\rangle\langle ij|$ is the flip (or swap) operator.

What is interesting to notice is that as the dimension of the underlying Hilbert space d increases, the optimal cloning map approaches the trivial approach to quantum cloning. The trivial approach is represented by the quantum channel

$$T_{\text{trivial}}(\rho)=\alpha\rho\otimes\frac{\mathbb{1}}{d}+(1-\alpha)\frac{\mathbb{1}}{d}\otimes\rho,$$

where $\alpha\in[0,1]$. Instead of cloning the quantum state ρ, an identity channel is applied and an additional state is prepared, the maximally mixed state. Thus, in the limit as the dimension of the underlying Hilbert space increases, $d\to\infty$, even approximate quantum cloning is not possible.

5.1 Determining achievable quantum clone qualities numerically

In order to support our findings, it is also possible to rewrite the optimization problem given by Eq. (19) as a semidefinite program (SDP) [40]. To solve this semidefinite program, we used cvx, a package for specifying and solving convex programs [41, 42]. Let

$$z = \begin{pmatrix} z_1 \\ z_2 \end{pmatrix} \in \mathbb{R}^2 \quad \text{and} \quad D = \begin{pmatrix} d^k(T_1, \mathrm{id}) \\ d^k(T_2, \mathrm{id}) \end{pmatrix},$$

for $k = F, 1, 2, \infty, \diamond$ with $T_i(\cdot) = \alpha_i^2 \frac{\mathbb{1}}{d} \mathrm{Tr}[\cdot] + \left(1 - \alpha_i^2\right) \mathrm{id}(\cdot)$.
Maximise

$$z^T \cdot D$$

subject to

$$s_0 \oplus s_+ \oplus s_- \oplus \begin{pmatrix} s_0 + s_3 & s_1 + is_2 \\ s_1 + is_2 & s_0 - s3 \end{pmatrix} \oplus s_+ + s_- + s_0 - 1 \\ \oplus 1 - (s_+ + s_- + s_0) \oplus \alpha_1^2 \oplus \alpha_2^2 \oplus \frac{d^2}{d^2-1} - \alpha_1^2 \oplus \frac{d^2}{d^2-1} - \alpha_2^2 \geq 0.$$

The corresponding analytical results for different figures of merit are shown in Figure B.1 up to Figure B.5 in the Appendix B.

6 The Set of all achievable Single Quantum Clone Qualities

In the previous part, we were only interested in the optimal asymmetric quantum cloning channel describing the boundary of the set of all achievable single quantum clone qualities. In this chapter, we analytically derive this set using different figures of merit. Let us, however, turn to the fidelity $d^F(T_i, \mathrm{id}) = \langle\Omega| \tau_{0i} |\Omega\rangle$ for $i = 1, 2$, where $\tau_{0i} = \mathrm{id} \otimes \tilde{T}_i(|\Omega\rangle\langle\Omega|)$, first, such that the optimization problem is given by

$$\sup_{\tilde{T}} \left[z_1 d^F(\tilde{T}_1, \mathrm{id}) + z_2 d^F(\tilde{T}_2, \mathrm{id}) \right] = \sup_{\substack{\tau \geq 0 \\ \mathrm{Tr}_{12}[\tau] = \frac{\mathbb{1}_0}{d}}} \left[z_1 \langle\Omega| \tau_{01} |\Omega\rangle + z_2 \langle\Omega| \tau_{02} |\Omega\rangle \right]. \tag{22}$$

This is visualized in Figure 3, which shows the set of all attainable qualities of the two quantum clones,

$$\mathcal{C}^F = \left\{ z \in \mathbb{R}^2 \middle| z = \begin{pmatrix} \langle\Omega| \tau_{01} |\Omega\rangle \\ \langle\Omega| \tau_{02} |\Omega\rangle \end{pmatrix} \right\}. \tag{23}$$

First of all, we notice that for $i = 1, 2$ and $\bar{i} = 2, 1$, if the overlap of τ_{0i} with the maximally entangled state is $\langle\Omega| \tau_{0i} |\Omega\rangle = 1$ yielding $\tau_{0i} = |\Omega\rangle\langle\Omega|_{0i}$, then the overall state must be $\tau_{012} = |\Omega\rangle\langle\Omega|_{0i} \otimes \frac{\mathbb{1}_{\bar{i}}}{d}$, such that the other marginal state turns out to be $\tau_{0\bar{i}} = \frac{\mathbb{1}_0}{d} \otimes \frac{\mathbb{1}_{\bar{i}}}{d}$. This gives $\langle\Omega| \tau_{0\bar{i}} |\Omega\rangle = \frac{1}{d^2}$. Furthermore, if the overlap of τ_{0i} with the maximally entangled state is $\langle\Omega| \tau_{0i} |\Omega\rangle = 0$ yielding $\tau_{0i} = \frac{d^2}{d^2-1} \frac{\mathbb{1}}{d^2} - \frac{1}{d^2-1} |\Omega\rangle\langle\Omega|$, then the other marginal state must be $\tau_{0\bar{i}} = \frac{1}{d^2-1} \frac{\mathbb{1}}{d^2} + \frac{d^2-2}{d^2-1} |\Omega\rangle\langle\Omega|$, such that its overlap with the maximally entangled state is $\langle\Omega| \tau_{0\bar{i}} |\Omega\rangle = \frac{d^2-1}{d^2}$. This gives four extreme points of our set $\mathcal{C}^F$ as illustrated in Figure 3.

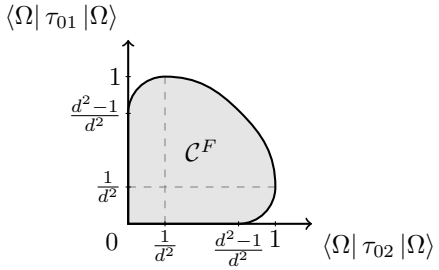


Fig. 3. Set of all attainable single quantum clone fidelities $\mathcal{C}^F$. The figure of merit is the single clone fidelity $d^F(T_i, \mathrm{id}) = \langle\Omega|\,\tau_{0i}\,|\Omega\rangle$ for $i = 1, 2$, where $\tau_{0i} := \mathrm{id} \otimes T_i\,(|\Omega\rangle\langle\Omega|)$ is the Choi-Jamiolkowski state.

Lemma 4. *Let τ be the Choi-Jamiolkowski operator of a symmetrized quantum channel. Then*

$$\tau_{012} \geq 0 \text{ and } \mathrm{Tr}_{12}\left[\tau_{012}\right] = \frac{\mathbb{1}_0}{d}$$

is equivalent to

$$\tau_{012} \geq 0 \text{ and } \mathrm{Tr}\left[\tau_{012}\right] = 1.$$

Proof. If $\mathrm{Tr}_{12}\left[\tau_{012}\right] = \frac{\mathbb{1}_0}{d}$, then taking the full trace gives $\mathrm{Tr}_{012}\left[\tau_{012}\right] = 1$. The other direction follows from the form of the Choi-Jamiolkowsi state of a symmetrized quantum channel. Using $\tau_{012} = \sum_{\pi\in \mathrm{S}_3} a_\pi V_\pi^{t_0}$ gives

$$\begin{aligned} &\mathrm{Tr}_{012}\left[\tau_{012}\right] = 1 \\ \Leftrightarrow &\sum_{\pi\in \mathrm{S}_3} a_\pi \,\mathrm{Tr}_{012}\left[V_\pi^{t_0}\right] = 1 \\ \Leftrightarrow &\left(a_1 d^2 + (a_2 + a_3 + a_4)d + a_5 + a_6\right) d = 1. \end{aligned}$$

Now

$$\mathrm{Tr}_{12}\left[\tau_{012}\right] = \left(a_1 d^2 + (a_2 + a_3 + a_4)d + a_5 + a_6\right) \mathbb{1}_0 = \frac{\mathbb{1}_0}{d}$$

□.

In order to describe the set of all achievable single clone qualities, consider

$$H_z = z_1\,|\Omega\rangle\langle\Omega|_{01} \otimes \mathbb{1}_2 + z_2\,|\Omega\rangle\langle\Omega|_{02} \otimes \mathbb{1}_1, \quad \text{with } z_1, z_2 \in \mathbb{R}.$$

Then, one notices that our optimization problem given by Eq. (22) may be rewritten using

Lemma 4 as

$$\begin{aligned}
&\sup_{\substack{\tau\geq 0\\ \mathrm{Tr}_{12}[\tau]=\frac{1_0}{d}}} \left[z_1 \langle\Omega|\tau_{01}|\Omega\rangle + z_2 \langle\Omega|\tau_{02}|\Omega\rangle\right] \\
&= \sup_{\substack{\tau\geq 0\\ \mathrm{Tr}[\tau]=1}} \left[z_1 \langle\Omega|\tau_{01}|\Omega\rangle + z_2 \langle\Omega|\tau_{02}|\Omega\rangle\right] \\
&= \sup_{\substack{\tau\geq 0\\ \mathrm{Tr}[\tau]=1}} \mathrm{Tr}\left[z_1 |\Omega\rangle\langle\Omega|_{01} \otimes \mathbb{1}_2\tau_{012} + z_2 |\Omega\rangle\langle\Omega|_{02} \otimes \mathbb{1}_1\tau_{012}\right] \\
&= \sup_{\substack{\tau\geq 0\\ \mathrm{Tr}[\tau]=1}} \mathrm{Tr}\left[H_z\tau_{012}\right] \\
&= \lambda_{\max}\left(H_z\right),
\end{aligned}$$

where $\lambda_{\max}(H_z)$ is the maximum eigenvalue of H_z. We are thus interested in the largest eigenvalue of H_z, denoted as $\lambda_{\max}(H_z)$, i.e.

$$\lambda_{\max}\left(H_z\right) = \sup_{\substack{\tau\geq 0\\ \mathrm{Tr}[\tau]=1}} \left[z_1 \langle\Omega|\tau_{01}|\Omega\rangle + z_2 \langle\Omega|\tau_{02}|\Omega\rangle\right] = \sup_{x\in\mathcal{C}^F} \langle z, x\rangle .$$

The set $\mathcal{C}^F$ defined in Eq. (23) may then be expressed using the notion of a polar, which is defined as follows.

Definition 3 (Polar [43, Definition 5.101])**.** *Consider a finite dimensional vector space X and its dual vector space X^*. The one-sided polar $A^\odot$ of a nonempty subset A of X, is the subset of X^* defined by*

$$A^\odot := \{x' \in X^* : \langle x, x'\rangle \leq 1 \text{ for all } x \in A\}.$$

Likewise, if B is a nonempty subset of X^*, then its one-sided polar is the subset of X defined by

$$B^\odot := \{x \in X : \langle x, x'\rangle \leq 1 \text{ for all } x' \in B\}.$$

The one-sided bipolar of a subset A of X is the set $(A^\odot)^\odot$ written simply as $A^{\odot\odot}$. The bipolar of a subset of X^* is defined in a similar manner.

Lemma B.1, which can be found in the appendix, gives some properties of the one-sided polar, in order to allow a more intuitive handling of this Definition 3. With this definition at hand, we may state the Bipolar Theorem 4.

Theorem 4 (Bipolar Theorem [43, Theorem 5.103])**.** *Consider a finite dimensional vector space X and its dual vector space X^* and let A be a nonempty subset of X. The one-sided bipolar $A^{\odot\odot}$ is the convex closed hull of $A \cup \{0\}$. Hence if A is convex, closed, and contains zero, then $A = A^{\odot\odot}$. Corresponding results hold for subsets of X^*.*

The Bipolar Theorem 4 has numerous applications in functional analysis.[b] In quantum information it always presents a very powerful tool when one wishes to fully characterize a

[b] Further information about the concept of a polar and a more general statement of the Bipolar theorem can be found in [43].

closed convex set, which is exactly what we would like to do here. The one-sided polar $\left(\mathcal{C}^F\right)^{\odot}$ of the non-empty set $\mathcal{C}^F \subseteq \mathbb{R}^2$, defined in Eq. (23) is therefore given as

$$\begin{aligned}\left(\mathcal{C}^F\right)^{\odot} &= \left\{z \in \mathbb{R}^2 \middle| \forall\ x \in \mathcal{C}^F : \langle z, x\rangle \leq 1\right\} \\ &= \left\{z \in \mathbb{R}^2 \middle| \sup_{x \in \mathcal{C}^F} \langle z, x\rangle \leq 1\right\} \\ &= \left\{z \in \mathbb{R}^2 \middle| \lambda_{\max}\left(H_z\right) \leq 1\right\}.\end{aligned}$$

Since the one-sided bipolar $\left(\mathcal{C}^F\right)^{\odot\odot}$ is just the one-sided polar of the one-sided polar, we get

$$\begin{aligned}\left(\mathcal{C}^F\right)^{\odot\odot} &= \left\{x \in \mathbb{R}^2 \middle| \forall\ z \in \left(\mathcal{C}^F\right)^{\odot} : \langle x, z\rangle \leq 1\right\} \\ &= \left\{x \in \mathbb{R}^2 \middle| \forall\ z \in \mathbb{R}^2 :\ \text{if } \lambda_{\max}\left(H_z\right) \leq 1 \text{ then } \langle x, z\rangle \leq 1\right\} \\ &= \left\{x \in \mathbb{R}^2 \middle| \forall\ z \in \mathbb{R}^2 : \left\langle x, \frac{z}{\lambda_{\max}\left(H_z\right)}\right\rangle \leq 1\right\} \\ &= \left\{x \in \mathbb{R}^2 \middle| \sup_{z \in \mathbb{R}^2} \left\langle x, \frac{z}{\lambda_{\max}\left(H_z\right)}\right\rangle \leq 1\right\}.\end{aligned}$$

In order to analyze this even further, one may now realize that every vector $\left(\begin{smallmatrix} z_1 \\ z_2 \end{smallmatrix}\right) \in \mathbb{R}^2$ can be written as $b\left(\begin{smallmatrix} v \\ 1-v \end{smallmatrix}\right)$, with $b \in \mathbb{R}_+$ and $v \in \mathbb{R}$, if $z_1 + z_2 > 0$, or $b\left(\begin{smallmatrix} -v \\ v-1 \end{smallmatrix}\right)$, with $b \in \mathbb{R}_+$ and $v \in \mathbb{R}$, if $z_1 + z_2 < 0$, or $\left(\begin{smallmatrix} \pm v \\ \mp v \end{smallmatrix}\right)$, with $v \in \mathbb{R}$, if $z_1 + z_2 = 0$. This is helpful, because $b \in \mathbb{R}_+$ and $\lambda_{\max}\left(H_{bz}\right) = b\lambda_{\max}\left(H_z\right)$. Now differentiating these three cases, the one-sided bipolar is

$$\begin{aligned}\left(\mathcal{C}^F\right)^{\odot\odot} = \Bigg\{ x \in \mathbb{R}^2 \Bigg| & \sup_{v \in \mathbb{R}^2} \frac{\left\langle x, \begin{pmatrix} v \\ 1-v \end{pmatrix}\right\rangle}{\lambda_{\max}\left(H_{\left(\begin{smallmatrix} v \\ 1-v \end{smallmatrix}\right)}\right)} \leq 1 \\ & \wedge \sup_{v \in \mathbb{R}^2} \frac{\left\langle x, \begin{pmatrix} -v \\ v-1 \end{pmatrix}\right\rangle}{\lambda_{\max}\left(H_{\left(\begin{smallmatrix} -v \\ v-1 \end{smallmatrix}\right)}\right)} \leq 1 \\ & \wedge \sup_{v \in \mathbb{R}^2} \frac{\left\langle x, \begin{pmatrix} \pm v \\ \mp v \end{pmatrix}\right\rangle}{\lambda_{\max}\left(H_{\left(\begin{smallmatrix} \pm v \\ \mp v \end{smallmatrix}\right)}\right)} \leq 1 \Bigg\}.\end{aligned}$$

Note that by analyzing the rank of H_z we always expect an eigenvalue equal to zero. Therefore, the maximum eigenvalue must always be non-negative, i.e. $\lambda_{\max}\left(H_z\right) \geq 0$. It turns out that

$$\begin{aligned}\lambda_{\max}\left(H_{\left(\begin{smallmatrix} v \\ 1-v \end{smallmatrix}\right)}\right) &= \frac{1}{2d}\left(d + \sqrt{d^2 + 4(d^2-1)(v-1)v}\right), \\ \lambda_{\max}\left(H_{\left(\begin{smallmatrix} -v \\ v-1 \end{smallmatrix}\right)}\right) &= \begin{cases} 0 & \text{if } 0 \geq v \geq 1, \\ \frac{1}{2d}\left(-d + \sqrt{d^2 + 4(d^2-1)(v-1)v}\right) & \text{otherwise,} \end{cases} \\ \lambda_{\max}\left(H_{\left(\begin{smallmatrix} \pm v \\ \mp v \end{smallmatrix}\right)}\right) &= v\sqrt{\frac{d^2-1}{d^2}}.\end{aligned}$$

Proposition 2 (Convexity)**.** *The set*

$$\mathcal{C}^F = \left\{ z \in \mathbb{R}^2 \middle| z = \begin{pmatrix} \langle\Omega|\,\tau_{01}\,|\Omega\rangle \\ \langle\Omega|\,\tau_{02}\,|\Omega\rangle \end{pmatrix} \right\}$$

is convex.

Proof. Let $z^A, z^B \in \mathcal{C}$, then for $\lambda \in [0,1]$, $z^C = \lambda z^A + (1-\lambda) z^B \in \mathcal{C}^F$, because

$$\begin{aligned} z^C &= \lambda z^A + (1-\lambda) z^B \\ &= \lambda \begin{pmatrix} \langle\Omega|\,\tau_{01}^A\,|\Omega\rangle \\ \langle\Omega|\,\tau_{02}^A\,|\Omega\rangle \end{pmatrix} + (1-\lambda) \begin{pmatrix} \langle\Omega|\,\tau_{01}^B\,|\Omega\rangle \\ \langle\Omega|\,\tau_{02}^B\,|\Omega\rangle \end{pmatrix} \\ &= \begin{pmatrix} \langle\Omega|\,\tau_{01}^C\,|\Omega\rangle \\ \langle\Omega|\,\tau_{02}^C\,|\Omega\rangle \end{pmatrix} \in \mathcal{C}^F \end{aligned}$$

□.

By using the Bipolar Theorem 4 together with the fact that $\mathcal{C}^F$ is convex, as shown in Proposition 2, closed and contains the origin, we see that $\mathcal{C}^F = \left(\mathcal{C}^F\right)^{\odot\odot}$. A cumbersome computation then shows that the boundary of this set is described by

$$\frac{1}{d+1}\left(\sqrt{\langle\Omega|\,\tau_{01}\,|\Omega\rangle} + \sqrt{\langle\Omega|\,\tau_{02}\,|\Omega\rangle}\right)^2 + \frac{1}{d-1}\left(\sqrt{\langle\Omega|\,\tau_{01}\,|\Omega\rangle} - \sqrt{\langle\Omega|\,\tau_{02}\,|\Omega\rangle}\right)^2 = \frac{2}{d}, \quad (24)$$

which is illustrated in Figure B.1. We may therefore state the following Theorem 5, summarizing the main result.

Theorem 5 (Set of all attainable single clone fidelities within universal $1 \to 2$ asymmetric quantum cloning)**.** *The set of all attainable clone qualities in terms of single clone fidelities* $d^F(T_i, \mathrm{id}) = \langle\Omega|\,\tau_{0i}\,|\Omega\rangle$ *with* $i = 1, 2$*, where* $\tau_{0i} := \mathrm{id} \otimes \tilde{T}_i\left(|\Omega\rangle\langle\Omega|\right)$ *is the Choi-Jamiolkowski state of the marginals of the optimal quantum cloning channel, given by Eq. (21), with* $|\Omega\rangle\langle\Omega| := \frac{1}{d}\sum_{i,j=1}^{d} |ii\rangle\langle jj|$ *being the maximally entangled state, is given by*

$$\begin{aligned} \mathcal{C}^F = \Bigg\{ x \in \mathbb{R}^2 \Bigg| & \sup_{v\in\mathbb{R}^2} \frac{\left\langle x, \begin{pmatrix} v \\ 1-v \end{pmatrix} \right\rangle}{\lambda_{max}\left(H_{\binom{v}{1-v}}\right)} \leq 1 \\ & \wedge \sup_{v\in\mathbb{R}^2} \frac{\left\langle x, \begin{pmatrix} -v \\ v-1 \end{pmatrix} \right\rangle}{\lambda_{max}\left(H_{\binom{-v}{v-1}}\right)} \leq 1 \\ & \wedge \sup_{v\in\mathbb{R}^2} \frac{\left\langle x, \begin{pmatrix} \pm v \\ \mp v \end{pmatrix} \right\rangle}{\lambda_{max}\left(H_{\binom{\pm v}{\mp v}}\right)} \leq 1 \Bigg\}, \end{aligned} \quad (25a)$$

where $\lambda_{max}(H_z)$ *denotes the maximum eigenvalue of*

$$H_z = z_1\,|\Omega\rangle\langle\Omega|_{01} \otimes \mathbb{1}_2 + z_2\,|\Omega\rangle\langle\Omega|_{02} \otimes \mathbb{1}_1,$$

given by

$$\lambda_{max}\left(H_{\binom{v}{1-v}}\right) = \frac{1}{2d}\left(d+\sqrt{d^2+4(d^2-1)(v-1)v}\right),$$
$$\lambda_{max}\left(H_{\binom{-v}{v-1}}\right) = \begin{cases} 0 & \text{if } 0 \geq v \geq 1, \\ \frac{1}{2d}\left(-d+\sqrt{d^2+4(d^2-1)(v-1)v}\right) & \text{otherwise,} \end{cases}$$
$$\lambda_{max}\left(H_{\binom{\pm v}{\mp v}}\right) = v\sqrt{\frac{d^2-1}{d^2}}.$$

The upper boundary of this set is described by

$$\frac{1}{d+1}\left(\sqrt{\langle\Omega|\,\tau_{01}\,|\Omega\rangle}+\sqrt{\langle\Omega|\,\tau_{02}\,|\Omega\rangle}\right)^2+\frac{1}{d-1}\left(\sqrt{\langle\Omega|\,\tau_{01}\,|\Omega\rangle}-\sqrt{\langle\Omega|\,\tau_{02}\,|\Omega\rangle}\right)^2=\frac{2}{d}, \quad (25b)$$

and illustrated in Figure B.1, which can be found in the appendix.

This theorem is in agreement with previously established results [26, 28]. Here, the authors have used a group theoretic approach, whereas our main technique comes from convex analysis.

A similar theorem may be stated for different figures of merit, since the set of attainable single clone qualities is convex for any $d(T_i, \mathrm{id})$, satisfying the properties discussed earlier, as shown in the following proposition.

Proposition 3 (Convexity)**.** *The set*

$$\mathcal{C} = \left\{ z \in \mathbb{R}^2 \middle| z = \begin{pmatrix} d\left(T_1, \mathrm{id}\right) \\ d\left(T_2, \mathrm{id}\right) \end{pmatrix} \right\}$$

is convex.

Proof. Let $z^A, z^B \in \mathcal{C}$, then for $\lambda \in [0,1]$, we get

$$\begin{aligned} &\lambda \begin{pmatrix} d\left(T_1^A, \mathrm{id}\right) \\ d\left(T_2^A, \mathrm{id}\right) \end{pmatrix} + (1-\lambda) \begin{pmatrix} d\left(T_1^B, \mathrm{id}\right) \\ d\left(T_2^B, \mathrm{id}\right) \end{pmatrix} \\ &= \begin{pmatrix} \lambda d\left(T_1^A, \mathrm{id}\right) + (1-\lambda)\, d\left(T_1^B, \mathrm{id}\right) \\ \lambda d\left(T_2^A, \mathrm{id}\right) + (1-\lambda)\, d\left(T_2^B, \mathrm{id}\right) \end{pmatrix} \\ &\leq \begin{pmatrix} d\left(\lambda T_1^A + (1-\lambda)\, T_1^B, \mathrm{id}\right) \\ d\left(\lambda T_2^A + (1-\lambda)\, T_2^B, \mathrm{id}\right) \end{pmatrix} \\ &= \begin{pmatrix} d\left(T_1^C, \mathrm{id}\right) \\ d\left(T_2^C, \mathrm{id}\right) \end{pmatrix} \in \mathcal{C}. \end{aligned}$$

In the case, where the figure of merit is given by the fidelity, we even get equality. In all other cases for the lower boundary consider the following quantum channel,

$$T\left(\rho\right) = |\psi\rangle\langle\psi| \otimes \left(\lambda\rho + (1-\lambda)\,|\psi\rangle\langle\psi|\right),$$

such that

$$\begin{aligned} T_i\left(\rho\right) &= |\psi\rangle\langle\psi| \text{ and} \\ T_{\bar{i}}(\rho) &= \lambda\rho + (1-\lambda)\,|\psi\rangle\langle\psi|\,, \end{aligned}$$

for $i = 1, 2$ with $\lambda \in [0,1]$ and $|\psi\rangle \in \mathcal{D}_d$ some pure quantum state $\square$.

This immediately gives rise to the following corollary.

Corollary 2 (Set of all attainable single quantum clone qualities within universal $1 \to 2$ asymmetric quantum cloning using different figures of merit)**.** *The set of all attainable single quantum clone qualities in terms of the different figures of merit $d^k(T_i, \mathrm{id})$ with $i = 1, 2$ for $k = F, 1, 2, \infty, \diamond$, is given by*

$$\mathcal{C}^k = conv\left(\{0\} \cup \left\{x_{max}^{(k)}\right\} \cup \left\{x^{(k)} \middle| g\left(f^k\left(x_1^{(k)}\right), f^k\left(x_2^{(k)}\right)\right) = 0\right\}\right), \tag{26a}$$

where $\{0\}$ is the origin, $\left\{x_{max}^{(k)}\right\}$ are the two points where $d^k(T_i, \mathrm{id})$ reaches its maximum for $i = 1, 2$ and where the function $g : \mathbb{R}^2 \to \mathbb{R}$ is

$$g(x_1, x_2) = \frac{1}{d+1}\left(\sqrt{x_1} + \sqrt{x_2}\right)^2 + \frac{1}{d-1}\left(\sqrt{x_1} - \sqrt{x_2}\right)^2 - \frac{2}{d}, \tag{26b}$$

with the functions $f^k : \mathbb{R} \to \mathbb{R}$ specified by

$$f^F\left(x_i^{(F)}\right) = x_i^{(F)}, \tag{26c}$$

$$f^1\left(x_i^{(1)}\right) = 1 + \frac{1+d}{d}\left(x_i^{(1)} - 1\right), \tag{26d}$$

$$f^2\left(x_i^{(2)}\right) = 1 + \frac{d^2-1}{d^2}\sqrt{\frac{d}{d-1}}\left(x_i^{(2)} - 1\right), \tag{26e}$$

$$f^\infty\left(x_i^{(\infty)}\right) = 1 + \frac{1+d}{d}\left(x_i^{(\infty)} - 1\right) \text{ and} \tag{26f}$$

$$f^\diamond\left(x_i^{(\diamond)}\right) = x_i^{(\diamond)}. \tag{26g}$$

The sets are depicted in Figure B.1 up to Figure B.5 in the appendix.

7 Summary

This paper revisits the universal asymmetric $1 \to 2$ quantum cloning problem. We derived the optimal universal $1 \to 2$ asymmetric quantum cloning channel using its symmetry properties in Theorem 3. Additionally, we noticed that its inherent optimization problem can be recast as a semidefinite program. This result has been derived previously by [21, 22].

Furthermore, we completely characterize the set of all attainable single quantum clone qualities within universal asymmetric $1 \to 2$ quantum cloning for different figures of merit in Theorem 5 and Corollary 2 using the concept of a one-sided polar together with the famous Bipolar Theorem 4 from convex analysis. This is an alternative approach to the one chosen in [28], where the authors use a general group representation approach and only study the fidelity.

Acknowledgements

I would like to thank Michael M. Wolf for his constant support and insightful discussions as well as René Schwonnek and David Reeb for all their help. Furthermore, I would like to thank Sabina Alazzawi and Daniel Stilck França for all the useful comments.

This work is supported by the Elite Network of Bavaria through the PhD programme of excellence *Exploring Quantum Matter*.

References

1. W. K. Wootters and W. H. Zurek (1982), *A single quantum cannot be cloned*, Nature, vol. 299, pp. 802–803.
2. M. Keyl (2003), *Aspects of quantum information theory*, Habilitation thesis, Technische Universität Carolo-Wilhelmina zu Braunschweig.
3. A. Lamas-Linares, C. Simon, J. C. Howell and D. Bouwmeester (2002), *Experimental Quantum Cloning of Single Photons*, Science, vol. 296, pp. 712–714.
4. H. K. Cummins, C. Jones, A. Furze, N. F. Soffe, M. Mosca, J. M. Peach and J. A. Jones (2002), *Approximate Quantum Cloning with Nuclear Magnetic Resonance*, Phys. Rev. Lett., vol. 88, p. 187901.
5. Z. Zhao, A. Zhang, X. Zhou, Y. Chen, C. Lu, A. Karlsson and J. Pan (2005), *Experimental Realization of Optimal Asymmetric Cloning and Telecloning via Partial Teleportation*, Phys. Rev. Lett., vol. 95, p. 030502.
6. E. Nagali, D. Giovannini, L. Marrucci, S. Slussarenko, E. Santamato and F. Sciarrino (2010), *Experimental Optimal Cloning of Four-Dimensional Quantum States of Photons*, Phys. Rev. Lett., vol. 105, p. 073602.
7. N. J. Cerf and J. Fiurášek (2006), *Optical quantum cloning*, in E. Wolf, editor, *Progress in Optics*, vol. 49, pp. 455 – 545, Elsevier.
8. V. Scarani, S. Iblisdir, N. Gisin and A. Acín (2005), *Quantum cloning*, Rev. Mod. Phys., vol. 77, pp. 1225–1256.
9. V. Bužek and M. Hillery (1996), *Quantum copying: Beyond the no-cloning theorem*, Phys. Rev. A, vol. 54, p. 1844.
10. M. Hillery and V. Bužek (1997), *Quantum copying: Fundamental inequalities*, Phys. Rev. A, vol. 56, pp. 1212–1216.
11. D. Bruß, D. P. DiVincenzo, A. Ekert, C. A. Fuchs, C. Macchiavello and J. A. Smolin (1998), *Optimal universal and state-dependent quantum cloning*, Phys. Rev. A, vol. 57, pp. 2368–2378.
12. N. Gisin and S. Massar (1997), *Optimal Quantum Cloning Machines*, Phys. Rev. Lett., vol. 79, pp. 2153–2156.
13. D. Bruß, A. Ekert and C. Macchiavello (1998), *Optimal Universal Quantum Cloning and State Estimation*, Phys. Rev. Lett., vol. 81, pp. 2598–2601.
14. V. Bužek and M. Hillery (1998), *Universal Optimal Cloning of Arbitrary Quantum States: From Qubits to Quantum Registers*, Phys. Rev. Lett., vol. 81, pp. 5003–5006.
15. N. J. Cerf (1998), *Asymmetric Quantum Cloning Machines*, Acta Phys. Slovaca, vol. 48, pp. 115–132.
16. R. F. Werner (1998), *Optimal cloning of pure states*, Phys. Rev. A, vol. 58, pp. 1827–1832.
17. M. Keyl and R. F. Werner (1999), *Optimal cloning of pure states, testing single clones*, J. Math. Phys., vol. 40, pp. 3283–3299.
18. V. Bužek, M. Hillery and R. Bednik (1998), *Controlling the flow of information in quantum cloners: Asymmetric cloning*, Acta Phys. Slovaca, vol. 48, p. 177.
19. C. Niu and R. B. Griffiths (1998), *Optimal copying of one quantum bit*, Phys. Rev. A, vol. 58, pp. 4377–4393.
20. N. J. Cerf (2000), *Pauli cloning of a quantum bit*, Phys. Rev. Lett., vol. 84, pp. 4497–4500.
21. S. L. Braunstein, V. Bužek and M. Hillery (2001), *Quantum-information distributors: Quantum network for symmetric and asymmetric cloning in arbitrary dimension and continuous limit*, Phys. Rev. A, vol. 63, p. 052313.
22. N. J. Cerf (2000), *Asymmetric quantum cloning in any dimension*, J. Mod. Opt., vol. 47, pp. 187–209.

23. S. Iblisdir, A. Acín, N. J. Cerf, R. Filip, J. Fiurášek and N. Gisin (2005), *Multipartite asymmetric quantum cloning*, Phys. Rev. A, vol. 72, p. 042328.
24. S. Iblisdir, A. Acín and N. Gisin (2006), *Generalised Asymmetric Quantum Cloning Machines*, Quant. Inform. Comp., vol. 6, pp. 410–435.
25. J. Fiurášek, R. Filip and N. J. Cerf (2005), *Highly asymmetric quantum cloning in arbitrary dimension*, Quant. Inform. Comp., vol. 5, pp. 583–592.
26. M. Jiang and S. Yu (2010), *Extremal asymmetric universal cloning machines*, J. Math. Phys., vol. 51, 052306.
27. P. Ćwikliński, M. Horodecki and M. Studziński (2012), *Region of fidelities for a* $1 \to N$ *universal qubit quantum cloner*, Phys. Lett. A, vol. 376, pp. 2178–2187.
28. M. Studziński, P. Ćwikliński, M. Horodecki and M. Mozrzymas (2014), *Group-representation approach to* $1 \to N$ *universal quantum cloning machines*, Phys. Rev. A, vol. 89, p. 052322.
29. A. Kay, D. Kaszlikowski and R. Ramanathan (2009), *Optimal Cloning and Singlet Monogamy*, Phys. Rev. Lett., vol. 103, p. 050501.
30. A. Kay, R. Ramanathan and D. Kaszlikowshi (2013), *Optimal Asymmetric Quantum Cloning for Quantum Information and Computation*, Quant. Inform. Comp., vol. 13, pp. 880–900.
31. A. Kay (2016), *Optimal Universal Quantum Cloning: Asymmetries and Fidelity Measures*, Quant. Inform. Comp., vol. 16, p. 991.
32. T. Eggeling and R. F. Werner (2001), *Separability properties of tripartite states with* $U \otimes U \otimes U$*-symmetry*, Phys. Rev. A, vol. 63, p. 042111.
33. K. G. H. Vollbrecht and R. F. Werner (2001), *Entanglement measures under symmetry*, Phys. Rev. A, vol. 64, p. 062307.
34. H. Weyl (1997), *The Classical Groups: Their Invariants and Representations*, Princeton landmarks in mathematics and physics, Princeton University Press.
35. B. Simon (1996), *Representations of Finite and Compact Groups*, *Graduate studies in mathematics*, vol. 10, American Mathematical Society.
36. T. Heinosaari and M. Ziman (2012), *The Mathematical Language of Quantum Theory: From Uncertainty to Entanglement*, Cambridge University Press.
37. D. A. Lidar and T. A. Brun (2013), *Quantum error correction*, Cambridge University Press.
38. O. Bratteli and D. Robinson (2002), *Operator Algebras and Quantum Statistical Mechanics 1: C*- and W*-Algebras. Symmetry Groups. Decomposition of States.*, Operator Algebras and Quantum Statistical Mechanics, Springer-Verlag Berlin Heidelberg New York, 2nd edn.
39. O. Bratteli and D. Robinson (2002), *Operator Algebras and Quantum Statistical Mechanics: Equilibrium States. Models in Quantum Statistical Mechanics.*, Operator Algebras and Quantum Statistical Mechanics, Springer-Verlag Berlin Heidelberg New York, 2nd edn.
40. L. Vandenberghe and S. Boyd (1994), *Semidefinite Programming*, SIAM REVIEW, vol. 38, pp. 49–95.
41. I. CVX Research (2012), *CVX: Matlab Software for Disciplined Convex Programming, version 2.0*, `http://cvxr.com/cvx`.
42. M. Grant and S. Boyd (2008), *Graph implementations for nonsmooth convex programs*, in V. Blondel, S. Boyd and H. Kimura, editors, *Recent Advances in Learning and Control*, Lecture Notes in Control and Information Sciences, pp. 95–110, Springer-Verlag Limited.
43. C. D. Aliprantis and K. C. Border (2007), *Infinite Dimensional Analysis: A Hitchhiker's Guide*, Springer Berlin Heidelberg, 3rd edn.
44. M. Keyl (2002), *Fundamentals of quantum information theory*, Phys. Rep., vol. 369, pp. 431–548.
45. R. T. Rockafellar (1996), *Convex Analysis*, Princeton Landmarks in Mathematics and Physics, Princeton University Press.
46. The MathWorks, Inc. (2014), *MATLAB and Statistics Toolbox Release R2014b*, Natick, Massachusetts, United States.

Appendix A Proofs

Appendix A.1. *Proof of Lemma 3*

Lemma 3. *For a Choi-Jamiolkowski state* $\tau \in \mathcal{M}_{d^3}$,

$$\left[\tau, \bar{U} \otimes U \otimes U\right] = 0 \ \ \forall\ U \in \mathcal{U}\left(d\right)$$

is equivalent to

$$\int_{\mathcal{U}(d)} \left(\bar{U} \otimes U \otimes U\right) \tau \left(\bar{U} \otimes U \otimes U\right)^* \, \mathrm{d}U = \tau.$$

Proof. If

$$\left[\tau, \bar{U} \otimes U \otimes U\right] = 0$$

then

$$\begin{aligned}
&\int_{\mathcal{U}(d)} \left(\bar{U} \otimes U \otimes U\right) \tau \left(\bar{U} \otimes U \otimes U\right)^* \, \mathrm{d}U \\
=&\int_{\mathcal{U}(d)} \tau \left(\bar{U} \otimes U \otimes U\right) \left(\bar{U} \otimes U \otimes U\right)^* \, \mathrm{d}U \\
=&\int_{\mathcal{U}(d)} \tau \, \mathrm{d}U \\
=&\tau.
\end{aligned}$$

The other direction follows from

$$\int_{\mathcal{U}(d)} \left(\bar{U} \otimes U \otimes U\right) \tau \left(\bar{U} \otimes U \otimes U\right)^* \, \mathrm{d}U = \tau,$$

because then for unitary $V \in \mathcal{U}\left(d\right)$ we have

$$\begin{aligned}
&\tau \left(\bar{V} \otimes V \otimes V\right) \\
=&\int_{\mathcal{U}(d)} \left(\bar{U} \otimes U \otimes U\right) \tau \left(\bar{U} \otimes U \otimes U\right)^* \, \mathrm{d}U \left(\bar{V} \otimes V \otimes V\right) \\
=&\left(\bar{V} \otimes V \otimes V\right) \left(\bar{V} \otimes V \otimes V\right)^* \\
&\qquad \int_{\mathcal{U}(d)} \left(\bar{U} \otimes U \otimes U\right) \tau \left(\bar{U} \otimes U \otimes U\right)^* \, \mathrm{d}U \left(\bar{V} \otimes V \otimes V\right) \\
=&\left(\bar{V} \otimes V \otimes V\right) \int_{\mathcal{U}(d)} \left(\bar{U} \otimes U \otimes U\right) \tau \left(\bar{U} \otimes U \otimes U\right)^* \, \mathrm{d}U \\
=&\left(\bar{V} \otimes V \otimes V\right) \tau,
\end{aligned}$$

where we have used the invariance property □.

Appendix A.2. *Proof of Proposition 1*

Proposition 1. *If for a Choi-Jamiolkowski state* $\tau \in \mathcal{M}_{d^3}$ *we have that*

$$\left[\tau, \bar{U} \otimes U \otimes U\right] = 0 \ \ \forall\ U \in \mathcal{U}\left(d\right),$$

then

$$\left[\tau^{t_0}, U \otimes U \otimes U\right] = 0,$$

where t_0 denotes the partial transpose on the first system.

Proof.

$$\begin{aligned}
&\int_{\mathcal{U}(d)} (U \otimes U \otimes U)\, \tau^{t_0}\, (U \otimes U \otimes U)^*\, \mathrm{d}U \\
=&\int_{\mathcal{U}(d)} (U \otimes U \otimes U) \left(\mathrm{id} \otimes \tilde{T}\, |\Omega\rangle\langle\Omega|\right)^{t_0} (U \otimes U \otimes U)^*\, \mathrm{d}U \\
=&\int_{\mathcal{U}(d)} \left[\left((U^*)^T \otimes U \otimes U\right) \mathrm{id} \otimes \tilde{T}\, |\Omega\rangle\langle\Omega|\, \left(U^T \otimes U^* \otimes U^*\right)\right]^{t_0} \mathrm{d}U \\
=&\int_{\mathcal{U}(d)} \left[\left(\bar{U} \otimes U \otimes U\right) \mathrm{id} \otimes \tilde{T}\, |\Omega\rangle\langle\Omega|\, \left(\bar{U} \otimes U \otimes U\right)^*\right]^{t_0} \mathrm{d}U
\end{aligned}$$

remembering that $U \otimes \bar{U}\, |\Omega\rangle = |\Omega\rangle$ gives

$$\begin{aligned}
=&\int_{\mathcal{U}(d)} \left[(\mathbb{1} \otimes U \otimes U) \,\mathrm{id} \otimes \tilde{T}\, (\mathbb{1} \otimes U^*)\, |\Omega\rangle\langle\Omega|\, (\mathbb{1} \otimes U)\, (\mathbb{1} \otimes U^* \otimes U^*)\right]^{t_0} \mathrm{d}U \\
=&\int_{\mathcal{U}(d)} \tau^{t_0}\, \mathrm{d}U \\
=&\tau^{t_0}.
\end{aligned}$$

Application of Lemma 3 finishes the proof □.

Appendix B Appendix and Figures

Relation between different notation used in Chapter 5.

$$\begin{aligned}
s_0 &= a_1 2d + a_2 d^2 + a_3 d^2 + a_5 d + a_6 d \\
s_1 &= a_2 d + a_3 d + a_4 2d + a_5 d^2 + a_6 d^2 \\
s_2 &= a_2 d\sqrt{d^2-1} + a_3(-d)\sqrt{d^2-1} \\
s_3 &= a_5(-id)\sqrt{d^2-1} + a_6 id\sqrt{d^2-1} \\
s_+ &= a_1 \frac{1}{2} d(d+2)(d-1) + a_4 \frac{1}{2} d(d+2)(d-1) \\
s_- &= a_1 \frac{1}{2} d(d-2)(d+1) + a_4 \frac{1}{2} d(-d+2)(d+1)
\end{aligned}$$

Lemma B.1 (Properties of polars [43, Lemma 5.102])**.** *Consider a finite dimensional vector space X and its dual vector space X^*. Let A, B be nonempty subsets of X and let $\{A_i\}$ be a family of nonempty subsets of X. Then the following properties hold:*

1. *If $A \subset B$, then $A^\odot \supset A^\odot$.*
2. *If $\varepsilon \neq 0$, then $(\varepsilon A)^\odot = \frac{1}{\varepsilon} A^\odot$.*

3. $\cap \left(A_i^{\odot}\right) = (\cup A_i)^{\odot}$.

4. The one-sided polar $A^{\odot}$ *is nonempty, convex, closed and contains the origin.*

The corresponding dual statements are true for subsets of X^*.

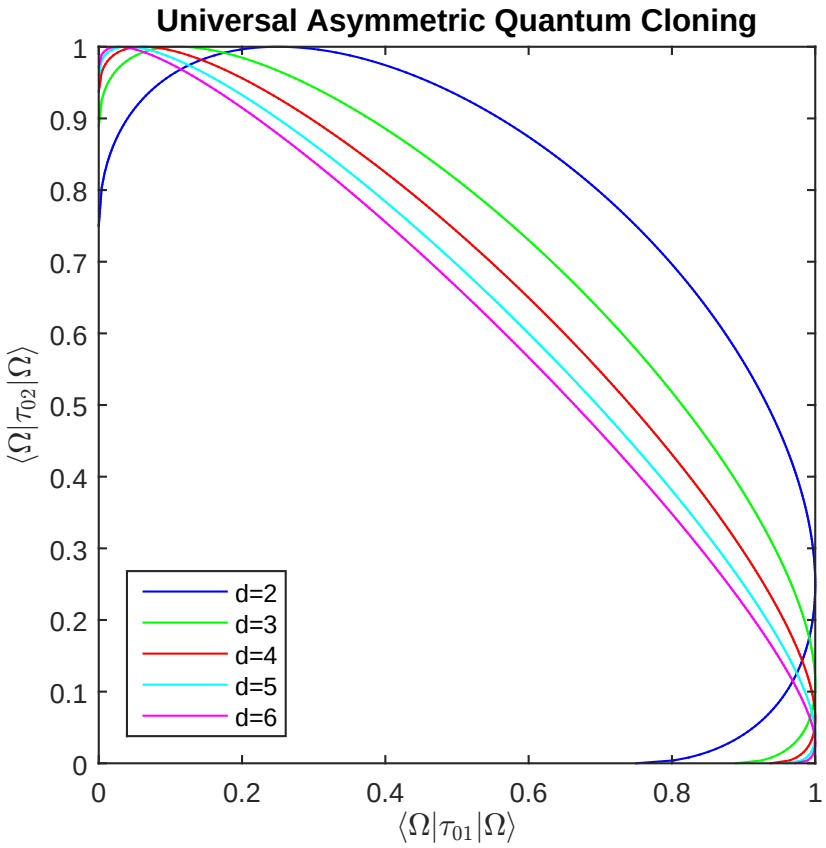


Fig. B.1. The set of all attainable single quantum clone qualities in terms of $d^F(T_i, \mathrm{id}) = \langle\Omega|\,\tau_{0i}\,|\Omega\rangle$, $i = 1, 2$, given by Eq. (25a) for different dimensions of the underlying Hilbert space, using MATLAB [46]. The upper boundary of this set is given by $\frac{1}{d+1}\left(\sqrt{x_1^{(F)}} + \sqrt{x_2^{(F)}}\right)^2 + \frac{1}{d-1}\left(\sqrt{x_1^{(F)}} - \sqrt{x_2^{(F)}}\right)^2 = \frac{2}{d}$.

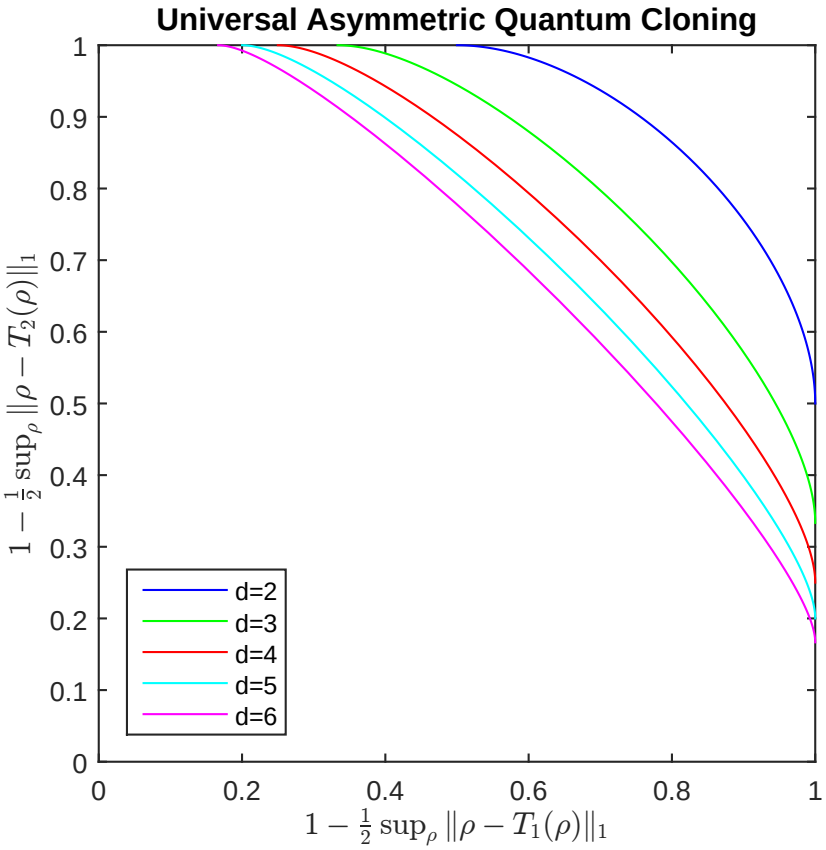


Fig. B.2. The set of all attainable single quantum clone qualities in terms of $d^1(T_i, \mathrm{id}) = 1 - \frac{1}{2}\sup_\rho \|T_i(\rho) - \rho\|_1$, $i = 1, 2$, given by Eq. (26a) for different dimensions of the underlying Hilbert space, using MATLAB [46]. The upper boundary of this set is given by

$$\frac{1}{d+1}\left(\sqrt{1+\frac{1+d}{d}\left(x_1^{(1)}-1\right)}+\sqrt{1+\frac{1+d}{d}\left(x_2^{(1)}-1\right)}\right)^2 + \frac{1}{d-1}\left(\sqrt{1+\frac{1+d}{d}\left(x_1^{(1)}-1\right)}-\sqrt{1+\frac{1+d}{d}\left(x_2^{(1)}-1\right)}\right)^2 = \frac{2}{d}.$$

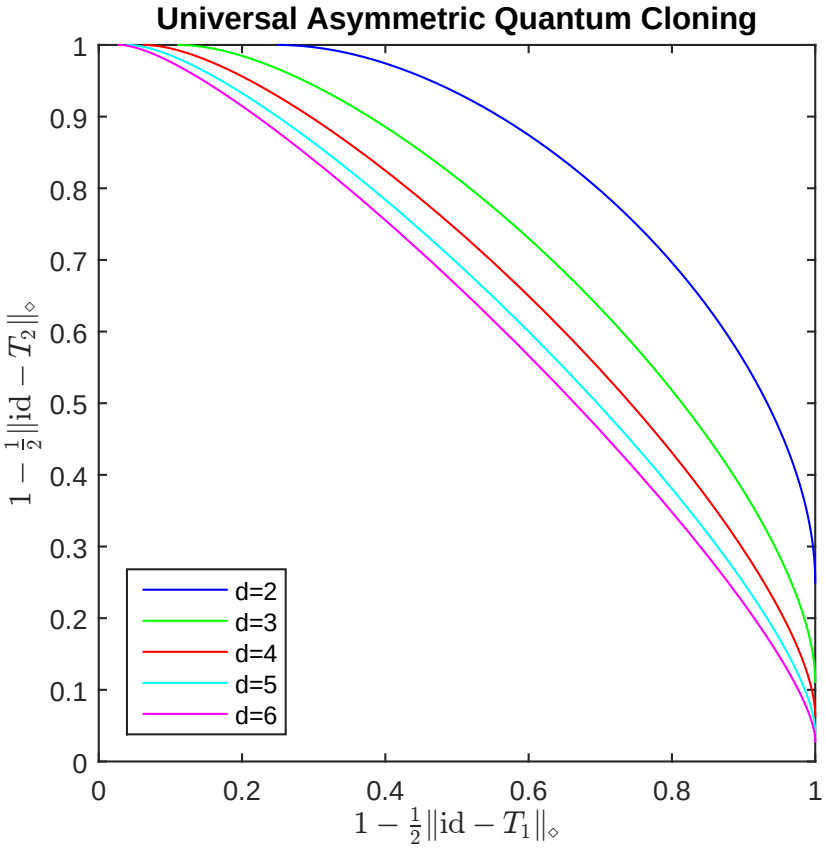


Fig. B.3. The set of all attainable single quantum clone qualities in terms of $d^\diamond(T_i, \mathrm{id}) = 1 - \frac{1}{2}\|T_i - \mathrm{id}\|_\diamond$, $i = 1, 2$, given by Eq. (26a) for different dimensions of the underlying Hilbert space, using MATLAB [46]. The upper boundary of this set is given by $\frac{1}{d+1}\left(\sqrt{x_1^{(\diamond)}} + \sqrt{x_2^{(\diamond)}}\right)^2 + \frac{1}{d-1}\left(\sqrt{x_1^{(\diamond)}} - \sqrt{x_2^{(\diamond)}}\right)^2 = \frac{2}{d}$.

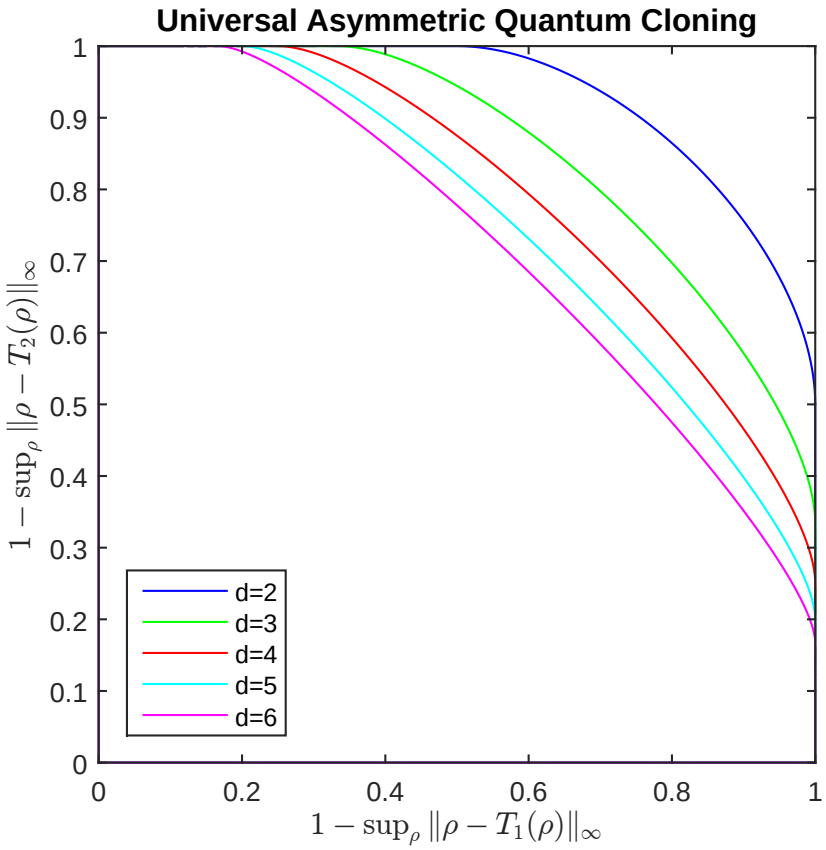


Fig. B.4. The set of all attainable single quantum clone qualities in terms of $d^\infty(T_i, \mathrm{id}) = 1 - \sup_\rho \|T_i(\rho) - \rho\|_\infty$, $i = 1, 2$, given by Eq. (26a) for different dimensions of the underlying Hilbert space, using MATLAB [46]. The upper boundary of this set is given by

$$\frac{1}{d+1}\left(\sqrt{1+\frac{1+d}{d}\left(x_1^{(\infty)}-1\right)}+\sqrt{1+\frac{1+d}{d}\left(x_2^{(\infty)}-1\right)}\right)^2 + \frac{1}{d-1}\left(\sqrt{1+\frac{1+d}{d}\left(x_1^{(\infty)}-1\right)}-\sqrt{1+\frac{1+d}{d}\left(x_2^{(\infty)}-1\right)}\right)^2 = \frac{2}{d}.$$

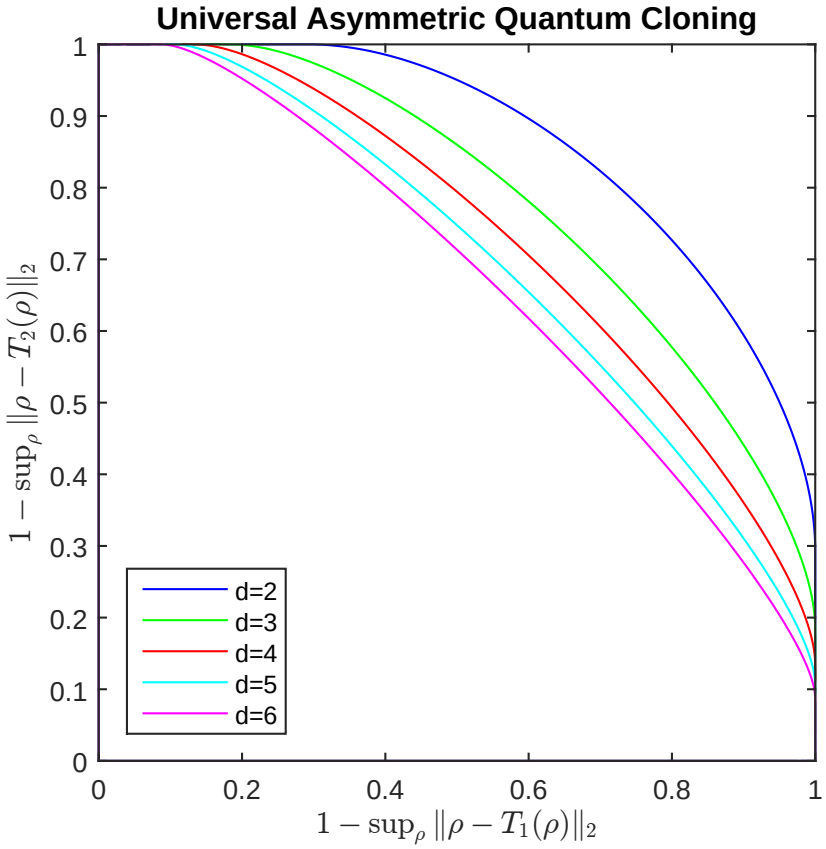


Fig. B.5. The set of all attainable single quantum clone qualities in terms of $d^2(T_i, \mathrm{id}) = 1 - \sup_\rho \|T_i(\rho) - \rho\|_2$, $i = 1, 2$, given by Eq. (26a) for different dimensions of the underlying Hilbert space, using MATLAB [46]. The upper boundary of this set is given by

$$\frac{1}{d+1}\left(\sqrt{1+\frac{d^2-1}{d^2}\sqrt{\frac{d}{d-1}}\left(x_1^{(2)}-1\right)}+\sqrt{1+\frac{d^2-1}{d^2}\sqrt{\frac{d}{d-1}}\left(x_2^{(2)}-1\right)}\right)^2$$
$$+\frac{1}{d-1}\left(\sqrt{1+\frac{d^2-1}{d^2}\sqrt{\frac{d}{d-1}}\left(x_1^{(2)}-1\right)}-\sqrt{1+\frac{d^2-1}{d^2}\sqrt{\frac{d}{d-1}}\left(x_2^{(2)}-1\right)}\right)^2=\frac{2}{d}.$$

B Contributed core article: Article 2

A. K. Hashagen and M. M. Wolf
Universality and optimality in the information-disturbance tradeoff
Accepted in Annales Henri Poincaré, 2018

Summary of article 2: Universality and Optimality in the Information-Disturbance Tradeoff [2]

In this article, we investigate the tradeoff between the quality of an approximate measurement and the disturbance this measurement induces to the quantum system. Even though this phenomena of necessary disturbance when extracting information from a quantum system has been known for decades (cf. proposition 2.4), its quantitative description has given rise to many debates and obscurities. Due to recent practical applications to quantum information processing tasks, such as quantum cryptography [59, 88–90], a thorough mathematical analysis is needed.

This article differs to known quantitative bounds on the disturbance derived in other papers in two respects. Firstly, the performed measurement is considered to approximate a fixed target measurement; this yields the measurement error. The disturbance, however, is not quantified with respect to another observable. This is in stark contrast to other papers, in which either the measurement error as well as the disturbance are quantified with respect to a second reference measurement [91–103] or no reference observable is used at all [104–111]. This classification is illustrated in figure B.1.

Secondly, instead of considering specific distance measures in order to derive the tradeoff, in this work we prove the existence of a two-parameter family of quantum instruments that are (almost) universally optimal as long as the distance measures exhibit a set of properties that are shared by many distance measures considered in the literature such as convexity and basis-independence. These include for example norm-based measures as used in [96, 103, 107], fidelities as used in [95, 111] or for example transport-cost functions as used in [97, 102]. The following result is proven in the case of a non-degenerate von Neumann target measurement. In this scenario it is possible to use symmetry methods to reduce the number of optimization parameters, such that an analytic solution can be obtained. It is shown that, without loss of generality, the optimum is obtained for a twirled quantum channel, which gives rise to a rich advantageous structure.

Theorem B.1 ((Almost universal) optimal instruments [2, theorem 1])**.** *Let* Δ *and* δ *be distance-measures for quantifying disturbance and measurement-error that satisfy assumpti-*

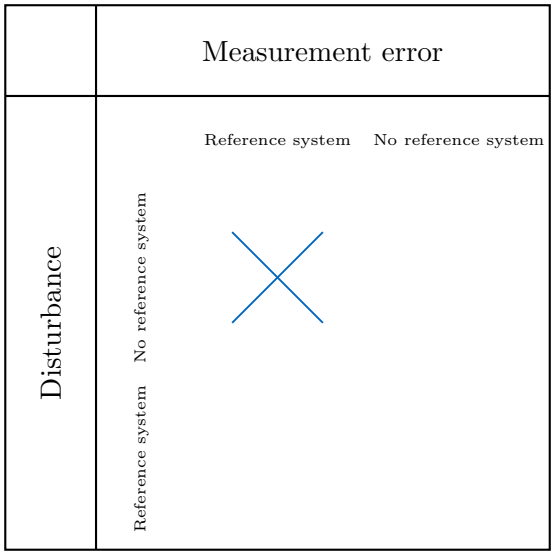


Figure B.1: Classification of information-disturbance tradeoffs w.r.t. whether a reference measurement is used to quantify the information gain and the disturbance. The blue cross shows the intermediate approach taken in [2].

ons 4.3 and 4.4 (see section 4.2), respectively. Then the optimal $\Delta - \delta$-tradeoff w.r.t. a target measurement that is given by an orthonormal basis $\{|i\rangle \in \mathbb{C}^d\}_{i=1}^d$ is attained within the two-parameter family of instruments defined by

$$I_i(\rho) := z \langle i| \rho |i\rangle \frac{\mathbb{1}_d - |i\rangle\langle i|}{d-1} + (1-z) K_i \rho K_i, \quad K_i := \mu \mathbb{1}_d + \nu |i\rangle\langle i|, \tag{B.1}$$

where $z \in [0,1]$ and $\mu, \nu \in \mathbb{R}$ satisfy $d\mu^2 + \nu^2 + 2\mu\nu = 1$ (which makes $\sum_i I_i$ trace preserving).

Furthermore, we show that for many common disturbance measures, $z = 0$ is optimal. This is the case if, for example, Δ is the worst-case or average-case fidelity, or it is the worst-case Schatten $1-1$-norm or the diamond norm. This is, however, not true in general. We construct an explicit example for Δ satisfying all assumptions 4.3 for which $z = 0$ is not optimal.

We give explicit tradeoffs in the case of non-degenerate von Neumann measurements for a variety of distance measures.

Theorem B.2 (Total variation - fidelity tradeoff [2, theorem 2])**.** *Consider a non-degenerate von Neumann measurement, given by an orthonormal basis in $\mathbb{C}^d$, and an instrument with d corresponding outcomes. Then the worst-case total variational distance δ_{TV} and the worst-case fidelity f satisfy*

$$\delta_{TV} \geq \begin{cases} \frac{1}{d} \left| \sqrt{f(d-1)} - \sqrt{1-f} \right|^2 & \text{if } f \geq \frac{1}{d}, \\ 0 & \text{if } f \leq \frac{1}{d}. \end{cases} \tag{B.2}$$

The inequality is tight and equality is attainable within the one-parameter family of instruments in equation (B.1) with $z = 0$.

Theorem B.3 (Total variation - average fidelity tradeoff [2, theorem 3])**.** *Consider a non-degenerate von Neumann measurement, given by an orthonormal basis in $\mathbb{C}^d$, and an instrument*

with d corresponding outcomes. Then the worst-case total variational distance δ_{TV} and the average-case fidelity $\bar{f}$ satisfy

$$\delta_{TV} \geq \begin{cases} \frac{1}{d}\left|\sqrt{\left(\bar{f}-\frac{1}{d+1}\right)\frac{d^2-1}{d}}-\sqrt{\left(1-\bar{f}\right)\frac{d+1}{d}}\right|^2 & \text{if } \bar{f}\geq\frac{2}{d+1}, \\ 0 & \text{if } \bar{f}\leq\frac{2}{d+1}. \end{cases} \tag{B.3}$$

The inequality is tight and equality is attainable within the one-parameter family of instruments in equation (B.1) with $z = 0$.

Theorem B.4 (Total variation - trace norm tradeoff [2, corollary 3])**.** *Consider a non-degenerate von Neumann measurement, given by an orthonormal basis in $\mathbb{C}^d$, and an instrument with d corresponding outcomes. Then the worst-case total variational distance δ_{TV} and its trace-norm analogue Δ_{TV} satisfy*

$$\delta_{TV} \geq \begin{cases} \frac{1}{d}\left|\sqrt{(1-\Delta_{TV})(d-1)}-\sqrt{\Delta_{TV}}\right|^2 & \text{if } \Delta_{TV}\leq 1-\frac{1}{d}, \\ 0 & \text{if } \Delta_{TV}\geq 1-\frac{1}{d}. \end{cases} \tag{B.4}$$

The inequality is tight and equality is attainable within the one-parameter family of instruments in equation (B.1) with $z = 0$.

The diamond norm, which is operationally the most relevant distance measure when it comes to distinguishing quantum channels, is treated in a more general setting. In this case, we allow the target measurement to be a possibly degenerate von Neumann measurement. This does, however, not affect the optimal tradeoff curve in case of the diamond norm. We prove that the optimal tradeoff only depends on the number of outcomes and is independent of the dimensions of the projections.

Theorem B.5 (Total variation - diamond norm tradeoff [2, theorem 4])**.** *If an instrument is considered approximating a (possibly degenerate) von Neumann measurement with m outcomes, then the worst-case total variational distance δ_{TV} and the diamond norm distance $\Delta_\diamond$ satisfy*

$$\delta_{TV} \geq \begin{cases} \frac{1}{2m}\left(\sqrt{(2-\Delta_\diamond)(m-1)}-\sqrt{\Delta_\diamond}\right)^2 & \text{if } \Delta_\diamond\leq 2-\frac{2}{m}, \\ 0 & \text{if } \Delta_\diamond > 2-\frac{2}{m}. \end{cases} \tag{B.5}$$

The inequality is tight in the sense that for every choice of the von Neumann measurement there is an instrument achieving equality.

Furthermore, we examine the more general case when the target measurement is given by an arbitrary positive operator-valued measure (POVM). We can then characterize the achievable region in the $\Delta - \delta$-plane as the feasible set of some semidefinite program (SDP), if Δ and δ are convex semialgebraic.

Theorem B.6 (SDP solution for arbitrary target measurements [2, theorem 7])**.** *If Δ and δ are both convex and semialgebraic, then the accessible region in the $\Delta - \delta$-plane is the feasible set of a SDP.*

We show that in the case of interest, if we consider a Schatten p-to-q-norm distance, with p and q rational, to describe the disturbance caused to the quantum system and a worst-case l_p-norm distance, with rational p, to quantify the measurement error, the accessible region in the $\Delta - \delta$-plane is the feasible set of a SDP.

Theorem B.7 ([2, corollary 5])**.** *The Schatten p-to-q norm-distances of a quantum channel $\Phi \in \mathcal{T}_d$ to the identity channel*

$$\Phi \ \mapsto ||\Phi - \mathrm{id}\,||_{p\to q,n} := \sup_{\rho\in\mathcal{D}_{dn}} \frac{||(\Phi - \mathrm{id}) \otimes \mathrm{id}_n(\rho)||_q}{||\rho||_p}, \quad n \in \mathbb{N}, \tag{B.6}$$

are semialgebraic for all $p, q \in [1, \infty) \cap \mathbb{Q}$ and $p, q = \infty$.

The worst-case fidelity distance of a quantum channel $\Phi \in \mathcal{T}_d$ to the identity channel

$$\Phi \ \mapsto \inf_{\rho\in\mathcal{D}_d} F\left(\Phi(\rho), \rho\right)^2 \tag{B.7}$$

is semialgebraic.

The worst-case l_p-distances of a POVM $E' \in \mathcal{E}_{d,m}$ to the target POVM $E \in \mathcal{E}_{d,m}$

$$E' \ \mapsto \sup_{\rho\in\mathcal{D}_d} ||\left(\mathrm{Tr}[\rho E_i] - \mathrm{Tr}[\rho E_i']\right)_{i=1}^m||_p, \tag{B.8}$$

are semialgebraic for all $p \in [1, \infty) \cap \mathbb{Q}$ and $p = \infty$.

In the special case of the worst-case l_∞-distance and the diamond norm, we explicitly state the SDP yielding the optimal tradeoff curve if considering a general POVM as target measurement. We apply this result to the example of a qubit symmetric, informationally complete (SIC) POVM and a qutrit SIC POVM.

Statement of individual contribution

This project is a follow up project to the previous one on universal asymmetric quantum cloning [1]. Instead of looking at the tradeoff between two quantum outputs of a channel, this project involves studying the tradeoff between a quantum output and a classical output of a channel. I, Anna-Lena Karolyn Hashagen, had the idea for this project while visiting the quantum information theory research group of Prof. Dr. Reinhard F. Werner at the Leibniz Universität Hannover. It was the result of many discussions with some of the group members, in particular with René Schwonnek. The concrete research question was then a joint effort of my doctoral supervisor, Prof. Dr. Michael M. Wolf, and me.

We started to examine the research question by looking at concrete measures, namely the worst-case l_∞-norm distance to quantify the measurement error and the diamond norm distance to quantify the disturbance caused to the system. It was immediately clear that the symmetry methods, discussed in chapter 3, apply to this particular research question. I proved proposition [2, proposition 1] and lemma [2, lemma 1] in close consultation with my doctoral supervisor, Prof. Dr. Michael M. Wolf. Even though at that time, we were still looking at

these concrete examples of distance measures, it was obvious that the same proof holds for any distance measure fulfilling some specific assumptions.

In a joint effort, we then wrote down the explicit SDP given in theorem [2, theorem 8] and we proved proposition [2, proposition 3]. I then decided to perform some numerical work to get an Ansatz for an analytical solution. I was solely responsible for the numerical work throughout this project.

I then lifted all our results to general distance measures satisfying some specific assumptions. My doctoral supervisor, Prof. Dr. Michael M. Wolf, had the idea to construct a von Neumann algebra isomorphism to obtain a manageable representation of the research problem. Many discussions via e-mail and video conferences then resulted in theorem [2, theorem 1]. Finding the optimal tradeoffs for specific distance measures was then possible by using the explicit form of the optimal quantum instruments found in theorem [2, theorem 1].

The numerical work that I performed suggested that the optimal tradeoff curve w.r.t. the worst-case total variational distance and the diamond norm distance is independent of the dimension of the projections. My doctoral supervisor, Prof. Dr. Michael M. Wolf, and I were able to prove this in lemma [2, lemma 7] and derived and proved the optimal tradeoff in theorem [2, theorem 4].

I studied the Helton-Nie conjecture in one of the group's seminars at the beginning of my doctoral studies [130–132]. We then had the idea that this might be applicable to the characterization of the achievable region in the $\Delta - \delta$-plane. I pursued this direction of thought and derived and proved all results stated in section [2, section 6]. I was solely responsible for all of section [2, section 6] and all of the appendix [2, appendix].

Furthermore, I was fully responsible for writing this article and finishing up the final version.

I, Anna-Lena Karolyn Hashagen, am the principal author of this article and I was extensively involved in all parts of it.

Journal permission and article

Subject **AHP final decision**
From AHP C. Delongeas <annaleshenripoincare@gmail.com>
To Anna-Lena Hashagen <hashagen@ma.tum.de>
Copy Pillet Claude-Alain <pillet@univ-tln.fr>
Date 2018-09-11 07:52

Dear Dr. Hashagen,

The revised version of your article: “Universality and Optimality in the Information-Disturbance Tradeoff" in collaboration with Michael M. Wolf, is definitely accepted for publication in AHP.

You will receive the proofs for reading as soon as possible.

Best regards,

Chantal Delongeas
Annales Henri Poincaré

Copyright Information: Annales Henri Poincaré

Copyright Information

For Authors

Submission of a manuscript implies: that the work described has not been published before (except in form of an abstract or as part of a published lecture, review or thesis); that it is not under consideration for publication elsewhere; that its publication has been approved by all co-authors, if any, as well as – tacitly or explicitly – by the responsible authorities at the institution where the work was carried out.

Author warrants (i) that he/she is the sole owner or has been authorized by any additional copyright owner to assign the right, (ii) that the article does not infringe any third party rights and no license from or payments to a third party is required to publish the article and (iii) that the article has not been previously published or licensed. The author signs for and accepts responsibility for releasing this material on behalf of any and all co-authors. Transfer of copyright to Springer (respective to owner if other than Springer) becomes effective if and when a Copyright Transfer Statement is signed or transferred electronically by the corresponding author. After submission of the Copyright Transfer Statement signed by the corresponding author, changes of authorship or in the order of the authors listed will not be accepted by Springer.

The copyright to this article, including any graphic elements therein (e.g. illustrations, charts, moving images), is assigned for good and valuable consideration to Springer effective if and when the article is accepted for publication and to the extent assignable if assignability is restricted for by applicable law or regulations (e.g. for U.S. government or crown employees).

The copyright assignment includes without limitation the exclusive, assignable and sublicensable right, unlimited in time and territory, to reproduce, publish, distribute, transmit, make available and store the article, including abstracts thereof, in all forms of media of expression now known or developed in the future, including pre- and reprints, translations, photographic reproductions and microform. Springer may use the article in whole or in part in electronic form, such as use in databases or data networks for display, print or download to stationary or portable devices. This includes interactive and multimedia use and the right to alter the article to the extent necessary for such use.

Authors may self-archive the Author's accepted manuscript of their articles on their own websites. Authors may also deposit this version of the article in any repository, provided it is only made publicly available 12 months after official publication or later. He/she may not use the publisher's version (the final article), which is posted on SpringerLink and other Springer websites, for the purpose of self-archiving or deposit. Furthermore, the Author may only post his/her version provided acknowledgement is given to the original source of publication and a link is inserted to the published article on Springer's website. The link must be accompanied by the following text: "The final publication is available at link.springer.com".

Prior versions of the article published on non-commercial pre-print servers like arXiv.org can remain on these servers and/or can be updated with Author's accepted version. The final published version (in pdf or html/xml format) cannot be used for this purpose. Acknowledgement needs to be given to the final publication and a link must be inserted to the published article on Springer's website, accompanied by the text "The final publication is available at link.springer.com". Author retains the right to use his/her article for his/her further scientific career by including the final published journal article in other publications such as dissertations and postdoctoral qualifications provided acknowledgement is given to the original source of publication.

Author is requested to use the appropriate DOI for the article. Articles disseminated via link.springer.com are indexed, abstracted and referenced by many abstracting and information services, bibliographic networks, subscription agencies, library networks, and consortia.

For Readers

While the advice and information in this journal is believed to be true and accurate at the date of its

UNIVERSALITY AND OPTIMALITY IN THE INFORMATION-DISTURBANCE TRADEOFF

ANNA-LENA K. HASHAGEN[1] AND MICHAEL M. WOLF[1,2]

Abstract. We investigate the tradeoff between the quality of an approximate version of a given measurement and the disturbance it induces in the measured quantum system. We prove that if the target measurement is a non-degenerate von Neumann measurement, then the optimal tradeoff can always be achieved within a two-parameter family of quantum devices that is independent of the chosen distance measures. This form of almost universal optimality holds under mild assumptions on the distance measures such as convexity and basis-independence, which are satisfied for all the usual cases that are based on norms, transport cost functions, relative entropies, fidelities, etc. for both worst-case and average-case analysis. We analyze the case of the cb-norm (or diamond norm) more generally for which we show dimension-independence of the derived optimal tradeoff for general von Neumann measurements. A SDP solution is provided for general POVMs and shown to exist for arbitrary convex semialgebraic distance measures.

Contents

1. Introduction

The idea that measurements inevitably disturb a quantum system is so much folklore and so deeply routed in the foundations of quantum mechanics that it is difficult to trace back historically. It is certainly present in Heisenberg's original exposition of the uncertainty relation. However, it only became amenable to mathematical analysis after the 'projection postulate' was replaced by a more refined theory of the quantum measurement process [1, 2]. With the emergence of the field of quantum information theory, the interest in a quantitative analysis of the information-disturbance tradeoff has intensified. At the same time, it became an issue of practical significance for many quantum information processing tasks, most notably for quantum cryptography [3, 4, 5, 6].

In the last two decades numerous papers derived quantitative bounds on the disturbance induced by a quantum measurement. A coarse way to categorize the existing approaches is depending on whether or not there are reference measurements w.r.t. which information gain on one side and disturbance on the other side are quantified. In [7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19] disturbance and information gain are both considered w.r.t. reference measurements. In [20, 21, 22, 23, 24, 25, 26, 27], in contrast, no reference observable is used on either side. In the present paper, we follow an intermediate route: we consider the performed measurement as an approximation of a given reference measurement, but we quantify the disturbance without specifying a second observable.

Another way of classifying previous works is in terms of the measures that are used to mathematically formalize and quantify disturbance and information gain: for instance, [7, 21, 25, 16, 22, 17] use various entropic measures, [23, 12, 19] use norm-based measures, [20, 21, 25, 26] use fidelities, [11, 27] use Fisher information, and [13, 18] use transport-cost functions. Many other measures are conceivable and most of them come in two flavors: a worst-case and an average-case variant, where the latter again calls for the choice of an underlying distribution.

A central point of the present work is to show that the information-disturbance problem has a core that is largely independent of the measures chosen. More specifically, we prove the existence of a small set of devices that are (almost) universally optimal independent of the chosen measures, as long as these exhibit a set of elementary properties that are shared by the vast majority of distance measures found in the literature. Based on this universality result, we then derive optimal tradeoff bounds for specific choices of measures. These include the diamond norm and its classical counterpart the total variation distance. In this case, the reachability of the optimal tradeoff has been demonstrated experimentally in a parallel work [28].

Organization of the paper. Sec. 2 starts off with introducing the setup and summarizes the paper's main results. In Sec. 3, we discuss distance measures that quantify the measurement error and the disturbance caused to the system. We give a brief overview of common measures found in the literature that fulfill the assumptions we make, necessary to derive the universality theorem. In Sec. 4, for the case of a non-degenerate von Neumann target measurement, we derive a universal two-parameter family of optimal devices that yield the best information-disturbance tradeoff. In Sec. 5, still for the case of a non-degenerate von Neumann target measurement, we use the universal optimal devices derived in the previous section to compute the optimal tradeoff for a variety of distance measures. In the special case where we consider the diamond norm for quantifying disturbance, we derive the optimal tradeoff also for the case of degenerate von Neumann target measurements. In the last section, Sec. 6, we show that the optimal tradeoff can always be represented as a SDP if the distance measures under consideration are convex semialgebraic. We give the explicit SDP that represents the tradeoff between the diamond norm and the worst-case l_∞-distance and apply it to the special case of qubit as well as qutrit SIC POVMs.

2. Summary

This section will briefly introduce some notation, specify the considered setup, and summarize the main results. More details and proofs will then be given in the following sections.

Notation. Throughout we will consider finite dimensional Hilbert spaces $\mathbb{C}^d$, write $\mathcal{M}_d$ for the set of complex $d \times d$ matrices and $\mathcal{S}_d \subseteq \mathcal{M}_d$ for the subset of density operators, usually denoted by ρ. An m-outcome measurement on this space will be described by a *positive operator valued measure* (POVM) $E = (E_1, \ldots, E_m)$ whose elements $E_i \in \mathcal{M}_d$ are positive semidefinite and sum up to the identity operator $\sum_{i=1}^m E_i = \mathbb{1}$. The set of all such POVM's will be denoted by $\mathcal{E}_{d,m}$ and we will set $\mathcal{E}_d := \mathcal{E}_{d,d}$. We will call E a *von Neumann measurement* if the E_i's are mutually orthogonal projections and further call it *non-degenerate* if those are one-dimensional, i.e., characterized by an orthonormal basis. A completely positive, trace-preserving linear map will be called a *quantum channel* and the set of quantum channels from $\mathcal{M}_d$ into $\mathcal{M}_d$ will be denoted by $\mathcal{T}_d$.

Setup. We will fix a *target measurement* $E \in \mathcal{E}_{d,m}$ and investigate the tradeoff between the quality of an approximate measurement of E, say by $E' \in \mathcal{E}_{d,m}$, and the disturbance the measurement process induces in the system. The evolution of the latter will be described by some channel $T_1 \in \mathcal{T}_d$. To this end, we will have to choose two suitable functionals $E' \mapsto \delta(E')$ and $T_1 \mapsto \Delta(T_1)$ that quantify the deviation of E' and T_1 from the target measurement E and the ideal channel id, respectively.

For a given triple (E, δ, Δ) the question will then be: what is the accessible region in the $\delta - \Delta$-plane when running over all possible measurement devices and, in particular, what is the optimal tradeoff curve and how can it be achieved?

Clearly, E' and T_1 are not independent. The framework of *instruments* allows to describe all pairs (E', T_1) that are compatible within the rules of quantum theory. An *instrument* assigns to each possible outcome i of a measurement a completely positive map $I_i : \mathcal{M}_d \to \mathcal{M}_d$ so that the corresponding POVM element is $E'_i := I_i^*(\mathbb{1})$ and the evolution of the remaining quantum system is governed by $T_1 := \sum_{i=1}^m I_i$. Normalization requires that this sum is trace-preserving.

Main results. There are zillions of possible choices for the measures Δ and δ. If one had to choose one pair that stands out for operational significance this would probably be the *diamond norm* and its classical counterpart, the *total variational distance* (defined and discussed in Sec. 3 and Sec. 5). One of our results is the derivation of the optimal tradeoff curve for this pair (Thm. 4 in Sec. 5.5):

Theorem (Total variation - diamond norm tradeoff)**.** *If an instrument is considered approximating a (possibly degenerate) von Neumann measurement with m outcomes, then the worst-case total variational distance δ_{TV} and the diamond norm distance $\Delta_\diamond$ satisfy*

$$\delta_{TV} \geq \begin{cases} \frac{1}{2m}\left(\sqrt{(2-\Delta_\diamond)(m-1)} - \sqrt{\Delta_\diamond}\right)^2 & \text{if } \Delta_\diamond \leq 2 - \frac{2}{m}, \\ 0 & \text{if } \Delta_\diamond > 2 - \frac{2}{m}. \end{cases} \tag{1}$$

The inequality is tight in the sense that for every choice of the von Neumann measurement there is an instrument achieving equality.

Note that the tradeoff depends solely on the number m of outcomes and is independent of the dimension of the underlying Hilbert space (apart from $d \geq m$). Also note that the accessible region shrinks with increasing m and in the limit $m \to \infty$ becomes a triangle, determined by $\delta_{TV} \geq 1 - \Delta_\diamond/2$.

In Sec. 5 we derive similar results for the worst-case as well as average-case fidelity and trace-norm. In all cases, the bounds are tight and we show how the optimal tradeoff can be achieved. Instead of going through these and more examples one-by-one we follow a different approach. We provide a general tool for obtaining optimal tradeoffs for *all pairs* (δ, Δ) that exhibit a set of elementary properties that are shared by the vast majority of distance measures that can be found in the literature. These properties, which are discussed in Sec. 3, are essentially convexity and suitable forms of basis-(in)dependence. For the case of a non-degenerate von Neumann target measurement Thm. 1 in Sec. 4 shows that optimal devices can always be found within a universal two-parameter family, independent of the specific choice of δ and Δ:

Theorem ((Almost universal) optimal instruments)**.** *Let Δ and δ be distance-measures for quantifying disturbance and measurement-error that satisfy Assumptions 1 and 2 (cf. Sec. 3), respectively. Then the optimal $\Delta-\delta$-tradeoff w.r.t. a target measurement that is given by an orthonormal basis $\{|i\rangle \in \mathbb{C}^d\}_{i=1}^d$ is attained within the two-parameter family of instruments defined by*

$$I_i(\rho) := z\langle i|\rho|i\rangle \frac{\mathbb{1}_d - |i\rangle\langle i|}{d-1} + (1-z)K_i\rho K_i, \quad K_i := \mu\mathbb{1}_d + \nu|i\rangle\langle i|, \qquad (2)$$

where $z \in [0,1]$ and $\mu, \nu \in \mathbb{R}$ satisfy $d\mu^2 + \nu^2 + 2\mu\nu = 1$ (which makes $\sum_i I_i$ trace preserving).

While the parameter z can be eliminated for instance in all cases mentioned above, we show in Cor. 2 that this is not possible in general.

If the target measurement itself is not a von Neumann measurement but a general POVM, then closed-form expressions like the ones above should not be expected. For the important case of the diamond norm, we show in Sec. 6 how the optimal tradeoff curve can still be obtained via a semidefinite program (SDP). This is an instance of the following more general fact (Thm. 7):

Theorem (SDP solution for arbitrary target measurements)**.** *If Δ and δ are both convex and semialgebraic, then the accessible region in the $\Delta-\delta$-plane is the feasible set of a SDP.*

Note that no assumptions on the chosen measures are made other than being convex and semialgebraic.

3. Distance measures

In this section we have a closer look at the functionals $\Delta : \mathcal{T}_d \to [0,\infty]$ and $\delta : \mathcal{E}_{d,m} \to [0,\infty]$ that quantify how much E' and T_1 differ from E and id, respectively. We will not assume that they arise from metrics and use the notion of a 'distance' merely in the colloquial sense. We will state the assumptions that we will use in Sec. 4 and discuss some of the most common measures that appear in the literature.

Quantifying disturbance. For the universality theorem (Thm. 1) we will need the following assumption on Δ:[1]

Assumption 1 (on the distance measure to the identity channel)**.**
For $\Delta : \mathcal{T}_d \to [0,\infty]$ we assume that (a) $\Delta(\mathrm{id}) = 0$, (b) Δ is convex, and (c) Δ is basis-independent in the sense that for every unitary $U \in \mathcal{M}_d$ and every channel $\Phi \in \mathcal{T}_d$:

$$\Delta\Big(U\Phi(U^* \cdot U)U^*\Big) = \Delta(\Phi). \qquad (3)$$

[1]In fact, slightly less is required since Eq. (3) will only be used for unitaries that are products of diagonal and permutation matrices.

In the usually considered cases, Δ arises from a distance measure on the set of density operators $\mathcal{S}_d \subseteq \mathcal{M}_d$. In fact, if $\tilde{\Delta} : \mathcal{S}_d \times \mathcal{S}_d \to [0, \infty]$ is convex in its first argument, unitarily invariant and satisfies $\tilde{\Delta}(\rho, \rho) = 0$, then considering the worst case as well as the average case w.r.t. the input state both lead to functionals that satisfy Assumption 1. More precisely, if μ is a unitarily invariant measure on $\mathcal{S}_d$ and $S \subseteq \mathcal{S}_d$ a unitarily closed subset (e.g., the set of all pure states), then the following two definitions can easily be seen to satisfy Assumption 1, see the appendix:

$$\begin{aligned} \Delta_\infty(\Phi) &:= \sup_{\rho \in S} \tilde{\Delta}\big(\Phi(\rho), \rho\big), \\ \Delta_\mu(\Phi) &:= \int_{\mathcal{S}_d} \tilde{\Delta}\big(\Phi(\rho), \rho\big)\, \mathrm{d}\mu(\rho). \end{aligned}$$

While Δ_∞ quantifies the distance between Φ and id in the worst case in terms of $\tilde{\Delta}$, Δ_μ does the same for the average case.

Concrete examples for $\tilde{\Delta}$ are (i) $\tilde{\Delta}(\rho, \sigma) = 1 - F(\rho, \sigma)$, where $F(\rho, \sigma) := ||\sqrt{\rho}\sqrt{\sigma}||_1$ is the fidelity, (ii) the relative entropy and many other quantum f-divergences [29] including the Chernoff- and Hoeffding-distance and (iii) $\tilde{\Delta}(\rho, \sigma) = |||\rho - \sigma|||$, where $||| \cdot |||$ is any unitarily invariant norm such as the Schatten p-norms.

The latter can, in a similar vein, be used to define Schatten p-to-q norm-distances to the identity channel

$$\Phi \mapsto ||\Phi - \mathrm{id}||_{p \to q, n} := \sup_{\rho \in \mathcal{S}_{dn}} \frac{||(\Phi - \mathrm{id}) \otimes \mathrm{id}_n(\rho)||_q}{||\rho||_p}, \quad q, p \in [1, \infty], n \in \mathbb{N},$$

which also fulfill Assumption 1. Special cases are given by the *diamond norm* $|| \cdot ||_\diamond := || \cdot ||_{1 \to 1, d}$, which we discuss in more detail in Sec. 5.5, and its dual, the *cb-norm* (with $p = q = \infty, n = d$).

Quantifying measurement error. The following assumptions that we need for the universality theorem on the functional δ refer to the case of a non-degenerate von Neumann target measurement that is given by an orthonormal basis $(|i\rangle\langle i|)_{i=1}^d$.

Assumption 2 (on the distance measure to the target measurement)**.** *For $\delta : \mathcal{E}_d \to [0, \infty]$ we assume that (a) $\delta\big((|i\rangle\langle i|)_{i=1}^d\big) = 0$, (b) δ is convex, (c) δ is permutation-invariant in the sense that for every permutation $\pi \in S_d$ and any $M \in \mathcal{E}_d$*

$$M_i' = U_\pi^* M_{\pi(i)} U_\pi \ \forall i \ \Rightarrow \ \delta(M') = \delta(M), \tag{4}$$

where U_π is the permutation matrix that acts as $U_\pi |i\rangle = |\pi(i)\rangle$, and (d) that for every diagonal unitary $D \in \mathcal{M}_d$ and any $M \in \mathcal{E}_d$

$$M_i' = D^* M_i D \ \forall i \ \Rightarrow \ \delta(M') = \delta(M). \tag{5}$$

Here, the most common cases arise from distance measures $\tilde{\delta} : \mathcal{P}_d \times \mathcal{P}_d \to [0, \infty]$ on the space of probability distributions $\mathcal{P}_d := \big\{q \in \mathbb{R}^d \big| \sum_{i=1}^d q_i =$

$1 \wedge \forall i : q_i \geq 0\big\}$ applied to the target distribution $p_i := \langle i|\rho|i\rangle$ and the actually measured distribution $p'_i := \operatorname{tr}[\rho E'_i]$. Suppose $\tilde{\delta}$ is convex in its second argument, invariant under joint permutations and satisfies $\tilde{\delta}(q,q) = 0$. Then the worst-case as well as the average-case construction

$$\begin{aligned} \delta_\infty(E') &:= \sup_{\rho\in S} \tilde{\delta}(p,p'), \\ \delta_\mu(E') &:= \int_{\mathcal{S}_d} \tilde{\delta}(p,p')\, \mathrm{d}\mu(\rho), \end{aligned}$$

both satisfy Assumption 2, see appendix. Concrete examples for $\tilde{\delta}$ are all l_p-norms for $p \in [1,\infty]$ and the Kullback-Leibler divergence as well as other f-divergences. Other examples for δ that satisfy Assumption 2 are transport cost functions like the ones used in [18].

Note that convexity of the two measures Δ and δ implies that the region in the $\Delta-\delta$-plane that is accessible by quantum instruments is a convex set. The boundary of this set is given by two lines that are parallel to the axes (and correspond to the maximal values of Δ and δ) and what we call the *optimal tradeoff curve*.

4. Universal optimal devices

There are three major steps towards proving the claimed universality theorem: the exploitation of symmetry, the construction of a von Neumann algebra isomorphism to obtain a manageable representation, and the final reduction to the envelope of a unit cone.

Throughout this section, the target measurement will be given by an orthonormal basis $E = (|i\rangle\langle i|)_{i=1}^d$. In this case, instead of working with instruments it turns out to be slightly more convenient to work with channels. More specifically, we will describe the entire process by a channel $T : \mathcal{M}_d \to \mathcal{M}_d \otimes \mathcal{M}_d$ with marginals $T_1, T_2 \in \mathcal{T}_d$. T_1 will then reflect the evolution of the 'disturbed' quantum system, whereas the output of T_2 is measured by E leading to $E'_i = T_2^*(E_i)$. This is clearly describable by an instrument and conversely, for every instrument I we can simply construct

$$T(\rho) := \sum_{i=1}^d I_i(\rho) \otimes |i\rangle\langle i|,$$

which shows that the two viewpoints are equivalent.

Proposition 1 (Reduction to symmetric channels)**.** *Let G be the group generated by all diagonal unitaries and permutation matrices in $\mathcal{M}_d$. If Δ and δ satisfy Assumptions 1 and 2, respectively, the optimal tradeoff between them can be attained within the set of channels $T : \mathcal{M}_d \to \mathcal{M}_d \otimes \mathcal{M}_d$ for which*

$$(U\otimes U)T\big(U^*\rho U\big)(U\otimes U)^* \;=\; T(\rho) \quad \forall U \in G, \rho \in \mathcal{S}_d. \tag{6}$$

Proof. We will show that for an arbitrary channel T, which does not necessarily satisfy Eq. (6), the symmetrization

$$\bar{T} := \int_G (U \otimes U) T\big(U^* \cdot U\big)(U \otimes U)^* \, \mathrm{d}U$$

w.r.t. the Haar measure of G performs at least as well as T. Let $\bar{T}_1$ and $\bar{T}_2$ be the marginals of $\bar{T}$. Then

$$\begin{aligned} \Delta\big(\bar{T}_1\big) &= \Delta\left(\int_G U T_1\big(U^* \cdot U\big) U^* \, \mathrm{d}U\right) \\ &\overset{(1b)}{\leq} \int_G \Delta\left(U T_1\big(U^* \cdot U\big) U^*\right) \, \mathrm{d}U \overset{(1c)}{=} \Delta(T_1), \end{aligned}$$

where the used assumption is indicated above the (in-)equality sign. Similarly, we obtain

$$\begin{aligned} \delta\left[\Big(\bar{T}_2^*\big(|i\rangle\langle i|\big)\Big)_{i=1}^d\right] &\overset{(2b)}{\leq} \int_G \delta\left[\Big(U^* T_2^*\big(U|i\rangle\langle i|U^*\big)U\Big)_{i=1}^d\right] \mathrm{d}U \\ &\overset{(2d)}{=} \int_G \delta\left[\Big(U_\pi^* T_2^*\big(|\pi(i)\rangle\langle\pi(i)|\big)U_\pi\Big)_{i=1}^d\right] \mathrm{d}U \\ &\overset{(2c)}{=} \delta\left[\Big(T_2^*\big(|i\rangle\langle i|\big)\Big)_{i=1}^d\right], \end{aligned}$$

where we have used that every $U \in G$ can be written as $U = U_\pi D$, where U_π is a permutation and D a diagonal unitary, both depending on U.

Consequently, when replacing T by its symmetrization $\bar{T}$, which satisfies Eq. (6) by construction, neither Δ nor δ is increasing. □

Lemma 1 (Structure of marginals of symmetric channels)**.** *Let G be the group generated by all diagonal unitaries and permutation matrices in $\mathcal{M}_d$ and $\Phi : \mathcal{M}_d \to \mathcal{M}_d$ a quantum channel. Then the following are equivalent:*

(1) $\Phi(\rho) = U\Phi\big(U^*\rho U\big)U^* \quad \forall U \in G, \rho \in \mathcal{S}_d$.

(2) There are $\alpha, \beta, \gamma \in \mathbb{R}$ with $\alpha + \beta + \gamma = 1$ so that

$$\Phi = \alpha \operatorname{tr}[\cdot]\frac{\mathbb{1}}{d} + \beta \operatorname{id} + \gamma \sum_{i=1}^d |i\rangle\langle i|\langle i| \cdot |i\rangle. \tag{7}$$

Proof. (2) ⇒ (1) can be seen by direct inspection. In order to prove the converse, we consider the Jamiolkowski-state (= normalized Choi-matrix) $J_\Phi := \frac{1}{d}\sum_{i,j=1}^d \Phi\big(|i\rangle\langle j|\big) \otimes |i\rangle\langle j|$. Then (1) is equivalent to the statement that J_Φ commutes with all unitaries of the form $U \otimes \bar{U}$, $U \in G$. Considering for the moment only the subgroup of diagonal unitaries, this requires that

$$\langle ij|J_\Phi|kl\rangle = (2\pi)^{-d} \int_0^{2\pi} \ldots \int_0^{2\pi} e^{i(\varphi_i - \varphi_j - \varphi_k + \varphi_l)} \langle ij|J_\Phi|kl\rangle \, \mathrm{d}\varphi_1 \ldots d\varphi_d,$$

which vanishes unless $(i = j \wedge k = l) \vee (i = k \wedge j = l)$. Hence, there are $A, B \in \mathcal{M}_d$ such that

$$J_\Phi = \sum_{i,j=1}^{d} A_{ij}|i\rangle\langle i| \otimes |j\rangle\langle j| + B_{ij}|i\rangle\langle j| \otimes |i\rangle\langle j|.$$

Next, we will exploit that J_Φ commutes in addition with permutations of the form $U_\pi \otimes U_\pi$ for all $\pi \in S_d$. For $i \neq j$ this implies that $A_{i,j} = A_{\pi(i),\pi(j)}$ and $B_{i,j} = B_{\pi(i),\pi(j)}$ so that there is only one independent off-diagonal element for each A and B. The case $i = j$ leads to a third parameter that is a coefficient in front of $\sum_i |ii\rangle\langle ii|$. Translating this back to the level of quantum channels then yields Eq. (7). The coefficients are real and sum up to one since Φ preserves hermiticity as well as the trace. □

If T is symmetric as in Prop. 1, then both marginal channels T_1 and T_2 are of the form derived in the previous Lemma. That is, each T_i, $i \in \{1, 2\}$, is specified by three parameters $\alpha_i, \beta_i, \gamma_i$ only two of which are independent.

The following Lemma shows that under Assumption 2 the error measure δ depends only on α_2 and does so in a non-decreasing way.

Lemma 2. *Let δ satisfy Assumption 2. There is a non-decreasing function $\hat{\delta} : [0, 1] \to [0, \infty]$ s.t. for all $T_2 : \mathcal{M}_d \to \mathcal{M}_d$ of the form in Eq. (7) with coefficients $\alpha_2, \beta_2, \gamma_2$ we have $\delta\big[\big(T_2^*(|i\rangle\langle i|)\big)_{i=1}^d\big] = \hat{\delta}(\alpha_2)$.*

Proof. The statement follows from convexity of δ together with the observation that β and γ only contribute jointly to δ and not individually. This is seen by composing T_2 with the projection onto the diagonal. This leads to a channel of the same form, but possibly different parameters. On the level of the latter the composition corresponds to $(\alpha_2, \beta_2, \gamma_2) \mapsto (\alpha_2, 0, \beta_2 + \gamma_2)$. The distance measure δ, however, does not change in this process and thus depends only on the sum $\beta_2 + \gamma_2$ and not on those two parameters individually. As this sum equals $1 - \alpha_2$ we see that δ can be regarded as a function of α_2 only. We formally denote this function by $\hat{\delta}$. Assumption (2b) then implies that $\hat{\delta}$ is convex. As it is in addition positive and satisfies $\hat{\delta}(0) = 0$ by Assumption (2a), we get that $\hat{\delta}$ is non-decreasing. □

For later investigation, it is useful to decompose the J_Φ that corresponds to Eq. (7) into its spectral projections:

$$\begin{aligned} J_\Phi &= aP_a + bP_b + cP_c, \quad \text{where} \quad P_a := \mathbb{1} - \sum_{i=1}^{d} |ii\rangle\langle ii|, \\ P_b &:= \frac{1}{d}\sum_{i,j=1}^{d} |ii\rangle\langle jj|, \quad P_c := \sum_{i=1}^{d} |ii\rangle\langle ii| - P_b. \end{aligned} \tag{8}$$

The coefficients a, b, c are the eigenvalues of J_Φ (and thus non-negative) and related to α, β, γ via $\alpha = d^2 a$, $\beta = b - c$, $\gamma = d(c - a)$. When considering

symmetric T, we will label the eigenvalues of J_{T_i} with a subscript $i \in \{1, 2\}$ to distinguish the two marginals.

Since the P's are mutually orthogonal projectors, we can obtain the eigenvalues from their expectation values. That is,

$$x_1 = \frac{\mathrm{tr}\left[(P_x \otimes \mathbb{1})J_T\right]}{\mathrm{tr}\left[P_x\right]} \quad \text{and} \quad x_2 = \frac{\mathrm{tr}\left[(\mathbb{1} \otimes P_x)J_T\right]}{\mathrm{tr}\left[P_x\right]}, \quad x \in \{a, b, c\}. \tag{9}$$

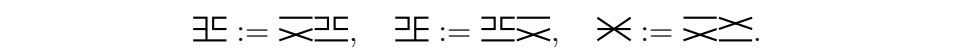
If we are aiming at identifying a subset of optimal channels, we can, according to Lemma 2, w.l.o.g. use a_2 as δ. Due to the monotonic relation between the two, optimality for one implies optimality for the other. The question we are going to address in the next step of the argumentation is then: which values of a_1, b_1 and c_1 are consistent with a given value of a_2? After all, due to Prop. 1, Δ and δ will be functions of those parameters only. Thus, we would like to know which is the accessible region in the space of these parameters, when we vary J_T over the set of all density matrices.

We tackle this question using an operator algebraic point of view: the operators $\mathbb{1} \otimes P_a, P_x \otimes \mathbb{1}$ together with the identity operator generate a von Neumann algebra $\mathcal{A}$ on which J_T acts as a state, i.e., as a normalized positive linear functional. This suggests the use of a von Neumann algebra isomorphism that simplifies the representation. To this end, we observe that $\mathcal{A}$ is generated by the following operators:

$\mathbb{1}_{d^3} =:$ $\qquad \mathbb{1}_d \otimes \sum_{i=1}^{d} |ii\rangle\langle ii| =:$ 

$\sum_{i,j=1}^{d} |ii\rangle\langle jj| \otimes \mathbb{1}_d =:$ $\qquad \sum_{i=1}^{d} |ii\rangle\langle ii| \otimes \mathbb{1}_d =:$ 

The introduced diagrammatic notation turns out be useful as it reflects that these operators are what one may call *contraction tensors*.[2] If we view an element in $\mathcal{M}_d \otimes \mathcal{M}_d \otimes \mathcal{M}_d$ as a tensor with three left and three right indices, then the diagrammatic notation indicates which of these indices get contracted together—by connecting them. Taking products of pairs of these four operators generates (up to scalar multiples, which arise from closed loops) three new contraction tensors:



The set of these seven tensors is, however, closed under multiplication (again ignoring scalar multiples). This is easily verified by using the diagrammatic notation and going through all cases. This observation is the core for constructing a simplifying isomorphism:

[2]Please note that these diagrams are not braid diagrams, but rather diagrammatically represent contraction tensors.

Lemma 3 (Isomorphic representation)**.** *Let $\mathcal{A}$ be the von Neumann algebra that is generated by the set $\{\mathbb{1}_{d^3}, \mathbb{1}_d \otimes P_a, P_a \otimes \mathbb{1}_d, P_b \otimes \mathbb{1}_d, P_c \otimes \mathbb{1}_d\}$. A unital map $\iota : \mathcal{A} \to \mathcal{M}_2 \oplus \mathbb{C}^3$ defined by*

$$\iota : \exists\sqsubset \mapsto d|e_1\rangle\langle e_1| \qquad \iota : \nsucc \mapsto |e_2\rangle\langle e_2| \tag{10}$$

$$\iota : \succcurlyeq \mapsto \mathbb{1}_2 \oplus f_2 \qquad \iota : \overline{\succ} \mapsto |e_2\rangle\langle e_2| \oplus f_1 \tag{11}$$

is an isomorphism if $|e_1\rangle, |e_2\rangle$ constitute unit vectors with $|\langle e_1|e_2\rangle|^2 = 1/d$ in the space of the non-abelian part (i.e., the corresponding projections as well as $\mathbb{1}_2$ are in $\mathcal{M}_2$) and $f_1 := (1,0,0), f_2 := (0,1,0)$ are elements of the abelian part.[3]

Proof. $\mathcal{A}$ is generated by the above set of seven contraction tensors. Since this set is closed under multiplication, *-operation and contains linear independent elements, we have $\dim(\mathcal{A}) = 7$. Moreover, $\mathcal{A}$ is non-commutative since $[\exists\sqsubset, \nsucc] \neq 0$. From the representation theory of finite-dimensional von Neumann algebras we known that every 7-dimensional non-commutative von Neumann algebra is isomorphic to $\mathcal{M}_2 \oplus \mathbb{C}^3$ [30, Thm. 5.6]. Hence, we can establish an isomorphism ι by representing a generating set of $\mathcal{A}$ in $\mathcal{M}_2 \oplus \mathbb{C}^3$. Due to unitality $\iota(\mathbb{1}_{d^3}) = \mathbb{1}_2 \oplus (1,1,1)$ has to hold. Moreover, since $\exists\sqsubset, \nsucc$ are (proportional to) non-commuting minimal projectors in $\mathcal{A}$, they need to be the same in $\mathcal{M}_2 \oplus \mathbb{C}^3$. Taking proportionality factors into account, this determines Eq. (10) and requires $|\langle e_1|e_2\rangle|^2 = 1/d$ in order to be consistent with the value of the trace $\mathrm{tr}\,[\exists\sqsubset\nsucc]$. From $\succcurlyeq\exists\sqsubset = \exists\sqsubset$ and $\succcurlyeq\nsucc = \nsucc$ we see that $\iota(\succcurlyeq)$ acts as identity on $\mathcal{M}_2$. Similarly, $\iota(\overline{\succ})$, when restricted to $\mathcal{M}_2$, has to be a projector that is not the identity and has $|e_2\rangle$ as eigenvector (due to $\overline{\succ}\nsucc = \nsucc$). This determines Eq. (11) when restricted to $\mathcal{M}_2$. Moreover, since $\mathcal{M}_2 \oplus \mathbb{C}^3$ has to be generated, both $\iota(\succcurlyeq)$ and $\iota(\overline{\succ})$ have to have non-zero parts on the abelian side. Since they are projectors, these parts need to be projectors as well. Finally, they have to be one-dimensional since otherwise the identity operator would become linearly dependent. □

Using this Lemma we can now express the accessible region within the space of parameters $\alpha_1, \beta_1, \gamma_1, \alpha_2$ by varying over all states on $\mathcal{M}_2 \oplus \mathbb{C}^3$, instead of over all states J_T on $\mathcal{M}_{d^3}$. To this end, we just have to unravel the linear maps from the parameters to the eigenvalues a_1, b_1, c_1, a_2, to the P's, to the contraction tensors, and finally to their representation in $\mathcal{M}_2 \oplus \mathbb{C}^3$. In this way, we obtain:

Corollary 1. *There exists a channel $T : \mathcal{M}_d \to \mathcal{M}_d \otimes \mathcal{M}_d$ with corresponding Jamiolkowski state J_T whose marginals give rise to the parameters α_1, β_1 and*

[3]Here we regard $\mathbb{C}^3$ as space $\mathcal{M}_1 \oplus \mathcal{M}_1 \oplus \mathcal{M}_1$ of diagonal matrices in $\mathcal{M}_3$.

a_2 iff there exists a state ϱ on $\mathcal{M}_2 \oplus \mathbb{C}^3$ such that

$$\alpha_1 = \frac{d}{d-1}\Big(1 - \mathrm{tr}\,[\mathbb{1}_2\varrho] - \mathrm{tr}\,[f_2\varrho]\Big), \tag{12}$$

$$\beta_1 = \langle e_1|\varrho|e_1\rangle - \frac{1}{d-1}\Big(\mathrm{tr}\,[\mathbb{1}_2\varrho] + \mathrm{tr}\,[f_2\varrho] - \langle e_1|\varrho|e_1\rangle\Big), \tag{13}$$

$$a_2 = \big(1 - \langle e_2|\varrho|e_2\rangle - \mathrm{tr}\,[f_1\varrho]\,\big)/(d^2-d), \tag{14}$$

where $\mathbb{C}^3$ is regarded as space of diagonal 3×3 matrices and e_1, e_2, f_1, f_2 are as in Lemma 3.

The proof of this corollary can be found in the appendix.

There is still unitary freedom in the choice of the vectors e_1, e_2. We utilize this and set

$$\langle e_1|\sigma_y|e_1\rangle = \langle e_2|\sigma_y|e_2\rangle = 0 \quad \text{and} \quad |e_2\rangle\langle e_2| = \frac{1}{2}(\mathbb{1}_2 + \sigma_x), \tag{15}$$

where the σ_i's are the usual Pauli matrices. So in particular, we choose the vectors such that the corresponding projectors lie in an equatorial plane of the Bloch sphere that is characterized by density matrices with real entries.

In order to simplify the problem further, we now focus more explicitly on minimizing a_2:

Proposition 2 (Reduction to the unit cone)**.** *Under the constraints given by Eqs. (12 – 15), the minimum value for a_2 for arbitrary fixed values of α_1, β_1 that is achievable by varying over all states ϱ is attained for a state of the form*

$$\varrho = \frac{1}{2}\Big((1-z)\mathbb{1}_2 + x\sigma_x + y\sigma_z\Big) \oplus (z,0,0), \tag{16}$$

where $(x,y,z) \in \mathbb{R}^3$ is an element of the envelope of the unit cone, i.e., $z \in [0,1], x^2+y^2 = (1-z)^2$.

Proof. We simplify the structure of ϱ in four steps, each of which eliminates one parameter. First, note that we can assume $\mathrm{tr}\,[f_3\varrho] = 0$, where f_3 is the diagonal matrix $(0,0,1)$. This is seen by considering the map $\varrho \mapsto \varrho + \mathrm{tr}\,[f_3\varrho]\,(f_1 - f_3)$, which decreases a_2, sets the f_3-component to zero, but leaves α_1 and β_1 unchanged.

Second, we claim that the f_2-component can be set to zero, as well. To this end, consider the map $\varrho \mapsto \varrho + \mathrm{tr}\,[f_2\varrho]\,(|e_1^\perp\rangle\langle e_1^\perp| - f_2)$ where $e_1^\perp$ is a unit vector in $\mathbb{C}^2$ that is orthogonal to e_1. By construction, this sets the f_2-component to zero, decreases a_2 and leaves α_1 and β_1 invariant. Taken together with the first step, this already shows that the abelian part of ϱ can be assumed to be of the form $(z,0,0)$ for some $z \in [0,1]$.

Third, observe that the σ_y-component of the non-abelian part of ϱ does not enter any of the equations so that we can as well set it to zero and thus assume that, restricted to $\mathcal{M}_2$, ϱ lies in the 'real' equatorial plane of the Bloch sphere.

Taking positivity and normalization into account, Eq. (16) summarizes these findings, so far with $x^2+y^2 \leq (1-z)^2$. What remains to show is that equality

can be assumed, here. Let $v_1, v_2, w \in \mathbb{R}^3$ be the Bloch vectors of e_1, e_2 and ϱ, respectively. Suppose $||w||_2 < 1$, which corresponds to a point that does not lie on the envelope of the cone and let $v_1^\perp \in \mathbb{R}^3$ be a unit vector in the equatorial plane that is orthogonal to v_1. Then the map $w \mapsto w + \epsilon v_1^\perp$, for sufficiently small ϵ of the right sign, leaves α_1 and β_2 unchanged, but decreases a_2. Hence, we can choose ϵ so that the Bloch vector reaches unit norm, which completes the proof of the proposition. □

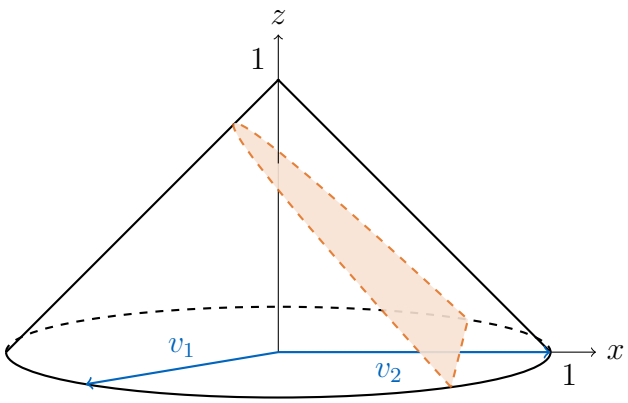


FIGURE 1. Sketch of the unit cone used in the construction of the proof in Prop. 2. The orange parabola corresponds to a fixed value of δ and the optimal device is contained within its boundary; its location depends on the chosen disturbance distance measure Δ.

This completes the list of ingredients that are needed for the main theorem of this section:

Theorem 1 ((Almost universal) optimal devices)**.** *Let Δ and δ be distance-measures for quantifying disturbance and measurement-error that satisfy Assumptions 1 and 2, respectively. Then the optimal $\Delta - \delta$-tradeoff is attained within the following two-parameter family of quantum channels:*

$$T(\rho) := \sum_{i=1}^{d} \left[z\langle i|\rho|i\rangle \frac{\mathbb{1}_d - |i\rangle\langle i|}{d-1} + (1-z) K_i \rho K_i \right] \otimes |i\rangle\langle i|, \qquad (17)$$
$$K_i := \mu \mathbb{1}_d + \nu |i\rangle\langle i|,$$

where $z \in [0,1]$ and $\mu, \nu \in \mathbb{R}$ are constrained by imposing T to be trace preserving.

Proof. What remains to do is to translate the two-parameter family of Eq. (16) into the world of channels. It suffices to consider the cases in which either $z = 0$ or $z = 1$ since these generate the general case by convex combination. In both cases the relevant von Neumann algebra is a factor on which the dual of ι becomes its inverse, up to a multiplicity factor. This means, we have to compute $\iota^{-1}(\varrho)$ and show that it equals J_T when normalized.

If $z = 1$, this is readily verified since in this case $\varrho = f_1$ for which Eqs. (10,11) give

$$\iota^{-1}(\varrho) = \overline{\times} - \times = \sum_{i=1}^{d} \big(\mathbb{1}_d - |i\rangle\langle i|\big) \otimes |ii\rangle\langle ii|.$$

If $z = 0$ then ϱ is a rank-one projection within the real algebra generated by the projections onto $|e_1\rangle$ and $|e_2\rangle$. That is,

$$\varrho = \mu^2 |e_1\rangle\langle e_1| + \frac{\nu^2}{d} |e_2\rangle\langle e_2| + \tau\big(|e_1\rangle\langle e_2| + |e_2\rangle\langle e_1|\big),$$

for some $\tau, \mu, \nu \in \mathbb{R}$. Having rank one requires vanishing determinant, which fixes $\tau^2 = \mu^2\nu^2/d$ while the remaining two parameters are constrained by the normalization $\mathrm{tr}\,[\varrho] = 1$. Please note that we choose $\tau = \mu\nu/\sqrt{d}$, since $\mu \in \mathbb{R}$, which thus includes the other case. Exploiting that ι^{-1} is again an isomorphism and that for instance $|e_1\rangle\langle e_2| = \sqrt{d}|e_1\rangle\langle e_1| \cdot |e_2\rangle\langle e_2|$, we obtain

$$\begin{aligned} \iota^{-1}(\varrho) &= \frac{1}{d}\big[\mu^2 \,\exists\!\sqsubset + \nu^2 \times + \mu\nu\big(\exists\!\sqsubset + \exists\!\sqsupset\big)\big] \\ &= \frac{1}{d} \sum_{i,k,l=1}^{d} K_i|k\rangle\langle l|K_i \otimes |k\rangle\langle l| \otimes |i\rangle\langle i|, \end{aligned}$$

which is, up to normalization, indeed the Choi matrix of the claimed channel. □

In the following section we will see that for many common disturbance measures Δ, in fact, one more parameter can be eliminated: $z = 0$ turns out to be optimal if Δ is for instance constructed from the average-case or worst-case fidelity, the worst-case Schatten $1-1$-norm or the diamond norm. This may not come as a surprise since a look at Eq. (12) reveals that for channels that correspond to elements of the unit cone we have

$$\alpha_1 = \frac{d}{d-1} z. \tag{18}$$

In other words, the contribution of the completely depolarizing channel to T_1 vanishes iff $z = 0$. This raises the question whether $z = 0$ is generally optimal under Assumptions 1 and 2. The following construction, whose only purpose is to enable the argument, shows that this is not true. Hence, without adding further assumptions about the distance measures (in particular about Δ) no further reduction is possible. On the set of quantum channels on $\mathcal{M}_d$ we define

$$\hat{\Delta}(\Phi) := \sup_{||\psi||=1} \langle\psi|\Phi\big(|\psi\rangle\langle\psi|\big)|\psi\rangle - \inf_{||\varphi||=1} \langle\varphi|\Phi\big(|\varphi\rangle\langle\varphi|\big)|\varphi\rangle.$$

This particular example yields zero disturbance for the depolarizing channel, and thus allows to show that $z = 0$ is not true in general.

Lemma 4. $\hat{\Delta}$ *satisfies Assumption 1.*

Proof. Evidently, $\hat{\Delta}(\mathrm{id}) = 0$ and $\hat{\Delta}$ is basis-independent. Convexity follows from the fact that $\hat{\Delta}$ is a supremum over linear functionals. □

Corollary 2 (Necessity of the second parameter)**.** *Let δ be any error-measure that satisfies Assumption 2 and that is faithful in the sense that $\delta = 0$ implies a perfect measurement. Then the optimal $\hat{\Delta} - \delta$-tradeoff cannot be attained within the family of channels in Eq. (17) with $z = 0$.*

Proof. Consider $\delta = 0$ in the $\hat{\Delta} - \delta$-plane. Within the full set of channels in Eq. (17) there is one that attains $\delta = 0$ while $T_1(\cdot) = \mathrm{tr}\,[\cdot]\,\mathbb{1}/d$, by choosing $\mu = 0$, $\nu = 1$ and $z = (d-1)/d$. The latter implies $\hat{\Delta}(T_1) = 0$. However, if we restrict ourselves to channels with $z = 0$, then the unique channel in Eq. (17) that achieves $\delta = 0$ has $T_1(\cdot) = \sum_i \langle i|\cdot|i\rangle |i\rangle\langle i|$ for which clearly $\hat{\Delta}(T_1) > 0$. □

Clearly, $\hat{\Delta}$ is not a 'natural' disturbance measure. For instance, it has the somewhat odd property that it vanishes for the ideal channel as well as for the projection onto the maximally mixed state. In particular, it is not faithful. Note, however, that adding the latter as an additional requirement to Assumption 1, would still not allow to eliminate the parameter z. In order to construct a new counterexample, we could just consider $\Phi \mapsto \hat{\Delta}(\Phi) + \epsilon||\Phi - \mathrm{id}||_\diamond$. This would be faithful and satisfy Assumption 1 for any $\epsilon > 0$, but for sufficiently small ϵ, the minimum Δ-value for $\delta = 0$ would, by continuity, again not be attainable for $z = 0$.

5. Optimal tradeoffs

In this section we will continue considering non-degenerate von Neumann measurements and exploit the universality theorem of the previous section in order to explicitly compute the optimal tradeoff for a variety of worst-case distance measures. We first discuss the total variational distance as a paradigm for the measurement error δ and then the fidelity and trace-norm as means for quantifying disturbance.

5.1. **Total variation.** We saw in Lemma 2 that all functionals quantifying the measurement error consistent with Assumption 2 are non-decreasing functions of the parameter α_2. In the following, we want to make this dependence explicit for one case that we regard as the most important one from an operational point of view — the worst-case total variational distance. Given two finite probability distributions p and p', their total variational distance is given by

$$||p - p'||_{TV} := \frac{1}{2}||p - p'||_1 = \frac{1}{2}\sum_i |p_i - p'_i|. \tag{19}$$

The significance of this distance stems from the fact that it displays the largest possible difference in probabilities that the two distributions assign to the same event. In our context the two probability distributions arise from an ideal and an approximate measurement on a quantum state. As $||p - p'||_{TV}$ has itself a

'worst-case interpretation' it is natural to also consider the worst case w.r.t. all quantum states and use the resulting functional as δ. That is,

$$\delta_{TV}\left(E'\right) = \sup_{\rho} \frac{1}{2} \sum_{i} \left| \mathrm{tr}\left[E'_i \rho\right] - \langle i|\rho|i\rangle \right|. \tag{20}$$

If $E'_i = T_2^*(|i\rangle\langle i|)$ with T_2 of the form in Eq. (7) so that we can regard δ_{TV} as a function of α_2, we will write $\hat{\delta}_{TV}(\alpha_2)$.

Lemma 5 (Total variational distance)**.** *In the symmetric setting discussed above, the worst-case total variational distance, regarded as a function of* α_2*, is given by* $\hat{\delta}_{TV}(\alpha_2) = \alpha_2(1-1/d)$*. Furthermore, if an instrument is parametrized by the unit cone coordinates of Eq. (16), then it leads to a worst-case total variational distance of* $(1-z-x)/2$*.*

Proof. Inserting $E'_i = T_2^*(|i\rangle\langle i|) = \alpha_2 \mathbb{1}/d + (1-\alpha_2)|i\rangle\langle i|$ into Eq. (20) we obtain

$$\begin{aligned} \hat{\delta}(\alpha_2) &= \alpha_2 \sup_{\rho} \frac{1}{2} \sum_{i} \left| \mathrm{tr}\left[\rho\big(\mathbb{1}/d - |i\rangle\langle i|\big)\right] \right| \\ &= \alpha_2 \left(1 - \frac{1}{d}\right), \end{aligned}$$

where the supremum is computed by first realizing that diagonal ρ's (i.e., classical probability distributions) suffice and then noting that convexity of the l_1-norm allows to restrict to the extreme points of the simplex of classical distributions, which all lead to the same, stated value.

The δ_{TV}-value of an instrument parametrized by the coordinates of the unit cone can then be obtained from Eq. (14) when using that $\alpha_2 = d^2 a_2$. □

An alternative way of quantifying the measurement error would be the worst-case l_∞-distance between the two probability distributions p and p'. In the present context, this measure turns out to have exactly the same value since

$$\begin{aligned} \sup_{\rho} \max_{i} \left| \mathrm{tr}\left[E'_i \rho\right] - \langle i|\rho|i\rangle \right| &= \max_{i} \left|\left| E'_i - |i\rangle\langle i| \right|\right|_\infty \\ &= \alpha_2 \left|\left| \mathbb{1}/d - |i\rangle\langle i| \right|\right|_\infty = \alpha_2 \left(1 - \frac{1}{d}\right). \end{aligned}$$

5.2. **Worst-case fidelity.** We consider the worst-case fidelity of a channel $T_1 : \mathcal{M}_d \to \mathcal{M}_d$

$$f := \inf_{||\psi||=1} \langle \psi | T_1\big(|\psi\rangle\langle\psi|\big) |\psi\rangle, \tag{21}$$

which is equal to $\inf_\rho F\big(T_1(\rho), \rho\big)^2$ due to joint concavity of the fidelity. The following states the optimal 'information-disturbance tradeoff' between f and the total variational distance:

Theorem 2 (Total variation - fidelity tradeoff)**.** *Consider a non-degenerate von Neumann measurement, given by an orthonormal basis in $\mathbb{C}^d$, and an instrument with d corresponding outcomes. Then the worst-case total variational distance δ_{TV} and the worst-case fidelity f satisfy*

$$\delta_{TV} \geq \begin{cases} \frac{1}{d}\left|\sqrt{f(d-1)} - \sqrt{1-f}\right|^2 & \text{if } f \geq \frac{1}{d}, \\ 0 & \text{if } f \leq \frac{1}{d}. \end{cases} \tag{22}$$

The inequality is tight and equality is attainable within the one-parameter family of instruments in Eq. (2) with $z = 0$.

Proof. We exploit that the optimal tradeoff is attainable for symmetric channels (Prop. 1) whose marginal is given in Eq. (7). Inserting this into the worst-case fidelity in Eq. (21) we obtain

$$\begin{aligned} f &= \min_{||\psi||=1} \left(\frac{\alpha_1}{d} + \beta_1 + \gamma_1 \sum_{i=1}^{d} |\langle\psi|i\rangle|^4 \right) \\ &= \frac{\alpha_1}{d} + \beta_1 + \begin{cases} \frac{\gamma_1}{d} & \text{if } \gamma_1 \geq 0, \\ \gamma_1 & \text{if } \gamma_1 < 0. \end{cases} \end{aligned} \tag{23}$$

Using Eqs. (12,13) together with $\gamma_1 = 1 - \alpha_1 - \beta_1$ we can express this in terms of the state ϱ. From the proof of Prop. 2 we know in addition that we can w.l.o.g. assume that $\mathrm{tr}\left[\varrho f_2\right] = 0$ and $\mathrm{tr}\left[\mathbb{1}_2 \varrho\right] = 1 - \mathrm{tr}\left[\varrho f_1\right]$. In this way, we obtain

$$f = \min\{1 - \mathrm{tr}\left[\varrho f_1\right], \langle e_1|\varrho|e_1\rangle + \mathrm{tr}\left[\varrho f_1\right]/d\}. \tag{24}$$

We aim at maximizing Eq. (24) for each value of the total variational distance, which by Lemma 5 and Eq. (14) can be expressed as

$$\delta_{TV} = 1 - \langle e_2|\varrho|e_2\rangle - \mathrm{tr}\left[\varrho f_1\right].$$

Considering the map $\varrho \mapsto \varrho + \epsilon|e_2\rangle\langle e_2| - \epsilon f_1$, $\epsilon \geq 0$, under which δ_{TV} is constant and f non-decreasing, we see that $\mathrm{tr}\left[\varrho f_1\right] = 0$ can be assumed. That is, $z = 0$ is indeed sufficient for the optimal tradeoff.

The remaining optimization problem can be solved in the equatorial plane of the Bloch sphere, where $\varrho, |e_2\rangle\langle e_2|$ and $|e_1\rangle\langle e_1|$ are represented by Bloch vectors $(x, y) =: w, (1, 0)$ and $(2/d - 1, 2\sqrt{d-1}/d) =: v$, respectively. Minimizing $\delta_{TV} = (1 - x)/2$ under the constraints

$$f \leq \frac{1}{2}\big(1 + \langle w, v\rangle\big), \quad \langle w, w\rangle = 1,$$

then amounts to a quadratic problem whose solution is stated in Eq. (22). □

5.3. **Average-case fidelity.** One prominent example of an average-case measure is the average-case fidelity of a quantum channel $T_1 : \mathcal{M}_d \to \mathcal{M}_d$

$$\bar{f} := \int_{\|\psi\|=1} \langle\psi|T_1(|\psi\rangle\langle\psi|)|\psi\rangle \, \mathrm{d}\psi. \tag{25}$$

The following theorem gives the optimal 'information-disturbance tradeoff' between the average-case fidelity and the worst-case total variational distance:

Theorem 3 (Total variation - average fidelity tradeoff)**.** *Consider a non-degenerate von Neumann measurement, given by an orthonormal basis in $\mathbb{C}^d$, and an instrument with d corresponding outcomes. Then the worst-case total variational distance δ_{TV} and the average-case fidelity $\bar{f}$ satisfy*

$$\delta_{TV} \geq \begin{cases} \frac{1}{d}\left|\sqrt{\left(\bar{f}-\frac{1}{d+1}\right)\frac{d^2-1}{d}}-\sqrt{\left(1-\bar{f}\right)\frac{d+1}{d}}\right|^2 & \text{if } \bar{f}\geq\frac{2}{d+1}, \\ 0 & \text{if } \bar{f}\leq\frac{2}{d+1}. \end{cases} \tag{26}$$

The inequality is tight and equality is attainable within the one-parameter family of instruments in Eq. (2) with $z = 0$.

Proof. We again use the fact that the optimal tradeoff is attainable for symmetric channels by Prop. 1 and its marginal is given in Eq. (7). The average-case fidelity given in Eq. (25) therefore yields

$$\begin{aligned} \bar{f} &= \int_{\|\psi\|=1} \langle\psi| \left(\alpha_1 \frac{\mathbb{1}}{d} + \beta_1 |\psi\rangle\langle\psi| + \gamma_1 \sum_{i=1}^{d} |i\rangle\langle i|\langle i|\psi\rangle\langle\psi|i\rangle\right) |\psi\rangle \, \mathrm{d}\psi \\ &= \frac{\alpha_1}{d} + \beta_1 + \gamma_1 \sum_{i=1}^{d} \int_{\|\psi\|=1} \langle\psi|i\rangle\langle i|\psi\rangle\langle i|\psi\rangle\langle\psi|i\rangle \, \mathrm{d}\psi. \end{aligned}$$

The integral can be rewritten to give

$$\begin{aligned} & \int_{\|\psi\|=1} \langle\psi\otimes\psi| \left(|i\rangle\langle i|\otimes|i\rangle\langle i|\right) |\psi\otimes\psi\rangle \, \mathrm{d}\psi \\ = & \int_{U(d)} \langle 00| \left(U\otimes U\right) \left(|i\rangle\langle i|\otimes|i\rangle\langle i|\right) \left(U\otimes U\right)^* |00\rangle \, \mathrm{d}U \\ = & \langle 00| \frac{\mathbb{1}+\mathbb{F}}{d(d+1)} |00\rangle \\ = & \frac{2}{d(d+1)}, \end{aligned}$$

where $\mathbb{F}$ is the flip operator defined as $\mathbb{F}|ij\rangle = |ji\rangle$ and dU denotes the normalized Haar measure on the unitary group $U(d)$ acting on $\mathbb{C}^d$. Together with $\gamma_1 = 1 - \alpha_1 - \beta_1$, this gives an average fidelity

$$\bar{f} = \frac{2}{d+1} - \alpha_1 \frac{d-1}{d(d+1)} + \beta_1 \frac{d-1}{d+1}.$$

Using Eqs. (12,13) we can express this in terms of the state ϱ. We can again w.l.o.g. assume that $\mathrm{tr}\left[\varrho f_2\right] = 0$ and $\mathrm{tr}\left[\mathbb{1}_2\varrho\right] = 1 - \mathrm{tr}\left[\varrho f_1\right]$ from the proof of Prop. 2. Therefore, we obtain

$$\bar{f} = \frac{1}{d+1}\left(1 + d\langle e_1|\varrho|e_1\rangle\right). \tag{27}$$

We would like to maximize Eq. (27) for each value of the worst-case total variational distance, which by Lemma 5 and Eq. (14) is

$$\delta_{TV} = 1 - \langle e_2|\varrho|e_2\rangle - \mathrm{tr}\left[\varrho f_1\right].$$

Similarly to the worst-case fidelity, we can again consider the map $\varrho \mapsto \varrho + \epsilon|e_2\rangle\langle e_2| - \epsilon f_1$, $\epsilon \geq 0$, under which δ_{TV} is constant and $\bar{f}$ non-decreasing, such that $\mathrm{tr}\left[\varrho f_1\right] = 0$ can be assumed. That is, $z = 0$ is sufficient for the optimal tradeoff.

The remaining optimization problem can be solved by realizing that $(\bar{f}(d+1)-1)/d = \langle e_1|\varrho|e_1\rangle$ and using the solution to the quadratic problem stated and solved in the worst-case fidelity tradeoff. This yields the solution stated in Eq. (26). □

5.4. **Trace norm.** The analogue of the total variational distance for density operators is (up to a factor of 2) the trace norm distance. The corresponding distance between a channel T_1 and the identity map is then given by half of the 1-to-1-norm distance

$$\Delta_{TV}(T_1) := \frac{1}{2}\sup_{\rho} ||T_1(\rho) - \rho||_1, \tag{28}$$

where the supremum is taken over all density operators. Δ_{TV} quantifies how well T_1 can be distinguished from id in a statistical experiment, if no ancillary system is allowed. For the two-parameter family of channels in Eq. (7) Δ_{TV} turns out to be a function of the worst-case fidelity f, which was defined in Eq. (21). This is in contrast to the case of general channels, which merely satisfy the Fuchs-van de Graaf inequalities

$$1 - f \leq \Delta_{TV} \leq \sqrt{1-f}. \tag{29}$$

Lemma 6. *For every channel of the form in Eq. (7), we have* $\Delta_{TV} = 1 - f$.

Proof. Due to convexity of the norm we can restrict the supremum in Eq. (28) to pure state density operators. The resulting operator $T_1\big(|\psi\rangle\langle\psi|\big) - |\psi\rangle\langle\psi|$ then has a single negative eigenvalue and vanishing trace. Hence, the trace-norm is twice the operator norm and we can write

$$\begin{aligned}
\Delta_{TV}(T_1) &= \max_{||\psi||=||\phi||=1} \langle\phi|\big[|\psi\rangle\langle\psi| - T_1\big(|\psi\rangle\langle\psi|\big)\big]|\phi\rangle \qquad (30)\\
&= \max_{||\psi||=||\phi||=1} \left[(1-\beta_1)|\langle\psi|\phi\rangle|^2 - \frac{\alpha_1}{d} - \gamma_1 \sum_{i=1}^{d} |\langle\phi|i\rangle|^2 |\langle\psi|i\rangle|^2\right]\\
&= \max_{||\psi||=||\phi||=1} \langle\psi\otimes\phi|R|\psi\otimes\phi\rangle - \frac{\alpha_1}{d},\\
&\qquad R := (1-\beta_1)\mathbb{F} - \gamma_1 \sum_{i=1}^{d} |ii\rangle\langle ii|.
\end{aligned}$$

Our aim is to prove that the maximum in Eq. (30) is attained for $\psi = \phi$ since then the Lemma follows from the definition of the worst-case fidelity f.

In order to achieve this, we exploit the symmetry properties of R, which is block-diagonal w.r.t. the decomposition of $\mathbb{C}^d \otimes \mathbb{C}^d$ into symmetric and anti-symmetric subspace. Moreover, if we denote by $P_+ := (\mathbb{1} + \mathbb{F})/2$ the projector onto the symmetric subspace, then $R \leq P_+ R P_+$. Defining $\mathcal{S}$ as the set of separable density operators and utilizing its convexity, we obtain

$$\begin{aligned} \max_{||\psi||=||\phi||=1} \langle \psi \otimes \phi | R | \psi \otimes \phi \rangle &= \max_{\rho \in \mathcal{S}} \operatorname{tr}[R\rho] \leq \max_{\rho \in \mathcal{S}} \operatorname{tr}[R P_+ \rho P_+] \\ &= \max_{\rho \in P_+ \mathcal{S} P_+} \operatorname{tr}[R\rho] \\ &= \max_{||\psi||=1} \langle \psi \otimes \psi | R | \psi \otimes \psi \rangle, \end{aligned}$$

where the last step follows from the fact that the extreme points of the convex set $P_+ \mathcal{S} P_+$ are pure, symmetric product states. □

Due to Prop. 1 we can now plug the previous Lemma into Thm. 2 and obtain:

Corollary 3 (Total variation - trace norm tradeoff)**.** *Consider a non-degenerate von Neumann measurement, given by an orthonormal basis in $\mathbb{C}^d$, and an instrument with d corresponding outcomes. Then the worst-case total variational distance δ_{TV} and its trace-norm analogue Δ_{TV} satisfy*

$$\delta_{TV} \geq \begin{cases} \frac{1}{d} \left| \sqrt{(1-\Delta_{TV})(d-1)} - \sqrt{\Delta_{TV}} \right|^2 & \text{if } \Delta_{TV} \leq 1 - \frac{1}{d}, \\ 0 & \text{if } \Delta_{TV} \geq 1 - \frac{1}{d}. \end{cases} \tag{31}$$

The inequality is tight and equality is attainable within the one-parameter family of instruments in Eq. (2) with $z = 0$.

5.5. **Diamond norm.** We treat the diamond norm separately, not only because it might be the operationally most relevant measure, but also because the corresponding tradeoff result will be proven in a more general setting: we will allow the target measurement to be a von Neumann measurement that may be degenerate. We will see that degeneracy, even if it varies among the measurement outcomes, does not affect the optimal tradeoff curve if the diamond norm is considered. For general distance measures Δ that satisfy Assumption 1 we do not expect this result to be true since, loosely speaking, they typically behave less benign w.r.t. extending the system than the diamond norm. Hence, assigning different dimensions to different measurement outcomes may, in general, affect the optimal information-disturbance relation. Before we prove that this is not the case for the tradeoff between the diamond norm and its classical counterpart, the total variational distance, let us recall its definition and basic properties.

For a hermiticity-preserving map $\Phi : \mathcal{M}_d \to \mathcal{M}_{d'}$ we define

$$||\Phi||_\diamond := \sup_\rho ||(\Phi \otimes \mathrm{id}_d)(\rho)||_1, \tag{32}$$

where the supremum is taken over all density operators in $\mathcal{M}_{d^2}$, which by convexity may be assumed to be pure. For a quantum channel $T_1 : \mathcal{M}_d \to \mathcal{M}_d$ we then define

$$\Delta_\diamond(T_1) := ||T_1 - \mathrm{id}_d||_\diamond. \tag{33}$$

$\Delta_\diamond(T_1)$ quantifies how well T_1 can be distinguished from the identity channel id in a statistical experiment, when arbitrary preparations, measurements and ancillary systems are allowed. There are two crucial properties of the diamond norm that we will exploit: 1) Monotonicity: for any quantum channel Ψ, neither $||\Psi \circ \Phi||_\diamond$ nor $||\Phi \circ \Psi||_\diamond$ can be larger than $||\Phi||_\diamond$. 2) Tensor stability: in particular, $||\Phi \otimes \mathrm{id}||_\diamond = ||\Phi||_\diamond$.

Lemma 7 (Dimension-independence of optimal tradeoff curve)**.** *Consider a von Neumann measurement with m outcomes, corresponding to m mutually orthogonal, non-zero projections of possibly different dimensions, as target. Then the optimal $\Delta_\diamond - \delta_{TV}$-tradeoff depends only on m and is independent of the dimensions of the projections.*

Proof. Let $(d_1, \ldots, d_m) \in \mathbb{N}^m$ be the dimensions of the projections (i.e., the dimensions of their ranges) and assume w.l.o.g. that d_m is the largest among them. We will consider three changes of those dimensions, namely

$$(1, \ldots, 1) \to (d_m \ldots, d_m) \to (d_1, \ldots, d_m) \to (1, \ldots, 1), \tag{34}$$

and show that in each of those three steps the accessible region in the $\Delta_\diamond - \delta_{TV}$-plane can only grow or stay the same. Since Eq. (34) describes a full circle, this means that the region, indeed, stays the same, which proves the claim of the Lemma.

For the starting point in Eq. (34) we consider an arbitrary instrument $\big(I_i : \mathcal{M}_m \to \mathcal{M}_m\big)_{i=1}^m$ that is supposed to approximate a von Neumann measurement given by $(|i\rangle\langle i|)_{i=1}^m$. From here, we construct an instrument that approximates $(|i\rangle\langle i| \otimes \mathbb{1}_{d_m})_{i=1}^m$ simply by taking $I_i \otimes \mathrm{id}_{d_m} =: \tilde{I}_i$. Then $\Delta_\diamond\big(\sum_i \tilde{I}_i\big) = \Delta_\diamond\big(\sum_i I_i\big)$ holds due to the tensor stability of the diamond norm and

$$\begin{aligned}
& \sup_\rho \sum_{i=1}^m \Big| \mathrm{tr}\Big[\rho\big(\tilde{I}_i^*(\mathbb{1}) - |i\rangle\langle i| \otimes \mathbb{1}_{d_m}\big)\Big] \Big| \\
= \; & \sup_\rho \sum_{i=1}^m \Big| \mathrm{tr}\Big[\rho\Big(\big(I_i^*(\mathbb{1}) - |i\rangle\langle i|\big) \otimes \mathbb{1}_{d_m}\Big)\Big] \Big| \\
= \; & \sup_\rho \sum_{i=1}^m \Big| \mathrm{tr}\Big[\rho\big(I_i^*(\mathbb{1}) - |i\rangle\langle i|\big)\Big] \Big|
\end{aligned}$$

shows that the value of δ_{TV} is preserved, as well.

Second and third step in Eq. (34) can be treated at once by realizing that in both cases the dimensions are pointwise non-increasing. So let us consider this scenario in general. Denote the projections corresponding to two von Neumann

measurements by $Q_i \in \mathcal{M}_D$ and $\tilde{Q}_i \in \mathcal{M}_{\tilde{D}}$ and assume that $\mathrm{tr}\,[Q_i] =: d_i \geq \tilde{d}_i := \mathrm{tr}\left[\tilde{Q}_i\right]$. Let $I_i : \mathcal{M}_D \to \mathcal{M}_D$ be the elements of an instrument that approximates the measurement in the larger space. In order to construct an instrument in the smaller space that is at least as good w.r.t. $\Delta_\diamond$ and δ, we introduce two isometries V and W as

$$\begin{aligned} V : \mathbb{C}^{\tilde{D}} \to \mathbb{C}^D \quad &\text{s.t.} \quad V^* Q_i V = \tilde{Q}_i \\ W : \mathbb{C}^D \to \mathbb{C}^k \otimes \mathbb{C}^{\tilde{D}} \quad &\text{s.t.} \quad \forall i \in \{1, \ldots, \tilde{D}\} : \; WV|i\rangle = |1\rangle \otimes |i\rangle, \end{aligned}$$

where $\{|i\rangle\}_i$ is an orthonormal basis in $\mathbb{C}^{\tilde{D}}$ and $k \in \mathbb{N}$ is sufficiently large so that W can be an isometry. The sought instrument in the smaller space can then be defined as

$$\tilde{I}_i(\rho) := \mathrm{tr}_{\mathbb{C}^k}\left[W I_i\big(V\rho V^*\big)W^*\right],$$

where $\mathrm{tr}_{\mathbb{C}^k}$ means the partial trace w.r.t. the first tensor factor. For the value of $\Delta_\diamond$ we obtain

$$\begin{aligned} \Big\|\mathrm{id} - \sum_i \tilde{I}_i\Big\|_\diamond &= \Big\|\mathrm{tr}_{\mathbb{C}^k}\left[WV \cdot V^*W^*\right] - \mathrm{tr}_{\mathbb{C}^k}\Big[W\Big(\sum_i I_i\big(V \cdot V^*\big)\Big)W^*\Big]\Big\|_\diamond \\ &\leq \Big\|V \cdot V^* - \sum_i I_i\big(V \cdot V^*\big)\Big\|_\diamond \leq \Big\|\mathrm{id} - \sum_i I_i\Big\|_\diamond, \end{aligned}$$

where we have used the monotonicity property of the diamond norm twice. Finally, using that $\tilde{I}_i^*(\mathbb{1}) = V^* I_i^*(\mathbb{1}) V$ we can show that also δ_{TV} is non-increasing when moving to the smaller space since

$$\begin{aligned} \sup_\rho \sum_i \Big|\mathrm{tr}\left[\rho\big(\tilde{I}_i^*(\mathbb{1}) - \tilde{Q}_i\big)\right]\Big| &= \sup_\rho \sum_i \Big|\mathrm{tr}\left[V\rho V^*\big(I_i^*(\mathbb{1}) - Q_i\big)\right]\Big| \\ &\leq \sup_\rho \sum_i \Big|\mathrm{tr}\left[\rho\big(I_i^*(\mathbb{1}) - Q_i\big)\right]\Big|, \end{aligned}$$

where the supremum in the first (second) line is taken over all density operators in the smaller (larger) space. □

Theorem 4 (Total variation - diamond norm tradeoff)**.** *If an instrument is considered approximating a (possibly degenerate) von Neumann measurement with m outcomes, then the worst-case total variational distance δ_{TV} and the diamond norm distance $\Delta_\diamond$ satisfy*

$$\delta_{TV} \geq \begin{cases} \frac{1}{2m}\left(\sqrt{(2-\Delta_\diamond)(m-1)} - \sqrt{\Delta_\diamond}\right)^2 & \text{if } \Delta_\diamond \leq 2 - \frac{2}{m}, \\ 0 & \text{if } \Delta_\diamond > 2 - \frac{2}{m}. \end{cases} \tag{35}$$

The inequality is tight in the sense that for every choice of the von Neumann measurement there is an instrument achieving equality.

Note: if the von Neumann measurement is non-degenerate, then equality is again attainable within the one-parameter family of instruments in Eq. (2) with $z = 0$. In the degenerate case, equality is attainable by such instruments when suitably embedded, as it is done in the proof of Lemma 7.

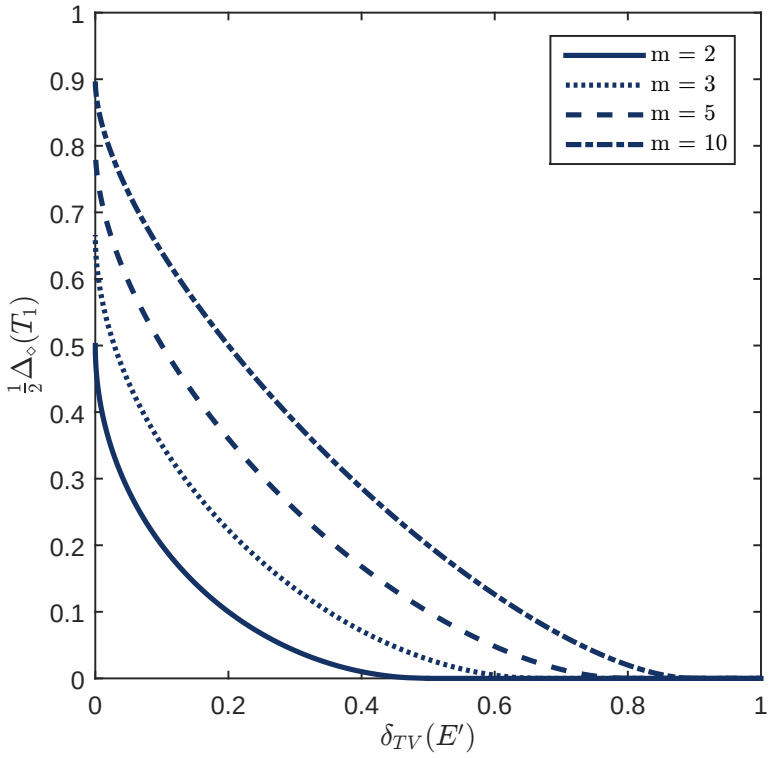


FIGURE 2. The optimal total variation - diamond norm tradeoff for different numbers of measurement outcome.

Proof. Due to Lemma 7 we can assume that the von Neumann measurement is non-degenerate and acts on a $d = m$ dimensional Hilbert space. We will prove that the accessible region stays the same when replacing $\Delta_\diamond$ with $2\Delta_{TV}$ so that the theorem follows from Cor. 3.

Since $\Delta_\diamond \geq 2\Delta_{TV}$ it suffices to show that this holds with equality for instruments that achieve the optimal $\Delta_{TV} - \delta_{TV}$ curve. Due to Eq. (18) and Cor. 3 we can restrict ourselves to symmetric channels T_1 of the form in Eq. (7) with $\alpha_1 = 0$. With $\mathcal{C}(\cdot) := \sum_{i=1}^{d} |i\rangle\langle i|\langle i| \cdot |i\rangle$ and using that $(1 - \beta_1) = \gamma_1$ we have

$$\begin{aligned}
\Delta_\diamond(T_1) &= \sup_{||\psi||=1} \left\| \left(T_1 \otimes \mathrm{id}_d - \mathrm{id}_{d^2}\right)\left(|\psi\rangle\langle\psi|\right) \right\|_1 \\
&= \sup_{||\psi||=1} \gamma_1 \left\| |\psi\rangle\langle\psi| - \left(\mathcal{C} \otimes \mathrm{id}_d\right)\left(|\psi\rangle\langle\psi|\right) \right\|_1 \\
&= 2\gamma_1 \sup_{||\psi||=||\phi||=1} |\langle\psi|\phi\rangle|^2 - \langle\phi|\left(\mathcal{C} \otimes \mathrm{id}_d\right)\left(|\psi\rangle\langle\psi|\right)|\phi\rangle \\
&= 2\gamma_1 \sup_{||\psi||=1} 1 - \langle\psi|\left(\mathcal{C} \otimes \mathrm{id}_d\right)\left(|\psi\rangle\langle\psi|\right)|\psi\rangle,
\end{aligned}$$

where the last two steps follow exactly the argumentation below Eq. (30). For the remaining optimization problem we write $|\psi\rangle = (\mathbb{1}_d \otimes X)\sum_{i=1}^d |ii\rangle$ where $X \in \mathcal{M}_d$ is s.t. $\sum_{i=1}^d \langle i|X^*X|i\rangle = ||\psi||^2 = 1$. Then

$$\langle\psi|\big(\mathcal{C}\otimes \mathrm{id}_d\big)\big(|\psi\rangle\langle\psi|\big)|\psi\rangle = \sum_{i=1}^d \big|\langle i|X^*X|i\rangle\big|^2 \geq \frac{1}{d}\Big(\sum_{i=1}^d \langle i|X^*X|i\rangle\Big)^2 = \frac{1}{d},$$

where the inequality is an application of Cauchy-Schwarz. Consequently,

$$\Delta_\diamond(T_1) \leq 2\gamma_1\Big(1-\frac{1}{d}\Big) = 2\Delta_{TV}(T_1), \tag{36}$$

where the last inequality uses that $\Delta_{TV} = 1 - f$ by Lemma 6 and $f = 1 - \gamma(1-1/d)$ by Eq. (23). As $\Delta_\diamond$ is also lower bounded by $2\Delta_{TV}$, equality has to hold in Eq. (36), which completes the proof. □

Note that equality in Eq. (36) means that entanglement assistance does not increase the distinguishability of the identity channel id and the channel T_1.

6. SDPs for general POVMs

In this section, we consider the most general case, when the target measurement E is given by an arbitrary POVM. It is then still possible to characterize the achievable region in the $\Delta - \delta$-plane as the set of solutions to some SDP if Δ and δ are convex semialgebraic. To this end, let us start with the definition of semialgebraicity.

A semialgebraic set is a set $S \subseteq \mathbb{R}^n$ defined by a finite sequence of polynomial equations and inequalities or any finite union of such sets. We mainly follow [31, 32].

Definition 1 (Semialgebraic set [32, Definition 3.1.1])**.** *A semialgebraic subset of* $\mathbb{R}^n$ *is an element of the Boolean algebra of subsets of* $\mathbb{R}^n$ *which is generated by the sets*

$$\{(x_1,\ldots,x_n) \in \mathbb{R}^n | p\,(x_1,\ldots,x_n) > 0\}\,, \quad p \in \mathbb{R}[X_1,\ldots,X_n], \tag{37}$$

where $\mathbb{R}[X_1,\ldots,X_n]$ *denotes the ring of real polynomials in the variables* X_1, …, X_n.

From this definition, it is immediately clear that sets of the form

$$\{(x_1,\ldots,x_n) \in \mathbb{R}^n | p\,(x_1,\ldots,x_n) \bullet 0\}\,,$$

where $\bullet \in \{<,>,\leq,\geq,=,\neq\}$, $p \in \mathbb{R}[X_1,\ldots,X_n]$, are semialgebraic and that the family of semialgebraic sets is closed under taking complements, finite unions and finite intersections. Moreover, by the Tarski-Seidenberg principle quantification over reals preserves the semialgebraic property [33, Appendix 1]:

Theorem 5 (Tarski-Seidenberg, quantifier elimination [34, Thm. 1])**.** *Given a finite set $\{p_i(x,z)\}_{i=1}^k$ of polynomial equalities and inequalities with variables $(x,z) \in \mathbb{R}^n \times \mathbb{R}^m$ and coefficients in $\mathbb{Q}$. Let $\phi(x,z)$ be a Boolean combination of the p_i's (using $\vee$, $\wedge$ and $\neg$) and*

$$\Psi(z) := \big(Q_1 x_1 \dots Q_n x_n : \phi(x,z)\big), \quad Q_j \in \{\exists, \forall\}\,. \tag{38}$$

Then there exists a formula $\psi(z)$ which is (i) a quantifier-free Boolean combination of finitely many polynomial (in-)equalities with rational coefficients, and (ii) equivalent in the sense

$$\forall z: \quad \big(\psi(z) \Leftrightarrow \Psi(z)\big). \tag{39}$$

Moreover, there exists an effective algorithm which constructs the quantifier-free equivalent ψ of any such formula Ψ.

Definition 2 (Semialgebraic function)**.** *Let $S_k \subseteq \mathbb{R}^{n_k}$ be non-empty semialgebraic sets, $k = 1, 2$. A function $f : S_1 \to S_2$ is said to be semialgebraic if its graph*

$$\{(x,z) \in S_1 \times S_2 | z = f(x)\} \tag{40}$$

is a semialgebraic subset of $\mathbb{R}^{n_1+n_2}$.

Using the Tarski-Seidenberg principle, Thm. 5, it is also possible to prove that the following functions, that are likely to appear in optimization problems, are semialgebraic [32, Sec. 3.1]:

- Real polynomial functions are semialgebraic.
- Compositions of semialgebraic functions are semialgebraic. Let $S_k \subseteq \mathbb{R}^{n_k}$, $k = 1, 2, 3$, be semialgebraic sets and let $f : S_1 \to S_2$ and $g : S_2 \to S_3$ be semialgebaric functions. Then their composition $g \circ f : S_1 \to S_3$ is semialgebraic.
- Let $f : S_1 \to S_2$ be a semialgebraic function, and let $A \subseteq S_1$ (resp. $B \subseteq S_2$) be a semialgebraic set. Then $f(A)$ (resp. $f^{-1}(B)$) is semialgebraic.
- Finite sums and products of semialgebraic functions are semialgebraic. Let $f_1, f_2 : S_1 \to \mathbb{R}$ be semialgebraic functions. Then $f_1 + f_2, f_1 f_2 : S_1 \to \mathbb{R}$ are semialgebraic.
- Let $f_1, f_2 : S_1 \to \mathbb{R}$ be semialgebraic functions. If $f_2^{-1}(\{0\}) \neq S_1$, then $f_1/f_2 : S_1 \backslash f_2^{-1}(\{0\}) \to \mathbb{R}$ is semialgebraic.
- Let $\mathcal{M}_n^{\mathrm{Herm}}$ denote the set of all Hermitian $n \times n$-matrices, and for $H \in \mathcal{M}_n^{\mathrm{Herm}}$ let $\lambda_k(H)$, $k \in \{1, \dots, n\}$, denote the eigenvalues of H in decreasing order. The functions $\lambda_k(\cdot) : \mathcal{M}_n^{\mathrm{Herm}} \to \mathbb{R}$ are semialgebraic.
- The singular value functions $\sigma_k : \mathbb{C}^{m \times n} \to [0, \infty)$, $1 \leq k \leq \min\{m, n\}$ are semialgebraic.

For the last point, we identify a subset of $\mathbb{C}^n$ with a subset of $\mathbb{R}^{2n}$ by separating the real and imaginary parts. Therefore, the notion of a semialgebraic subset of $\mathbb{C}^{m \times n}$ is well defined.

Furthermore, one can show the following regarding the supremum or infimum of a function:

Lemma 8 ([32, Cor. 3.1.15])**.** *Let $S_k \subseteq \mathbb{R}^{n_k}$ be non-empty semialgebraic sets, $k = 1, 2$, and $f : S_1 \times S_2 \to \mathbb{R}$ a semialgebraic function. Then $\hat{f}, \check{f} : S_1 \to \mathbb{R} \cup \{-\infty, \infty\}$,*

$$\hat{f}(x) := \sup_{y \in S_2} f(x, y) \quad \text{and} \tag{41}$$

$$\check{f}(x) := \inf_{y \in S_2} f(x, y) \tag{42}$$

are both semialgebraic.

Using the fact that singular value functions are semialgebraic, it is immediately possible to show the following corollary:

Corollary 4 ([32, Cor. 3.1.24])**.** *The Schatten p-norms $\|\cdot\|_p : \mathbb{C}^{n \times m} \to [0, \infty)$ are semialgebraic for all $p \in [1, \infty) \cap \mathbb{Q}$ and $p = \infty$.*

Proof. Please see [32, Cor. 3.1.23 and 3.1.19] for a full proof. The main idea is to establish that the function $x \mapsto x^{p/q}$, with $x > 0$ and p, q positive integers, is semialgebraic. Its graph is

$$\begin{aligned} & \left\{(x, z) \in \mathbb{R}_+^2 \middle| z = x^{\frac{p}{q}}\right\} \\ = \; & \left\{(x, z) \in \mathbb{R}^2 \middle| z^q - x^p = 0\right\} \cap \mathbb{R}_+^2, \end{aligned}$$

which is semialgebraic. □

Corollary 5. *The Schatten p-to-q norm-distances of a quantum channel $\Phi \in \mathcal{T}_d$ to the identity channel*

$$\Phi \mapsto ||\Phi - \mathrm{id}||_{p \to q, n} := \sup_{\rho \in \mathcal{S}_{dn}} \frac{||(\Phi - \mathrm{id}) \otimes \mathrm{id}_n(\rho)||_q}{||\rho||_p}, \quad n \in \mathbb{N},$$

are semialgebraic for all $p, q \in [1, \infty) \cap \mathbb{Q}$ and $p, q = \infty$.

The worst-case fidelity distance of a quantum channel $\Phi \in \mathcal{T}_d$ to the identity channel

$$\Phi \mapsto \inf_{\rho \in \mathcal{S}_d} F\left(\Phi(\rho), \rho\right)^2$$

is semialgebraic.

The worst-case l_p-distances of a POVM $E' \in \mathcal{E}_{d,m}$ to the target POVM $E \in \mathcal{E}_{d,m}$

$$E' \mapsto \sup_{\rho \in \mathcal{S}_d} || \left(\mathrm{tr}[\rho E_i] - \mathrm{tr}[\rho E_i']\right)_{i=1}^m ||_p,$$

are semialgebraic for all $p \in [1, \infty) \cap \mathbb{Q}$ and $p = \infty$.

Proof. Given that the set of all quantum states is semialgebraic [34, Lemma 1], Cor. 4 together with Lemma 8 immediately yields the statements. □

In particular, the special case of the *diamond norm* $||\cdot||_\diamond := ||\cdot||_{1 \to 1, d}$, which we discuss in more detail below, and its dual, the *cb-norm* (with $p = q = \infty, n = d$) are semialgebraic.

Theorem 6 (Helton-Nie conjecture in dimension two [35, Thm. 6.8.])**.** *Every convex semialgebraic subset S of $\mathbb{R}^2$ is the feasible set of a SDP. That is, it can be written as*

$$S = \left\{ \xi \in \mathbb{R}^2 \middle| \exists \eta \in \mathbb{R}^m : A + \sum_{i=1}^{2} \xi_i B_i + \sum_{j=1}^{m} \eta_j C_j \geq 0 \right\}, \tag{43}$$

where $m \geq 0$ and A, B_i as well as C_j are real symmetric matrices of the same size.

The proof of the Helton-Nie conjecture in dimension two can be found in [35].[4] The main observation of this section is a consequence of the previous theorem and the following simple Lemma:

Lemma 9. *If Δ and δ are both semialgebraic, then the accessible region in the $\Delta - \delta$-plane is a semialgebraic set.*

Proof. Let us denote the accessible region in the $\Delta - \delta$-plane by S, i.e.,

$$S = \left\{ x \in \mathbb{R}^2 \middle| \exists I = \{I_i\}_{i=1}^m : x_1 = \Delta\left(\sum_{i=1}^{m} I_i\right) \wedge x_2 = \delta\left((I_i^*(\mathbb{1}))_{i=1}^m\right) \right\}.$$

First note that the set of instruments is semialgebraic. The maps $I \mapsto \sum_{i=1}^m I_i$ as well as $I \mapsto (I_i^*(\mathbb{1}))_{i=1}^m$ are algebraic and therefore semialgebraic [31]. Given that the composition of two semialgebraic maps is semialgebraic [31, Prop. 2.2.6 (i)] and that the image of a semialgebraic set under a semialgebraic map is semialgebraic [31, Prop. 2.2.7.], $\Delta\left(\sum_{i=1}^m I_i\right)$ as well as $\delta\left((I_i^*(\mathbb{1}))_{i=1}^m\right)$ are semialgebraic. Using the Tarski-Seidenberg principle, Thm. 5, we arrive at the claim. □

Theorem 7 (SDP solution for arbitrary target measurements)**.** *If Δ and δ are both convex and semialgebraic, then the accessible region in the $\Delta - \delta$-plane is the feasible set of a SDP.*

Proof. If Δ and δ are convex and semialgebraic, then the whole region in the $\Delta - \delta$-plane that is accessible by quantum instruments is a convex semialgebraic subset of $\mathbb{R}^2$ by Lemma 9. By Thm. 6, it must thus be the feasible set of a SDP. □

In particular, if we consider a Schatten p-to-q-norm distance, with p and q rational, to describe the disturbance caused to the quantum system and a worst-case l_p-norm distances, with rational p, to quantify the measurement error, the accessible region in the $\Delta - \delta$-plane is the feasible set of a SDP.

Unfortunately, we do not know how to make the results of [35] constructive. That is while Thm. 7 proves the existence of a SDP, we do not have a way of making the SDP explicit.

[4] The conjecture for larger dimensions was shown to be false in general in [36].

SDP for the diamond norm tradeoff. We now explicitly state the SDP yielding the optimal tradeoff curve in the case of a general POVM for the worst-case l_∞-distance and the diamond norm. This particular example does not rely on the general result of Thm 7, since the l_∞-norm as well as the diamond norm are already well-suited to SDP formulation. Please note that on the measurement error side, we use the worst-case l_∞-norm to quantify the distance between the two probability distributions,

$$\delta_{l_\infty} := \sup_\rho \max_i \left| \mathrm{tr}\left[E_i'\rho\right] - \mathrm{tr}\left[E_i\rho\right] \right|. \tag{44}$$

In this setting the optimization problem, quantifying the information-disturbance tradeoff, is given as:
Compute for a given target POVM $E = \{E_i\}_{i=1}^m$ and $\lambda \in [0,1]$

$$\begin{aligned} \nu(E,\lambda) := \quad & \min_{\{I_i\}_{i=1}^m} \left\| \sum_{i=1}^m I_i - \mathrm{id} \right\|_\diamond \\ \text{such that} \quad & \left\| I_i^*(\mathbb{1}) - E_i \right\|_\infty \leq \lambda \quad \forall i, \\ & I_i \text{ is completely positive} \quad \forall i \text{ and} \\ & \sum_{i=1}^m I_i^*(\mathbb{1}) = \mathbb{1}. \end{aligned} \tag{45}$$

In the following, let us the define the Choi matrix for any linear map $T : \mathcal{M}_d \to \mathcal{M}_{d'}$ as

$$J(T) := (T \otimes \mathrm{id}_d) \left(\sum_{i,j=1}^d |ii\rangle\langle jj| \right). \tag{46}$$

Theorem 8. *For a given target POVM $E = \{E_i \in \mathcal{M}_d\}_{i=1}^m$ and $\lambda \in [0,1]$, the optimization problem $\nu(E,\lambda)$ given in Eq. (45), can be formulated as a SDP (ϕ, C, D), where $\phi : \mathcal{M}_{\hat{d}} \to \mathcal{M}_{\check{d}}$ is a hermiticity preserving map, $C = C^* \in \mathcal{M}_{\hat{d}}$ and $D = D^* \in \mathcal{M}_{\check{d}}$, with dimensions $\hat{d} = (m+4)d^2 + 2(m+2)d$ and $\check{d} = 2 + (m+2)d^2$. The primal and the dual SDP problem are given as follows:*

Primal SDP problem		Dual SDP problem	
maximize	$\mathrm{tr}\left[CX\right]$	*minimize*	$\mathrm{tr}\left[DY\right]$
subject to	$\phi(X) = D$	*subject to*	$\phi^*(Y) \geq C$
	$X \geq 0$		$Y = Y^\dagger$

where the hermiticity preserving map $\phi : \mathcal{M}_{\hat{d}} \to \mathcal{M}_{\check{d}}$ is

$$\begin{aligned} \phi(X) \quad = \quad & \mathrm{tr}\left[w_0\right] \oplus \mathrm{tr}\left[w_1\right] \oplus \left(A + Z_0 - \mathbb{1} \otimes w_0\right) \oplus \left(B + Z_1 - \mathbb{1} \otimes w_1\right) \oplus \\ & \bigoplus_{i=1}^m \left(M + M^* + \mathbb{1} \otimes \left(F_i - \widetilde{F}_i\right) + G_i + \mathbb{1} \otimes \left(H - \widetilde{H}\right) \right), \end{aligned} \tag{47}$$

with

$$X := \begin{pmatrix} A & M \\ M^* & B \end{pmatrix} \oplus w_0 \oplus w_1 \oplus Z_0 \oplus Z_1 \oplus \bigoplus_{i=1}^{m} F_i \oplus \bigoplus_{i=1}^{m} \widetilde{F}_i \oplus \bigoplus_{i=1}^{m} G_i \oplus H \oplus \widetilde{H}. \tag{48}$$

The adjoint of the map ϕ is

$$\begin{aligned} \phi^*(Y) := & \begin{pmatrix} Y_0 & \sum_{i=1}^m J(I_i) \\ \sum_{i=1}^m J(I_i) & Y_1 \end{pmatrix} \oplus (\lambda_0 \mathbb{1} - \mathrm{tr}_1[Y_0]) \oplus (\lambda_1 \mathbb{1} - \mathrm{tr}_1[Y_1]) \oplus \\ & Y_0 \oplus Y_1 \oplus \bigoplus_{i=1}^{m} \mathrm{tr}_1[J(I_i)] \oplus \bigoplus_{i=1}^{m} -\mathrm{tr}_1[J(I_i)] \oplus \bigoplus_{i=1}^{m} J(I_i) \oplus \\ & \sum_{i=1}^{m} \mathrm{tr}_1[J(I_i)] \oplus -\sum_{i=1}^{m} \mathrm{tr}_1[J(I_i)], \end{aligned} \tag{49}$$

with

$$Y := \lambda_0 \oplus \lambda_1 \oplus Y_0 \oplus Y_1 \oplus \bigoplus_{i=1}^{m} J(I_i). \tag{50}$$

Furthermore,

$$D := \frac{1}{2} \oplus \frac{1}{2} \oplus 0 \oplus 0 \oplus \bigoplus_{i=1}^{m} 0 \tag{51}$$

and

$$\begin{aligned} C := & \begin{pmatrix} 0 & J(\mathrm{id}) \\ J(\mathrm{id}) & 0 \end{pmatrix} \oplus 0 \oplus 0 \oplus 0 \oplus 0 \oplus \bigoplus_{i=1}^{m} \left(-\lambda \mathbb{1} + E_i^T\right) \oplus \\ & \bigoplus_{i=1}^{m} \left(-\lambda \mathbb{1} - E_i^T\right) \oplus \bigoplus_{i=1}^{m} 0 \oplus \mathbb{1} \oplus -\mathbb{1}. \end{aligned} \tag{52}$$

Proof. The diamond norm can be expressed as a SDP itself [37, 38],

$$\begin{aligned} \left\| \mathrm{id} - \sum_{i=1}^{m} I_i \right\|_\diamond = & \min_{Y_0, Y_1 \in \mathcal{M}_d \otimes \mathcal{M}_d} \frac{1}{2} \left[\left\| \mathrm{tr}_1[Y_0] \right\|_\infty + \left\| \mathrm{tr}_1[Y_1] \right\|_\infty \right] \\ & \text{such that} \quad \begin{pmatrix} Y_0 & J\left(\mathrm{id} - \sum_{i=1}^m I_i\right) \\ J\left(\mathrm{id} - \sum_{i=1}^m I_i\right) & Y_1 \end{pmatrix} \geq 0 \quad \text{and} \\ & Y_0, Y_1 \geq 0, \end{aligned}$$

where tr_1 denotes the partial trace over the first system. Using Watrous SDP for the diamond norm in the form of [38, p. 11] gives

$$\begin{aligned}
\nu(E,\lambda) = \text{minimize} \quad & \frac{1}{2}\left[\lambda_0 + \lambda_1\right] \\
\text{such that} \quad & \begin{pmatrix} Y_0 & \sum_{i=1}^m J(I_i) \\ \sum_{i=1}^m J(I_i) & Y_1 \end{pmatrix} \geq \begin{pmatrix} 0 & J(\mathrm{id}) \\ J(\mathrm{id}) & 0 \end{pmatrix} \\
& \lambda_0 \mathbb{1} - \mathrm{tr}_1\left[Y_0\right] \geq 0 \\
& \lambda_1 \mathbb{1} - \mathrm{tr}_1\left[Y_1\right] \geq 0 \\
& Y_0, Y_1 \geq 0 \\
& \mathrm{tr}_1\left[J(I_i)\right] \geq -\lambda \mathbb{1} + E_i^T \quad \forall i \\
& -\mathrm{tr}_1\left[J(I_i)\right] \geq -\lambda \mathbb{1} - E_i^T \quad \forall i \\
& J(I_i) \geq 0 \quad \forall i \\
& \sum_{i=1}^m \mathrm{tr}_1\left[J(I_i)\right] \geq \mathbb{1} \\
& -\sum_{i=1}^m \mathrm{tr}_1\left[J(I_i)\right] \geq -\mathbb{1}.
\end{aligned}$$

We would like to write this as a SDP in the form

$$\begin{aligned}
\text{minimize} \quad & \mathrm{tr}\left[DY\right] \\
\text{subject to} \quad & \phi^*(Y) \geq C, \\
& Y = Y^\dagger.
\end{aligned}$$

Collecting all variables that we optimize over yields $Y \in \mathbb{C} \oplus \mathbb{C} \oplus \mathcal{M}_{d^2} \oplus \mathcal{M}_{d^2} \oplus \bigoplus_{i=1}^m \mathcal{M}_{d^2}$ as

$$Y := \lambda_0 \oplus \lambda_1 \oplus Y_0 \oplus Y_1 \oplus \bigoplus_{i=1}^m J(I_i).$$

Furthermore, we set $D \in \mathbb{C} \oplus \mathbb{C} \oplus \mathcal{M}_{d^2} \oplus \mathcal{M}_{d^2} \oplus \bigoplus_{i=1}^m \mathcal{M}_{d^2}$ as

$$D := \frac{1}{2} \oplus \frac{1}{2} \oplus 0_{d^2} \oplus 0_{d^2} \oplus \bigoplus_{i=1}^m 0_{d^2}.$$

Similarly, set $\phi^*(Y) \in \mathcal{M}_{2d^2} \oplus \mathcal{M}_d \oplus \mathcal{M}_d \oplus \mathcal{M}_{d^2} \oplus \mathcal{M}_{d^2} \oplus \bigoplus_{i=1}^m \mathcal{M}_d \oplus \bigoplus_{i=1}^m \mathcal{M}_d \oplus \bigoplus_{i=1}^m \mathcal{M}_{d^2} \oplus \bigoplus_{i=1}^m \mathcal{M}_d \oplus \bigoplus_{i=1}^m \mathcal{M}_d$ to be

$$\begin{aligned}\phi^*(Y) &:= \begin{pmatrix} Y_0 & \sum_{i=1}^m J(I_i) \\ \sum_{i=1}^m J(I_i) & Y_1 \end{pmatrix} \oplus (\lambda_0 \mathbb{1}_d - \mathrm{tr}_1[Y_0]) \oplus (\lambda_1 \mathbb{1}_d - \mathrm{tr}_1[Y_1]) \oplus \\ & Y_0 \oplus Y_1 \oplus \bigoplus_{i=1}^m \mathrm{tr}_1[J(I_i)] \oplus \bigoplus_{i=1}^m -\mathrm{tr}_1[J(I_i)] \oplus \bigoplus_{i=1}^m J(I_i) \oplus \\ & \sum_{i=1}^m \mathrm{tr}_1[J(I_i)] \oplus -\sum_{i=1}^m \mathrm{tr}_1[J(I_i)],\end{aligned}$$

and we define $C \in \mathcal{M}_{2d^2} \oplus \mathcal{M}_d \oplus \mathcal{M}_d \oplus \mathcal{M}_{d^2} \oplus \mathcal{M}_{d^2} \oplus \bigoplus_{i=1}^m \mathcal{M}_d \oplus \bigoplus_{i=1}^m \mathcal{M}_d \oplus \bigoplus_{i=1}^m \mathcal{M}_{d^2} \oplus \bigoplus_{i=1}^m \mathcal{M}_d \oplus \bigoplus_{i=1}^m \mathcal{M}_d$ as

$$\begin{aligned}C &:= \begin{pmatrix} 0 & J(\mathrm{id}) \\ J(\mathrm{id}) & 0 \end{pmatrix} \oplus 0_d \oplus 0_d \oplus 0_{d^2} \oplus 0_{d^2} \oplus \bigoplus_{i=1}^m \left(-\lambda \mathbb{1} + E_i^T\right) \oplus \\ & \bigoplus_{i=1}^m \left(-\lambda \mathbb{1} - E_i^T\right) \oplus \bigoplus_{i=1}^m 0_{d^2} \oplus \mathbb{1}_d \oplus -\mathbb{1}_d.\end{aligned}$$

Therefore, the optimization problem $\nu(E, \lambda)$ is a SDP indeed. In order to state the dual SDP problem, define $X \in \mathcal{M}_{2d^2} \oplus \mathcal{M}_d \oplus \mathcal{M}_d \oplus \mathcal{M}_{d^2} \oplus \mathcal{M}_{d^2} \oplus \bigoplus_{i=1}^m \mathcal{M}_d \oplus \bigoplus_{i=1}^m \mathcal{M}_d \oplus \bigoplus_{i=1}^m \mathcal{M}_{d^2} \oplus \bigoplus_{i=1}^m \mathcal{M}_d \oplus \bigoplus_{i=1}^m \mathcal{M}_d$ to be

$$\begin{aligned}X &:= \begin{pmatrix} A & M \\ M^* & B \end{pmatrix} \oplus w_0 \oplus w_1 \oplus Z_0 \oplus Z_1 \oplus \\ & \bigoplus_{i=1}^m F_i \oplus \bigoplus_{i=1}^m \widetilde{F}_i \oplus \bigoplus_{i=1}^m G_i \oplus H \oplus \widetilde{H}.\end{aligned}$$

Using the fact that $\mathrm{tr}[\phi^*(Y)X] = \mathrm{tr}[Y\phi(X)]$ lets us construct ϕ such that $\phi(X) \in \mathbb{C} \oplus \mathbb{C} \oplus \mathcal{M}_{d^2} \oplus \mathcal{M}_{d^2} \oplus \bigoplus_{i=1}^m \mathcal{M}_{d^2}$ is

$$\begin{aligned}\phi(X) &= \mathrm{tr}[w_0] \oplus \mathrm{tr}[w_1] \oplus (A + Z_0 - \mathbb{1} \otimes w_0) \oplus (B + Z_1 - \mathbb{1} \otimes w_1) \oplus \\ & \bigoplus_{i=1}^m \left(M + M^* + \mathbb{1} \otimes \left(F_i - \widetilde{F}_i\right) + G_i + \mathbb{1} \otimes \left(H - \widetilde{H}\right)\right).\end{aligned}$$

□

Proposition 3. *For the above SDP (ϕ, C, D) the Slater-type strong duality holds, such that*

$$\sup_X \mathrm{tr}[CX] = \inf_Y \mathrm{tr}[DY]. \tag{53}$$

Proof. There is an interior point $X > 0$ that fulfills $\phi(X) = D$ and a $Y = Y^*$ such that $\phi^*(Y) \geq C$. By Slater's theorem strong duality holds for the SDP (ϕ, C, D). □

Using Thm. 8 it is therefore possible to explicitly state the SDP that yields the information-disturbance tradeoff curve for any general POVM in the case where the measurement-error is quantified by the worst-case l_∞-distance and the disturbance is quantified by the diamond norm.

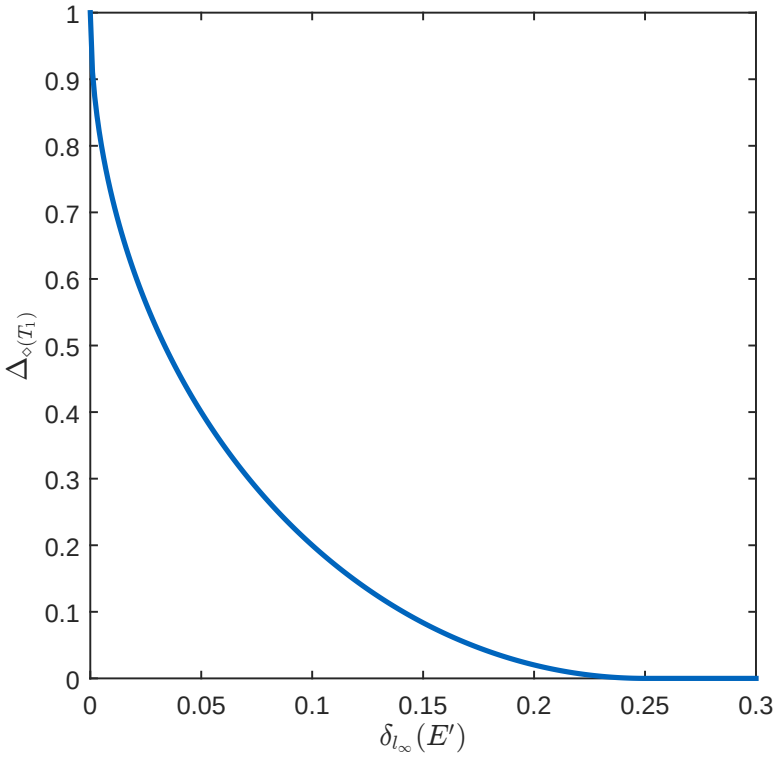


FIGURE 3. The information-disturbance tradeoff for a qubit SIC POVM target measurement.

SIC POVM. As it is a prominent application in various fields in quantum information theory, this section analyzes the example of a symmetric, informationally complete (SIC) POVM as target measurement. A SIC POVM is defined by a set of d^2 subnormalized rank-1 projectors $\{P_i/d\}_{i=1}^{d^2}$, which have equal pairwise Hilbert-Schmidt inner products, $\operatorname{tr}[P_iP_j]/d^2 = 1/d^2(d+1)$ for $i \neq j$. Figure 3 and 4 show the information-disturbance tradeoff for a qubit SIC POVM and qutrit SIC POVM as target measurement respectively. In two dimensions, we considered the following SIC POVM represented by the four Bloch vectors $(0,0,1)$, $(2\sqrt{2}/3, 0, -1/3)$, $(-\sqrt{2}/3, \sqrt{2/3}, -1/3)$ and $(-\sqrt{2}/3, -\sqrt{2/3}, -1/3)$. In dimension three, the nine explicit (unnormalized) vectors of the SIC POVM under consideration are $(0,1,-1)$, $(0,1,-\eta)$, $(0,1,-\eta^2)$, $(-1,0,1)$, $(-\eta,0,1)$, $(-\eta^2,0,1)$, $(1,-1,0)$, $(1,-\eta,0)$ and $(1,-\eta^2,0)$

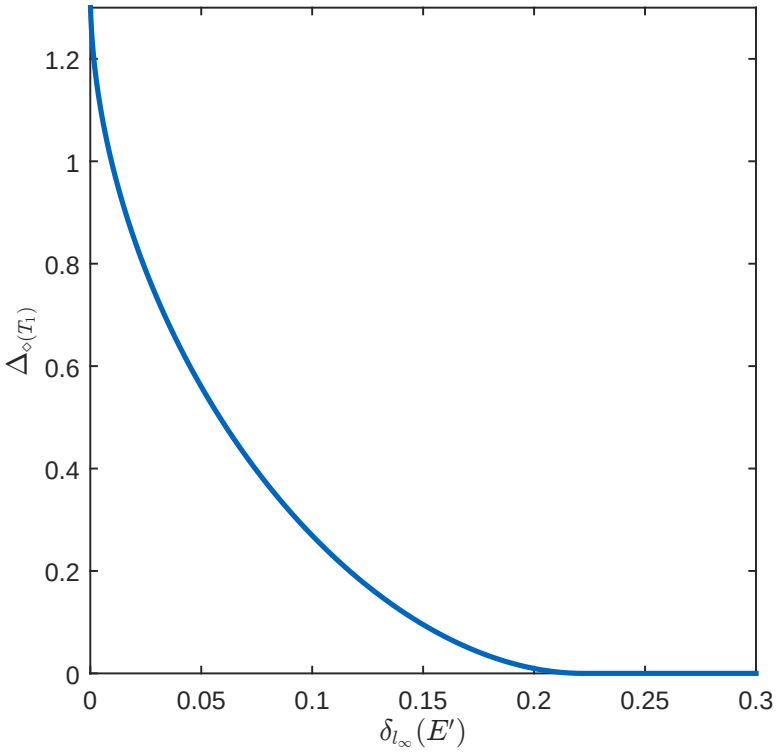


FIGURE 4. The information-disturbance tradeoff for a qutrit SIC POVM target measurement.

with $\eta = \exp 2\pi i/3$. To solve the SDP stated in Thm. 8 for this particular example, we used cvx, a package for specifying and solving convex programs [39, 40] in MATLAB [41].

The solution of the SDP is compared to an instrument similar to the one found in Thm. 1 consisting of an inherit POVM $E' = tE + (1-t)\mathbb{1}/d$, $t \in [0,1]$, together with the Lüders channel. The symmetry of the SIC POVM most likely leads to this agreement. However, further investigation would be necessary to get a better understanding of this observation.

ACKNOWLEDGMENT

The authors would like to thank Teiko Heinosaari for many useful comments. AKHs work is supported by the Elite Network of Bavaria through the PhD program of excellence *Exploring Quantum Matter*. This research was supported in part by the National Science Foundation under Grant No. NSF PHY11-25915.

Appendix

Proof that average- and worst-case construction satisfy Assumption 1 and Assumption 2.

Lemma 10. *If* $\tilde{\Delta} : \mathcal{S}_d \times \mathcal{S}_d \to [0, \infty]$ *satisfies*

(i) $\tilde{\Delta}(\rho, \rho) = 0$,
(ii) convexity in its first argument and
(iii) unitary invariance,

then the worst-case as well as the average-case construction

$$\begin{aligned} \Delta_\infty(\Phi) &:= \sup_{\rho \in S} \tilde{\Delta}\left(\Phi(\rho), \rho\right) \quad \textit{and} \\ \Delta_\mu(\Phi) &:= \int_{\mathcal{S}_d} \tilde{\Delta}\left(\Phi(\rho), \rho\right) \, \mathrm{d}\mu(\rho), \end{aligned}$$

with μ *a unitarily invariant measure on* $\mathcal{S}_d$ *and* $S \subseteq \mathcal{S}_d$ *a unitarily closed subset, satisfy Assumption 1.*

Proof. Let $\tilde{\Delta} : \mathcal{S}_d \times \mathcal{S}_d \to [0, \infty]$ be such that it

(i) satisfies $\tilde{\Delta}(\rho, \rho) = 0$,
(ii) is convex in its first argument, i.e., for any quantum state $\sigma, \sigma', \rho \in \mathcal{S}_d$

$$\tilde{\Delta}\left(\lambda\sigma + (1-\lambda)\sigma', \rho\right) \leq \lambda\tilde{\Delta}\left(\sigma, \rho\right) + (1-\lambda)\tilde{\Delta}\left(\sigma', \rho\right) \quad \forall \lambda \in [0, 1]\,,$$

(iii) and is unitarily invariant, i.e., for any quantum state $\sigma, \rho \in \mathcal{S}_d$

$$\tilde{\Delta}\left(U^*\sigma U, U^*\rho U\right) = \tilde{\Delta}\left(\sigma, \rho\right) \quad \forall \text{ unitaries } U \in \mathcal{M}_d.$$

Then its worst case Δ_∞ satisfies

(a) $\Delta_\infty(\mathrm{id}) = 0$, since

$$\Delta_\infty(\mathrm{id}) = \sup_{\rho \in S} \tilde{\Delta}\left(\mathrm{id}(\rho), \rho\right) = \sup_{\rho \in S} \tilde{\Delta}\left(\rho, \rho\right) = 0,$$

(b) is convex, i.e., for every quantum channel $\Phi, \Phi' \in \mathcal{T}_d$

$$\Delta_\infty\left(\lambda\Phi + (1-\lambda)\Phi'\right) \leq \lambda\Delta_\infty\left(\Phi\right) + (1-\lambda)\Delta_\infty\left(\Phi'\right) \quad \forall \lambda \in [0, 1]\,,$$

because

$$\begin{aligned} \Delta_\infty\left(\lambda\Phi + (1-\lambda)\Phi'\right) &= \sup_{\rho \in S} \tilde{\Delta}\left(\lambda\Phi(\rho) + (1-\lambda)\Phi'(\rho), \rho\right) \\ &\leq \lambda \sup_{\rho \in S} \tilde{\Delta}\left(\Phi(\rho), \rho\right) + (1-\lambda) \sup_{\rho \in S} \left(\Phi'(\rho), \rho\right) \\ &= \lambda\Delta_\infty\left(\Phi\right) + (1-\lambda)\Delta_\infty\left(\Phi'\right), \end{aligned}$$

(c) and is basis-independent, i.e., for every unitary $U \in \mathcal{M}_d$ and every channel $\Phi \in \mathcal{T}_d$, we have that

$$\Delta_\infty\left(U\Phi\left(U^* \cdot U\right)U^*\right) = \Delta_\infty(\Phi),$$

since

$$\begin{aligned}\Delta_\infty\left(U\Phi\left(U^*\cdot U\right)U^*\right) &= \sup_{\rho\in S}\tilde{\Delta}\left(U\Phi\left(U^*\rho U\right)U^*,\rho\right)\\ &= \sup_{\rho\in S}\tilde{\Delta}\left(\Phi\left(U^*\rho U\right),U^*\rho U\right)\\ &= \sup_{\rho\in S}\tilde{\Delta}\left(\Phi\left(\rho\right),\rho\right)\\ &= \Delta_\infty(\Phi).\end{aligned}$$

The average case Δ_μ satisfies

(a) $\Delta_\mu(\mathrm{id}) = 0$, since

$$\Delta_\mu(\mathrm{id}) = \int_{\mathcal{S}_d}\tilde{\Delta}\left(\mathrm{id}(\rho),\rho\right)\,\mathrm{d}\mu(\rho) = \int_{\mathcal{S}_d}\tilde{\Delta}\left(\rho,\rho\right)\,\mathrm{d}\mu(\rho) = 0,$$

(b) is convex, i.e., for every quantum channel $\Phi,\Phi'\in\mathcal{T}_d$

$$\Delta_\mu\left(\lambda\Phi+(1-\lambda)\Phi'\right)\leq\lambda\Delta_\mu\left(\Phi\right)+(1-\lambda)\Delta_\mu\left(\Phi'\right)\quad\forall\lambda\in[0,1]\,,$$

because

$$\begin{aligned}\Delta_\mu\left(\lambda\Phi+(1-\lambda)\Phi'\right) &= \int_{\mathcal{S}_d}\tilde{\Delta}\left(\lambda\Phi(\rho)+(1-\lambda)\Phi'(\rho),\rho\right)\,\mathrm{d}\mu(\rho)\\ &\leq \lambda\int_{\mathcal{S}_d}\tilde{\Delta}\left(\Phi(\rho),\rho\right)\,\mathrm{d}\mu(\rho)+(1-\lambda)\int_{\mathcal{S}_d}\tilde{\Delta}\left(\Phi'(\rho),\rho\right)\,\mathrm{d}\mu(\rho)\\ &= \lambda\Delta_\mu\left(\Phi\right)+(1-\lambda)\Delta_\mu\left(\Phi'\right),\end{aligned}$$

(c) and is basis-independent, i.e., for every unitary $U\in\mathcal{M}_d$ and every channel $\Phi\in\mathcal{T}_d$, we have that

$$\Delta_\mu\left(U\Phi\left(U^*\cdot U\right)U^*\right) = \Delta_\mu(\Phi),$$

since

$$\begin{aligned}\Delta_\mu\left(U\Phi\left(U^*\cdot U\right)U^*\right) &= \int_{\mathcal{S}_d}\tilde{\Delta}\left(U\Phi\left(U^*\rho U\right)U^*,\rho\right)\,\mathrm{d}\mu(\rho)\\ &= \int_{\mathcal{S}_d}\tilde{\Delta}\left(\Phi\left(U^*\rho U\right),U^*\rho U\right)\,\mathrm{d}\mu(\rho)\\ &= \int_{\mathcal{S}_d}\tilde{\Delta}\left(\Phi\left(\rho\right),\rho\right)\,\mathrm{d}\mu(\rho)\\ &= \Delta_\mu(\Phi),\end{aligned}$$

where we have used the fact that μ is a unitarily invariant measure on $\mathcal{S}_d$.

The worst-case construction as well as the average-case construction therefore satisfy Assumption 1 as claimed. □

Lemma 11. *If $\tilde{\delta}:\mathcal{P}_d\times\mathcal{P}_d\to[0,\infty]$ on the space of probability distributions $\mathcal{P}_d := \left\{q\in\mathbb{R}^d\middle|\sum_{i=1}^d q_i = 1\wedge\forall i: q_i\geq 0\right\}$ applied to the target distribution $p_i := \langle i|\rho|i\rangle$ and the actually measured distribution $p'_i := \mathrm{tr}\left[\rho E'_i\right]$ satisfies*

(i) $\tilde{\delta}(q,q) = 0$,
(ii) *convexity in its second argument and*
(iii) *invariance under joint permutations,*

then the worst-case as well as the average-case construction

$$\begin{aligned}\delta_\infty(E') &:= \sup_{\rho\in S} \tilde{\delta}(p,p'),\\ \delta_\mu(E') &:= \int_{\mathcal{S}_d} \tilde{\delta}(p,p')\,\mathrm{d}\mu(\rho),\end{aligned}$$

both satisfy Assumption 2.

Proof. Let $\tilde{\delta} : \mathcal{P}_d \times \mathcal{P}_d \to [0,\infty]$ be such that it

(i) satisfies $\tilde{\delta}(q,q) = 0$,
(ii) is convex in its second argument, i.e., for every probability distribution $p, q, q' \in \mathcal{P}_d$

$$\tilde{\delta}(p, \lambda q + (1-\lambda)q') \leq \lambda\tilde{\delta}(p,q) + (1-\lambda)\tilde{\delta}(p,q') \quad \forall\lambda \in [0,1],$$

(iii) and invariant under joint permutations,i.e., for every quantum state $\rho \in \mathcal{S}_d$ and every POVM $E, E' \in \mathcal{E}_d$

$$\tilde{\delta}\left(\left(\mathrm{tr}\left[\rho U_\pi^* E_{\pi(i)} U_\pi\right]\right)_{i=1}^d, \left(\mathrm{tr}\left[\rho U_\pi^* E'_{\pi(i)}\right] U_\pi\right)_{i=1}^d\right) = \tilde{\delta}\left(\left(\mathrm{tr}\left[\rho E_i\right]\right)_{i=1}^d, \left(\mathrm{tr}\left[\rho E'_i\right]\right)_{i=1}^d\right).$$

Then its worst case δ_∞ satisfies

(a) $\delta_\infty\left(\left(|i\rangle\langle i|\right)_{i=1}^d\right) = 0$, since

$$\delta_\infty\left(\left(|i\rangle\langle i|\right)_{i=1}^d\right) = \sup_{\rho\in S} \tilde{\delta}\left(\left(|i\rangle\langle i|\right)_{i=1}^d, \left(|i\rangle\langle i|\right)_{i=1}^d\right) = 0,$$

(b) is convex, i.e., for any POVM $Q, Q' \in \mathcal{E}_d$

$$\delta_\infty\left(\lambda Q + (1-\lambda)Q'\right) \leq \lambda\delta_\infty\left(Q\right) + (1-\lambda)\delta_\infty\left(Q'\right) \quad \forall\lambda \in [0,1],$$

because

$$\begin{aligned}\delta_\infty\left(\lambda Q + (1-\lambda)Q'\right) &= \sup_{\rho\in S} \tilde{\delta}(p, \lambda q + (1-\lambda)q')\\ &\leq \lambda \sup_{\rho\in S} \tilde{\delta}(p,q) + (1-\lambda)\sup_{\rho\in S} \tilde{\delta}(p,q')\\ &= \lambda\delta_\infty\left(Q\right) + (1-\lambda)\delta_\infty\left(Q'\right),\end{aligned}$$

where we have denoted the corresponding probability distribution as $q_i := \mathrm{tr}\left[\rho Q_i\right]$ and $q'_i := \mathrm{tr}\left[\rho Q'_i\right]$.

(c) is permutation-invariant, i.e., for every permutation $\pi \in S_d$ and any POVM $E \in \mathcal{E}_d$

$$\delta_\infty\left(\left(U_\pi^* E_{\pi(i)} U_\pi\right)_{i=1}^d\right) = \delta_\infty\left(\left(E_i\right)_{i=1}^d\right),$$

where U_π is the permutation matrix that acts as $U_\pi|i\rangle = |\pi(i)\rangle$, since

$$\begin{aligned}
\delta_\infty\left(\left(U_\pi^* E_{\pi(i)} U_\pi\right)_{i=1}^d\right) &= \sup_{\rho\in S} \tilde{\delta}\left(\left(\mathrm{tr}\left[\rho|i\rangle\langle i|\right]\right)_{i=1}^d, \left(\mathrm{tr}\left[\rho U_\pi^* E_{\pi(i)} U_\pi\right]\right)_{i=1}^d\right) \\
&= \sup_{\rho\in S} \tilde{\delta}\left(\left(\mathrm{tr}\left[\rho|i\rangle\langle i|\right]\right)_{i=1}^d, \left(\mathrm{tr}\left[U_\pi \rho U_\pi^* E_{\pi(i)}\right]\right)_{i=1}^d\right) \\
&= \sup_{\rho\in S} \tilde{\delta}\left(\left(\mathrm{tr}\left[U_\pi^* \rho U_\pi|i\rangle\langle i|\right]\right)_{i=1}^d, \left(\mathrm{tr}\left[\rho E_{\pi(i)}\right]\right)_{i=1}^d\right) \\
&= \sup_{\rho\in S} \tilde{\delta}\left(\left(\mathrm{tr}\left[\rho U_\pi|i\rangle\langle i|U_\pi^*\right]\right)_{i=1}^d, \left(\mathrm{tr}\left[\rho E_{\pi(i)}\right]\right)_{i=1}^d\right) \\
&= \sup_{\rho\in S} \tilde{\delta}\left(\left(\mathrm{tr}\left[\rho|\pi(i)\rangle\langle\pi(i)|\right]\right)_{i=1}^d, \left(\mathrm{tr}\left[\rho E_{\pi(i)}\right]\right)_{i=1}^d\right) \\
&= \sup_{\rho\in S} \tilde{\delta}\left(\left(\mathrm{tr}\left[\rho|i\rangle\langle i|\right]\right)_{i=1}^d, \left(\mathrm{tr}\left[\rho E_i\right]\right)_{i=1}^d\right) \\
&= \delta_\infty\left(\left(E_i\right)_{i=1}^d\right),
\end{aligned}$$

(d) and it satisfies for every diagonal unitary $D \in \mathcal{M}_d$ and any POVM $E \in \mathcal{E}_d$

$$\delta_\infty\left(\left(D^* E_i D\right)_{i=1}^d\right) = \delta_\infty\left(\left(E_i\right)_{i=1}^d\right),$$

because

$$\begin{aligned}
\delta_\infty\left(\left(D^* E_i D\right)_{i=1}^d\right) &= \sup_{\rho\in S} \tilde{\delta}\left(\left(\mathrm{tr}\left[\rho|i\rangle\langle i|\right]\right)_{i=1}^d, \left(\mathrm{tr}\left[\rho D^* E_i D\right]\right)_{i=1}^d\right) \\
&= \sup_{\rho\in S} \tilde{\delta}\left(\left(\mathrm{tr}\left[\rho|i\rangle\langle i|\right]\right)_{i=1}^d, \left(\mathrm{tr}\left[D\rho D^* E_i\right]\right)_{i=1}^d\right) \\
&= \sup_{\rho\in S} \tilde{\delta}\left(\left(\mathrm{tr}\left[D\rho D^*|i\rangle\langle i|\right]\right)_{i=1}^d, \left(\mathrm{tr}\left[\rho E_i\right]\right)_{i=1}^d\right) \\
&= \sup_{\rho\in S} \tilde{\delta}\left(\left(\mathrm{tr}\left[\rho D^*|i\rangle\langle i|D\right]\right)_{i=1}^d, \left(\mathrm{tr}\left[\rho E_i\right]\right)_{i=1}^d\right) \\
&= \sup_{\rho\in S} \tilde{\delta}\left(\left(\mathrm{tr}\left[\rho|i\rangle\langle i|\right]\right)_{i=1}^d, \left(\mathrm{tr}\left[\rho E_i\right]\right)_{i=1}^d\right) \\
&= \delta_\infty\left(\left(E_i\right)_{i=1}^d\right).
\end{aligned}$$

Similarly, its average case δ_μ satisfies

(a) $\delta_\mu\left(\left(|i\rangle\langle i|\right)_{i=1}^d\right) = 0$, since

$$\delta_\mu\left(\left(|i\rangle\langle i|\right)_{i=1}^d\right) = \int_{\mathcal{S}_d} \tilde{\delta}\left(\left(|i\rangle\langle i|\right)_{i=1}^d, \left(|i\rangle\langle i|\right)_{i=1}^d\right) \mathrm{d}\mu(\rho) = 0,$$

(b) is convex, i.e., for any POVM $Q, Q' \in \mathcal{E}_d$

$$\delta_\mu\left(\lambda Q + (1-\lambda)Q'\right) \leq \lambda\delta_\mu\left(Q\right) + (1-\lambda)\delta_\mu\left(Q'\right) \quad \forall\lambda \in [0,1],$$

because

$$\begin{aligned}\delta_\mu\left(\lambda Q+(1-\lambda)Q'\right) &= \int_{\mathcal{S}_d}\tilde{\delta}(p,\lambda q+(1-\lambda)q')\,\mathrm{d}\mu(\rho)\\ &\leq \lambda\int_{\mathcal{S}_d}\tilde{\delta}(p,q)\,d\mu(\rho)+(1-\lambda)\int_{\mathcal{S}_d}\tilde{\delta}(p,q')\,\mathrm{d}\mu(\rho)\\ &= \lambda\delta_\mu\left(Q\right)+(1-\lambda)\delta_\mu\left(Q'\right),\end{aligned}$$

where we have denoted the corresponding probability distribution as $q_i := \mathrm{tr}\left[\rho Q_i\right]$ and $q'_i := \mathrm{tr}\left[\rho Q'_i\right]$.

(c) is permutation-invariant, i.e. for every permutation $\pi \in S_d$ and any $E \in \mathcal{E}_d$

$$\delta_\mu\left(\left(U_\pi^* E_{\pi(i)} U_\pi\right)_{i=1}^d\right) = \delta_\mu\left(\left(E_i\right)_{i=1}^d\right)$$

where U_π is the permutation matrix that acts as $U_\pi|i\rangle = |\pi(i)\rangle$, since

$$\begin{aligned}\delta_\mu\left(\left(U_\pi^* E_{\pi(i)} U_\pi\right)_{i=1}^d\right) &= \int_{\mathcal{S}_d}\tilde{\delta}\left(\left(\mathrm{tr}\left[\rho|i\rangle\langle i|\right]\right)_{i=1}^d,\left(\mathrm{tr}\left[\rho U_\pi^* E_{\pi(i)} U_\pi\right]\right)_{i=1}^d\right)\,\mathrm{d}\mu(\rho)\\ &= \int_{\mathcal{S}_d}\tilde{\delta}\left(\left(\mathrm{tr}\left[\rho|i\rangle\langle i|\right]\right)_{i=1}^d,\left(\mathrm{tr}\left[U_\pi\rho U_\pi^* E_{\pi(i)}\right]\right)_{i=1}^d\right)\,\mathrm{d}\mu(\rho)\\ &= \int_{\mathcal{S}_d}\tilde{\delta}\left(\left(\mathrm{tr}\left[U_\pi^*\rho U_\pi|i\rangle\langle i|\right]\right)_{i=1}^d,\left(\mathrm{tr}\left[\rho E_{\pi(i)}\right]\right)_{i=1}^d\right)\,\mathrm{d}\mu(\rho)\\ &= \int_{\mathcal{S}_d}\tilde{\delta}\left(\left(\mathrm{tr}\left[\rho U_\pi|i\rangle\langle i|U_\pi^*\right]\right)_{i=1}^d,\left(\mathrm{tr}\left[\rho E_{\pi(i)}\right]\right)_{i=1}^d\right)\,\mathrm{d}\mu(\rho)\\ &= \int_{\mathcal{S}_d}\tilde{\delta}\left(\left(\mathrm{tr}\left[\rho|\pi(i)\rangle\langle\pi(i)|\right]\right)_{i=1}^d,\left(\mathrm{tr}\left[\rho E_{\pi(i)}\right]\right)_{i=1}^d\right)\,\mathrm{d}\mu(\rho)\\ &= \int_{\mathcal{S}_d}\tilde{\delta}\left(\left(\mathrm{tr}\left[\rho|i\rangle\langle i|\right]\right)_{i=1}^d,\left(\mathrm{tr}\left[\rho E_i\right]\right)_{i=1}^d\right)\,\mathrm{d}\mu(\rho)\\ &= \delta_\mu\left(\left(E_i\right)_{i=1}^d\right),\end{aligned}$$

(d) and it satisfies for every diagonal unitary $D \in \mathcal{M}_d$ and any $E \in \mathcal{E}_d$

$$\delta_\mu\left(\left(D^* E_i D\right)_{i=1}^d\right) = \delta_\mu\left(\left(E_i\right)_{i=1}^d\right),$$

because

$$\begin{aligned}
\delta_\mu\left((D^*E_iD)_{i=1}^d\right) &= \int_{\mathcal{S}_d} \tilde{\delta}\left(\left(\mathrm{tr}\left[\rho|i\rangle\langle i|\right]\right)_{i=1}^d, \left(\mathrm{tr}\left[\rho D^*E_iD\right]\right)_{i=1}^d\right)\,\mathrm{d}\mu(\rho)\\
&= \int_{\mathcal{S}_d} \tilde{\delta}\left(\left(\mathrm{tr}\left[\rho|i\rangle\langle i|\right]\right)_{i=1}^d, \left(\mathrm{tr}\left[D\rho D^*E_i\right]\right)_{i=1}^d\right)\,\mathrm{d}\mu(\rho)\\
&= \int_{\mathcal{S}_d} \tilde{\delta}\left(\left(\mathrm{tr}\left[D\rho D^*|i\rangle\langle i|\right]\right)_{i=1}^d, \left(\mathrm{tr}\left[\rho E_i\right]\right)_{i=1}^d\right)\,\mathrm{d}\mu(\rho)\\
&= \int_{\mathcal{S}_d} \tilde{\delta}\left(\left(\mathrm{tr}\left[\rho D^*|i\rangle\langle i|D\right]\right)_{i=1}^d, \left(\mathrm{tr}\left[\rho E_i\right]\right)_{i=1}^d\right)\,\mathrm{d}\mu(\rho)\\
&= \int_{\mathcal{S}_d} \tilde{\delta}\left(\left(\mathrm{tr}\left[\rho|i\rangle\langle i|\right]\right)_{i=1}^d, \left(\mathrm{tr}\left[\rho E_i\right]\right)_{i=1}^d\right)\,\mathrm{d}\mu(\rho)\\
&= \delta_\mu\left((E_i)_{i=1}^d\right).
\end{aligned}$$

The worst-case as well as the average-case construction therefore satisfy Assumption 2. □

Proof of Corollary 1.

Proof. The eigenvalues of J_{T_i}, $i = 1, 2$, can be obtained from the expectation values of the mutually orthogonal projectors, i.e.,

$$x_1 = \frac{\mathrm{tr}\left[(P_x \otimes \mathbb{1})J_T\right]}{\mathrm{tr}\left[P_x\right]} \quad \text{and} \quad x_2 = \frac{\mathrm{tr}\left[(\mathbb{1} \otimes P_x)J_T\right]}{\mathrm{tr}\left[P_x\right]}, \quad x \in \{a, b, c\}.$$

Since we know that a, b, c are related to α, β, γ via $\alpha = d^2a,\ \beta = b - c,\ \gamma = d(c - a)$, we get

$$\begin{aligned}
\alpha_1 &= d^2\frac{\mathrm{tr}\left[(P_a \otimes \mathbb{1})J_T\right]}{\mathrm{tr}\left[P_a\right]}\\
&= d^2\frac{\mathrm{tr}\left[\left(\mathbb{1}_{d^3} - \sum_{i=1}^d |ii\rangle\langle ii| \otimes \mathbb{1}_d\right)J_T\right]}{d^2 - d}.
\end{aligned}$$

Similarly,

$$\begin{aligned}
\beta_1 &= \frac{\mathrm{tr}\left[(P_b \otimes \mathbb{1})J_T\right]}{\mathrm{tr}\left[P_b\right]} - \frac{\mathrm{tr}\left[(P_c \otimes \mathbb{1})J_T\right]}{\mathrm{tr}\left[P_c\right]}\\
&= \frac{\mathrm{tr}\left[\left(\frac{1}{d}\sum_{i,j=1}^d |ii\rangle\langle jj| \otimes \mathbb{1}_d\right)J_T\right]}{1}\\
&\quad - \frac{\mathrm{tr}\left[\left(\sum_{i=1}^d |ii\rangle\langle ii| \otimes \mathbb{1}_d - \frac{1}{d}\sum_{i,j=1}^d |ii\rangle\langle jj| \otimes \mathbb{1}_d\right)J_T\right]}{d - 1},
\end{aligned}$$

and

$$\begin{aligned} a_2 &= \frac{\mathrm{tr}\left[(\mathbb{1}\otimes P_a)J_T\right]}{\mathrm{tr}\left[P_a\right]} \\ &= \frac{\mathrm{tr}\left[\left(\mathbb{1}_{d^3} - \mathbb{1}_d \otimes \sum_{i=1}^d |ii\rangle\langle ii|\right) J_T\right]}{d^2-d}. \end{aligned}$$

Using the diagrammatic notation introduced earlier, i.e.,

$$\mathbb{1}_{d^3} =: \quad \qquad \mathbb{1}_d \otimes \sum_{i=1}^d |ii\rangle\langle ii| =: \quad$$

$$\sum_{i,j=1}^d |ii\rangle\langle jj| \otimes \mathbb{1}_d =: \quad \qquad \sum_{i=1}^d |ii\rangle\langle ii| \otimes \mathbb{1}_d =: \quad$$

together with the isomorphic representation from Lemma 3, the claim follows immediately. □

References

[1] E. B. Davies and J. T. Lewis, "An operational approach to quantum probability," *Comm. Math. Phys.*, vol. 17, no. 3, pp. 239–260, 1970.

[2] G. Lüders, "Über die Zustandsänderung durch den Meßprozeß," *Ann. Phys.*, vol. 443, no. 5-8, pp. 322–328, 1950.

[3] C. H. Bennet and G. Brassard, "Quantum cryptography: Public key distribution and coin tossing," in *Proceedings of IEEE International Conference on Computers, Systems and Signal Processing*, pp. 175–179, Dec 1984.

[4] A. K. Ekert, "Quantum cryptography based on Bell's theorem," *Phys. Rev. Lett.*, vol. 67, pp. 661–663, Aug 1991.

[5] C. A. Fuchs and A. Peres, "Quantum-state disturbance versus information gain: Uncertainty relations for quantum information," *Phys. Rev. A*, vol. 53, pp. 2038–2045, Apr 1996.

[6] C. A. Fuchs, *Information Gain vs. State Disturbance in Quantum Theory*, ch. 13, pp. 229–259. Wiley-VCH Verlag GmbH & Co. KGaA, Jan 2005.

[7] H. Martens and W. M. de Muynck, "Disturbance, conservation laws and the uncertainty principle," *J. Phys. A*, vol. 25, no. 18, p. 4887, 1992.

[8] M. Ozawa, "Universally valid reformulation of the Heisenberg uncertainty principle on noise and disturbance in measurement," *Phys. Rev. A*, vol. 67, p. 042105, Apr 2003.

[9] M. Ozawa, "Uncertainty relations for noise and disturbance in generalized quantum measurements," *Ann. Phys.*, vol. 311, no. 2, pp. 350–416, 2004.

[10] T. Heinosaari and M. M. Wolf, "Nondisturbing quantum measurements," *J. Math. Phys.*, vol. 51, no. 9, p. 092201, 2010.

[11] Y. Watanabe and M. Ueda, "Quantum estimation theory of error and disturbance in quantum measurement," *ArXiv e-prints*, Jun 2011.

[12] A. C. Ipsen, "Error-disturbance relations for finite dimensional systems," *ArXiv e-prints*, Nov 2013.

[13] P. Busch, P. Lahti, and R. F. Werner, "Proof of Heisenberg's error-disturbance relation," *Phys. Rev. Lett.*, vol. 111, p. 160405, Oct 2013.

[14] P. Busch, P. Lahti, and R. F. Werner, "Colloquium: Quantum root-mean-square error and measurement uncertainty relations," *Rev. Mod. Phys.*, vol. 86, pp. 1261–1281, Oct 2014.
[15] C. Branciard, "How well can one jointly measure two incompatible observables on a given quantum state?," *Proc. Natl. Acad. Sci. USA*, vol. 110, pp. 6742–6747, Apr 2013.
[16] F. Buscemi, M. J. W. Hall, M. Ozawa, and M. M. Wilde, "Noise and disturbance in quantum measurements: An information-theoretic approach," *Phys. Rev. Lett.*, vol. 112, p. 050401, Feb 2014.
[17] P. J. Coles and F. Furrer, "State-dependent approach to entropic measurement-disturbance relations," *Phys. Lett. A*, vol. 379, pp. 105–112, Jan 2015.
[18] R. Schwonnek, D. Reeb, and R. F. Werner, "Measurement uncertainty for finite quantum observables," *Mathematics*, vol. 4, p. 38, Jun 2016.
[19] J. M. Renes, V. B. Scholz, and S. Huber, "Uncertainty relations: An operational approach to the error-disturbance tradeoff," *Quantum*, vol. 1, p. 20, Jul 2017.
[20] K. Banaszek, "Fidelity balance in quantum operations," *Phys. Rev. Lett.*, vol. 86, pp. 1366–1369, Feb 2001.
[21] H. Barnum, "Information-disturbance tradeoff in quantum measurement on the uniform ensemble," in *Proceedings of IEEE International Symposium on Information Theory*, p. 277, 2001.
[22] L. Maccone, "Entropic information-disturbance tradeoff," *Europhys. Lett.*, vol. 77, no. 4, p. 40002, 2007.
[23] D. Kretschmann, D. Schlingemann, and R. F. Werner, "The information-disturbance tradeoff and the continuity of Stinespring's representation," *IEEE Trans. Inf. Theory*, vol. 54, pp. 1708–1717, April 2008.
[24] F. Buscemi, M. Hayashi, and M. Horodecki, "Global information balance in quantum measurements," *Phys. Rev. Lett.*, vol. 100, p. 210504, May 2008.
[25] F. Buscemi and M. Horodecki, "Towards a unified approach to information-disturbance tradeoffs in quantum measurements," *Open Syst. Inf. Dyn.*, vol. 16, no. 01, pp. 29–48, 2009.
[26] A. Bisio, G. Chiribella, G. M. D'Ariano, and P. Perinotti, "Information-disturbance tradeoff in estimating a unitary transformation," *Phys. Rev. A*, vol. 82, p. 062305, Dec 2010.
[27] T. Shitara, Y. Kuramochi, and M. Ueda, "Trade-off relation between information and disturbance in quantum measurement," *Phys. Rev. A*, vol. 93, p. 032134, Mar 2016.
[28] L. Knips, J. Dziewior, A. K. Hashagen, J. Meinecke, H. Weinfurter, and M. M. Wolf, "Measurement-disturbance tradeoff outperforming optimal cloning," *In preparation*, 2018.
[29] F. Hiai, M. Mosonyi, D. Petz, and C. Bény, "Quantum f-divergences and error correction," *Rev. Math. Phys.*, vol. 23, no. 07, pp. 691–747, 2011.
[30] D. R. Farenick, *Algebras of Linear Transformations*. Universitext, Springer-Verlag New York, 1st ed., 2001.
[31] J. Bochnak, M. Coste, and M.-F. Roy, *Real algebraic geometry*, vol. 36 of *Ergebnisse der Mathematik und ihrer Grenzgebiete 3. Folge / A Series of Modern Surveys in Mathematics*. Springer-Verlag Berlin Heidelberg, 1st ed., 1998.
[32] M. Karow, *Geometry of spectral value sets*. PhD thesis, Universität Bremen, Jun 2003.
[33] M. Marshall, *Positive Polynomials and Sums of Squares*, vol. 146 of *Mathematical Surveys and Monographs*. American Mathematical Society, 2008.
[34] M. M. Wolf, T. S. Cubitt, and D. Perez-Garcia, "Are problems in quantum information theory (un)decidable?," *ArXiv e-prints*, Nov 2011.
[35] C. Scheiderer, "Semidefinite representation for convex hulls of real algebraic curves," *ArXiv e-prints*, Sep 2017.

[36] C. Scheiderer, "Spectrahedral shadows," *ArXiv e-prints*, Dec 2017.
[37] J. Watrous, "Semidefinite programs for completely bounded norms," *ArXiv e-prints*, Jan 2009.
[38] J. Watrous, "Simpler semidefinite programs for completely bounded norms," *ArXiv e-prints*, Jul 2012.
[39] I. CVX Research, "CVX: Matlab software for disciplined convex programming, version 2.0." `http://cvxr.com/cvx`, Aug 2012.
[40] M. Grant and S. Boyd, "Graph implementations for nonsmooth convex programs," in *Recent Advances in Learning and Control* (V. Blondel, S. Boyd, and H. Kimura, eds.), Lecture Notes in Control and Information Sciences, pp. 95–110, Springer-Verlag Limited, 2008.
[41] The MathWorks, Inc., *MATLAB and Statistics Toolbox Release R2014b*. Natick, Massachusetts, United States, 2014.

[1] Department of Mathematics, Technical University of Munich

[2] Kavli Institute for Theoretical Physics, University of California, Santa Barbara (Aug - Dec, 2017)

C Contributed further article: Article 3

L. Knips, J. Dziewior, A. K. Hashagen, J. D. A. Meinecke, H. Weinfurter and M. M. Wolf
Measurement-disturbance tradeoff outperforming optimal cloning
Arxiv e-prints: arXiv:1808.07882 [quant-ph], August 2018
Submitted to Physical Review Letters

Summary of article 3: Measurement-Disturbance Tradeoff outperforming Optimal Cloning [4]

In quantum mechanics, every measurement that extracts information, necessarily introduces some disturbance to the system. This idea is fundamental and emphasizes again the difference between classical and quantum mechanics. It, furthermore, finds applications within areas of quantum information theory, such as quantum communication [59, 88–90]. It is thus of special interest, whether the insights gained through a thorough mathematical analysis, are accessible in an experimental setting and really lead to an advantage in experimental applications.

In this article we experimentally realize the information-disturbance tradeoff derived in [2]. In the special case of binary measurements on qubits, which is the case considered here, we rederive results from [2] under specific assumptions, significantly simplifying the proofs. The derivation yields a tradeoff and specifies the optimal quantum instruments achieving it.

Theorem C.1 (Total variation - trace norm tradeoff [4, theorem 1])**.** *Consider a von Neumann target measurement given by an orthonormal basis* $\left\{|i\rangle \in \mathbb{C}^2\right\}_{i=1}^{2}$*, and an instrument with two corresponding outcomes. Then the worst-case total variational distance* δ *and its trace-norm analogue* Δ*, defined as*

$$\delta(E') := \sup_{\rho} \frac{1}{2} \sum_{i=1}^{2} \left| \mathrm{Tr}[E'_i \rho] - \mathrm{Tr}[E_i \rho] \right|, \tag{C.1}$$

and

$$\Delta(T_s) := \frac{1}{2} \sup_{\rho} \left\| T_s(\rho) - \rho \right\|_1, \tag{C.2}$$

quantifying measurement error and disturbance respectively, satisfy

$$\Delta \geq \begin{cases} \frac{1}{2}\left(\sqrt{1-\delta} - \sqrt{\delta}\right)^2 & \text{if } \delta \leq \frac{1}{2}, \\ 0 & \text{if } \delta \geq \frac{1}{2}. \end{cases} \tag{C.3}$$

The inequality is tight and equality is attained within the family of instruments defined by

$$I_j(\rho) := K_j \rho K_j, \tag{C.4}$$

with

$$K_j = \sqrt{1-\gamma}\,|j\rangle\langle j| + \sqrt{\gamma}(\mathbb{1} - |j\rangle\langle j|), \tag{C.5}$$

with $\gamma \in [0, 1/2]$, $j = 1, 2$.

Theorem C.1, therefore, gives the quantum instruments that need to be realized in the laboratory, in order to experimentally measure the information-disturbance tradeoff. In general, it is possible to realize every quantum channel acting on a quantum system by an interaction with an auxiliary system. In this article, we use an optical Mach-Zehnder interferometer to realize the quantum instrument. The quantum state is encoded in the polarization and the auxiliary system is provided by the path degree of freedom. Figure C.1 shows the experimental realizations of different quantum instruments with respect to the measurement error on one side and the state disturbance on the other side. The measurements agree well with theoretical predictions, which are represented by the blue line.

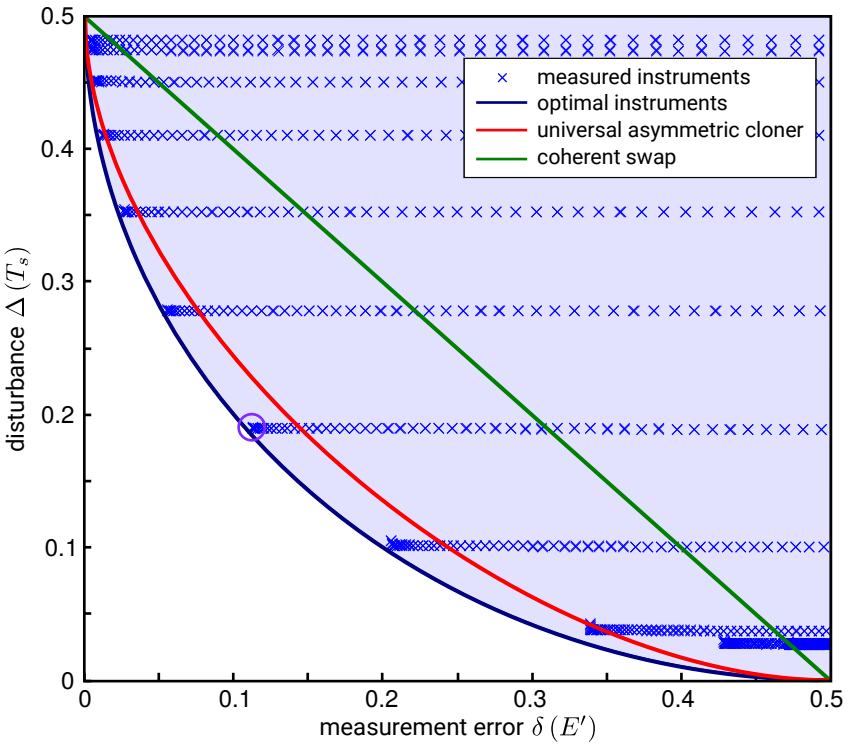


Figure C.1: [4, figure 1] The experimentally realized quantum instruments are represented by blue crosses. The optimal ones clearly outperform the theoretical prediction of the universal asymmetric quantum cloning protocol (red line) as well as the theoretical prediction of the coherent swap protocol (green line) with high significance. They, furthermore, turn out to be close to the theoretical optimal protocol (blue line).

These findings are compared to two other protocols, namely the universal asymmetric quantum cloning protocol and the coherent swap protocol. In the first one, a quantum state is asymmetrically cloned and the target measurement is performed on one of the clones. The asymmetry within the quality then yields the tradeoff regarding measurement error and disturbance. The resulting tradeoff is derived in the following theorem.

Theorem C.2 (Total variation - trace norm tradeoff using optimal universal asymmetric cloning [4, theorem 2])**.** *Consider a von Neumann measurement given by an orthonormal basis in* $\mathbb{C}^2$ *on one of the outputs of the optimal universal* $1 \to 2$ *asymmetric quantum cloning channel. Then the worst-case total variational distance* δ *and its trace-norm analogue* Δ *satisfy*

$$\Delta = \begin{cases} \frac{1}{4}\left(\sqrt{2-3\delta} - \sqrt{\delta}\right)^2 & \text{if } \delta \leq \frac{1}{2}, \\ 0 & \text{if } \delta \geq \frac{1}{2}. \end{cases} \tag{C.6}$$

The second protocol uses the coherent swap. The quantum state is coherently swapped with the maximally mixed state and the target measurement is then, again, performed on one of the systems. The resulting tradeoff is derived.

Theorem C.3 (Total variation - trace norm tradeoff using the coherent swap [4, theorem 3])**.** *Consider a von Neumman measurement given by an orthonormal basis in* $\mathbb{C}^2$ *on one of the outputs of a coherent swap channel. Then the worst-case total variational distance* δ *and its trace-norm analogue* Δ *satisfy*

$$\Delta = \frac{1}{2} - \delta. \tag{C.7}$$

Both theoretically derived tradeoffs are shown in figure C.1. The tradeoff derived using the universal asymmetric quantum cloning protocol is represented by the red line, whereas the tradeoff from the coherent swap protocol is shown using the blue line. These schemes clearly do not perform optimally and we have shown that this advantage is experimentally accessible and not just a mere theoretical improvement.

Statement of individual contribution

I, Anna-Lena Karolyn Hashagen, presented first results of my research project at the Munich Quantum Center poster session [2]. Jan Dziewior was very interested in this project on information-disturbance tradeoffs and said that he has been so for some time, but never really found access to the more mathematical papers on this topic. I, therefore, explained to him what the research question is that we were trying to answer and pointed out some of the methods we were using. We then realized that one of his group members, Lukas Knips, is actually a member of the Ph.D. program *Exploring Quantum Matter* of the Elite Network of Bavaria, of which I am also a member. After a second meeting, we then quickly realized that with their laboratory equipment, they can experimentally implement the theoretical findings of our article [2] in a qubit setting. Following up on this, I had numerous discussions with Lukas Knips, Jan Dziewior and Dr. Jasmin D. A. Meinecke on the experimental realization. I explained to them the mathematical preliminaries to understand the article [2] and suggested to focus on the binary qubit setting.

Moreover, I understood that the diamond norm as a measure of disturbance is not a feasible measure in practice and therefore suggested to focus on the worst-case trace norm distance.

Furthermore, this is the quantum analogue of the worst-case total variational distance, which is used as a figure of merit to assess the measurement error. Prof. Dr. Michael M. Wolf and I were able to prove that in the case considered in this article the allowed auxillary systems do not give an advantage and, for the optimal tradeoff curve, the trace norm equals the diamond norm distance.

Prof. Dr. Michael M. Wolf suggested to compare the tradeoff obtained by the optimal quantum instruments from [2] with other protocols. One natural candidate is the asymmetric quantum cloning protocol. Another protocol involves the coherent swap. In a meeting with all authors of this article present, we decided to pursue this line and compare our results to these two measurement protocols.

I proved and formulated all theorems in this article. That is, I was solely responsible for theorems [4, theorem 1, 2 and 3] as well as lemma [4, lemma 4 and 5].

Lukas Knips, Jan Dziewior, Dr. Jasmin D. A. Meinecke and Prof. Dr. Harald Weinfurter were solely responsible for the experimental implementation in the laboratory. They built the experiment, they ran the experiment and they recorded the data. My doctoral supervisor Prof. Dr. Michael M. Wolf and I were not involved in this process. I was, however, involved in the analysis of the recorded data.

Furthermore, I was responsible for writing the theoretical part of this article and I was extensively involved in polishing the final version.

Lukas Knips is the principal author of this paper. I, Anna-Lena Karolyn Hashagen, was extensively involved in all parts of this article, except in the experimental implementation in the laboratory.

Article

Measurement-Disturbance Tradeoff Outperforming Optimal Cloning

Lukas Knips,[1,2] Jan Dziewior,[1,2] Anna-Lena K. Hashagen,[3] Jasmin D. A. Meinecke,[1,2] Harald Weinfurter,[1,2] and Michael M. Wolf[3]

[1] *Max-Planck-Institut für Quantenoptik, Hans-Kopfermann-Straße 1, 85748 Garching, Germany*
[2] *Department für Physik, Ludwig-Maximilians-Universität, 80797 München, Germany*
[3] *Fakultät für Mathematik, Technische Universität München, Germany*

One of the characteristic features of quantum mechanics is that every measurement that extracts information about a general quantum system necessarily causes an unavoidable disturbance to the state of this system. A plethora of different approaches has been developed to characterize and optimize this tradeoff. Here, we apply the framework of quantum instruments to investigate the optimal tradeoff and to derive a class of procedures that is optimal with respect to most meaningful measures. We focus our analysis on binary measurements on qubits as commonly used in communication and computation protocols and demonstrate theoretically and in an experiment that the optimal universal asymmetric quantum cloner, albeit ideal for cloning, is not an optimal procedure for measurements and can be outperformed with high significance.

Introduction.—The work of Heisenberg, best visualized by the Heisenberg microscope [1], teaches us that every measurement is accompanied by a fundamental disturbance of a quantum system. The question about the precise relation between the information gained about the quantum system and the resulting disturbance has since inspired numerous studies [2–19]. A central problem is to find a tight, quantitative tradeoff relation, e.g., for the maximally achievable information for a given disturbance or, vice versa, for the minimal disturbance for a certain amount of extracted information. Obviously, this is not only relevant for quantum foundations, but also for many applications in quantum communication [20, 21] and quantum computation [22–24]. Initially studied in the context of which-path information and loss of visibility in interferometers [2, 3], quantifying the information-disturbance tradeoff was based on various measures such as the traditional root mean squared distance [4, 5], the distance of probability distributions [6], operation and estimation fidelities [7–9], entropic quantities [8–13], reversibility [13–15], stabilized operator norms [16, 17], state discrimination probability [10], probability distribution fidelity [18], and Fisher information [19]. In spite of all these distinct approaches, no clear candidate for a most fundamental framework for the analysis of the information-disturbance tradeoff in quantum mechanics has yet emerged.

Here we build upon a novel, comprehensive information-disturbance relation introduced recently by two of us [25]. There, optimal measurement devices have been proven to be independent of the chosen quality measures, as long as these fulfill some reasonable assumptions, such as convexity and basis-independence. This approach is unique with respect to the employment of reference observables. On one hand, since information eventually is obtained via measurements of observables, we base the quantification of the measurement error on a reference observable. On the other hand, the measurement induced disturbance is defined without relying on any reference observable in order not to restrict the further usage of the post-measurement state. For a finite-dimensional von Neumann measurement, the optimal tradeoff can be achieved with quantum instruments described by at most two parameters.

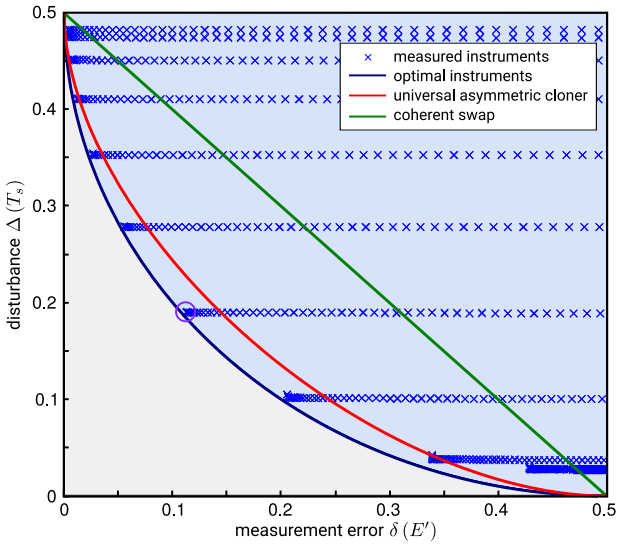


FIG. 1. The optimal quantum instruments in terms of measurement error and disturbance clearly outperform the optimal asymmetric cloner (red curve) and the coherent swap operation (green line). Our measurements (blue crosses) come close to the theoretical curve (blue curve). The violet marked instrument is discussed in Fig. 5 in more detail. The error bars are too small to be visible; for a detailed discussion see [26].

In this letter, we describe how optimal instruments can be derived for typical measures of measurement error, i.e., inverse information, and state disturbance and how they can be implemented in an experiment. Typically, quantum cloning is considered to be a good choice to achieve an optimal measurement disturbance tradeoff. Yet, here we show that the optimal instruments outper-

form all (asymmetric) quantum cloners [26]. We test the tradeoff relation experimentally using a tunable Mach-Zehnder-Interferometer and implement a large range of quantum instruments. We apply these instruments to a two-dimensional quantum system encoded in the photon polarization and investigate the relation between the error of the measurement and the disturbance of the qubit state. As distance measures we consider exemplarily some of the measures recommended in [16], i.e., the worst-case total variational distance and the worst-case trace norm. For other measures see supplemental material (SM) [26]. The experiment clearly shows that the optimal universal asymmetric cloner as well as the coherent swap scheme are suboptimal (Fig. 1).

Measurements as quantum instruments.—To generally quantify both the measurement error and the measurement induced disturbance, we describe the measurement of observables on a quantum system by means of quantum instruments [27, 28] as illustrated in Fig. 2. Formally, a quantum instrument I is defined as a set of completely positive linear maps $I := \{I_j\}_{j=1}^m$ that fulfills the normalization condition $\sum_{j=1}^m I_j^*(\mathbb{1}) = \mathbb{1}$, where I_j^* denotes the dual map to I_j with respect to the Hilbert-Schmidt inner product. This description naturally encompasses the connection between the observable given by a positive operator valued measure (POVM) $E' := \{E'_j\}_{j=1}^m$ and the quantum channel T_s, which describes the measurement induced change of the state.

In general, a quantum channel is a completely positive trace preserving linear map. In the context of quantum instruments, the channel is given by the sum of the linear maps with $T_s := \sum_{j=1}^m I_j$, where each map corresponds to one measurement operator E'_j of the POVM. The normalization condition of the quantum instrument ensures that the corresponding quantum channel is trace-preserving. Expressing the channel in terms of I as above reflects the decohering effect of the measurement on the quantum state of the measured system.

The measurement operators $\{E'_j\}_{j=1}^m$ themselves are fully determined by I via $E'_j := I_j^*(\mathbb{1})$, where the probability distribution for outcomes $\{j\}_{j=1}^m$ on state ρ is given by $\operatorname{tr}(I_j(\rho)) = \operatorname{tr}(I_j(\rho)\mathbb{1}) = \operatorname{tr}\left(\rho I_j^*(\mathbb{1})\right) = \operatorname{tr}\left(\rho E'_j\right)$. From this point of view, the normalization condition of the quantum instrument ensures that the distribution $\{\operatorname{tr}\left(E'_j\rho\right)\}_{j=1}^m$ is normalized. The instrument description based on the normalized set of maps I, which implies the pair (E', T_s), is sufficient to exhaustively describe all possible quantum measurement processes.

Distance measures.—From the notion of quantum instruments it becomes immediately clear that E' and T_s are not independent, i.e. the change of the state has a fundamental dependence on the information gained and vice versa. To enable a thorough quantitative analysis of this measurement-disturbance tradeoff, we use distance measures to assess the quality of the approximate measurement and to quantify the disturbance. We quantify the disturbance Δ caused to the system by the deviation of the channel T_s from the identity channel $T_{\mathrm{id}}(\rho) := \rho$.

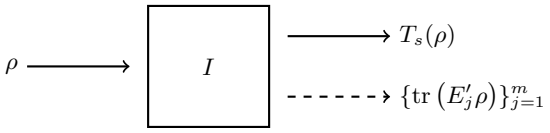


FIG. 2. General description of a measurement using a quantum instrument I. Obtaining information about the quantum state via the POVM E' (dashed line, classical output) induces a change of the quantum state described by the quantum channel T_s (solid line, quantum output).

The measurement error δ quantifies the deviation of the measurement E' from a reference measurement E. This approach utilizes a reference POVM E to quantify the measurement error, but not the disturbance, in contrast to all other approaches found in the literature, where either a reference system is used for both, measurement error and disturbance, or none is used at all.

The measurement error δ can be quantified by defining a worst-case total variational distance based on the l_1-distance between probability distributions. The l_1-distance, also called total variational distance, displays the largest possible difference between the probabilities that two probability distributions assign to the same event and therefore is the relevant distance measure for hypothesis testing [28, 29]. In our case, these two probability distributions stem from the target measurement E and the actual measurement E' for some quantum state. To generalize the measure for the measurement error to take into account all possible quantum states ρ of the system we additionally take the worst case w.r.t. all states, which is natural when considering the maximal difference, i.e., worst-case characteristic of the l_1-distance itself. Thus our worst-case total variational distance is defined as

$$\delta(E') := \sup_\rho \frac{1}{2} \sum_{i=1}^{2} \left|\operatorname{tr}(E'_i\rho) - \operatorname{tr}(E_i\rho)\right|. \tag{1}$$

The quantum analogue of the worst-case total variational distance is the worst-case trace norm distance, which we thus use to quantify the distance between the quantum channel T_s and the identity channel T_{id},

$$\Delta(T_s) := \frac{1}{2} \sup_\rho \left\|T_s(\rho) - \rho\right\|_1. \tag{2}$$

This disturbance measure quantifies how well the quantum channel T_s can be distinguished from the identity channel T_{id} in a statistical experiment, if no auxiliary systems are allowed [30].

Optimal instruments and tradeoff.—As reference measurement, we choose the ideal projective measurement of the qubit with $E = \{|j\rangle\langle j|\}_{j=1}^2$. As proven in [25] for the optimal quantum instruments each element I_j can be expressed by a single Kraus operator, agreeing with the intuition that additional Kraus operators introduce

noise to the system. In the case of a qubit this leads to

$$T_s(\rho) = \sum_{j=1}^{2} K_j \rho K_j^\dagger \quad \text{and} \quad \{E'_j = K_j^\dagger K_j\}_{j=1}^2. \tag{3}$$

The Kraus operators of an optimal instrument can be chosen diagonal in the basis $\{|j\rangle\}_{j=1}^2$ given by the target measurement [25]. Since for a qubit there are only two of them and they must satisfy the normalization condition, in general their form is

$$K_1 = \sqrt{1-b_2^2}\,|1\rangle\langle 1| + e^{i\beta_1} b_1\,|2\rangle\langle 2|\,, \tag{4a}$$

$$K_2 = b_2\,|1\rangle\langle 1| + e^{i\beta_2}\sqrt{1-b_1^2}\,|2\rangle\langle 2|\,, \tag{4b}$$

with $0 \leq b_1^2, b_2^2 \leq 1$ and two arbitrary phases β_1 and β_2.

As proven in [26], for such an instrument, the worst-case total variational distance δ and its trace-norm analogue Δ, Eqs. (1,2), quantifying measurement error and disturbance respectively, satisfy

$$\Delta \geq \begin{cases} \frac{1}{2}\left(\sqrt{1-\delta} - \sqrt{\delta}\right)^2 & \text{if } \delta \leq \frac{1}{2}, \\ 0 & \text{if } \delta \geq \frac{1}{2}. \end{cases} \tag{5}$$

The inequality is tight and cannot be exceeded by any quantum measurement procedure. Equality in Eq. (5) is attained for the family of optimal instruments defined by

$$K_1 = \frac{1}{\sqrt{2}}\left(\sqrt{1-\gamma}\,|1\rangle\langle 1| + \sqrt{1+\gamma}\,|2\rangle\langle 2|\right), \tag{6a}$$

$$K_2 = \frac{1}{\sqrt{2}}\left(\sqrt{1+\gamma}\,|1\rangle\langle 1| + \sqrt{1-\gamma}\,|2\rangle\langle 2|\right), \tag{6b}$$

with $\gamma \in [0,1]$, leading to $\delta(\gamma) = (1-\gamma)/2$.

Other known measurement schemes.—Let us evaluate common quantum measurement procedures in terms of their measurement-disturbance tradeoff. For perfect quantum cloning, there would be no measurement-disturbance tradeoff, as one of the perfect clones could be measured without error with the other clone staying undisturbed. Although perfect cloning is impossible [31], one can derive a protocol that is optimal for approximate quantum cloning. Hence, it is a manifest intuition that the optimal universal asymmetric quantum cloner provides a promising measurement protocol that naturally leads simultaneously to a small disturbance and a small measurement error. It is illustrated in Fig. 3. The quantum channel $T_s(\rho) = \mathrm{tr}_{s'}\left(T_{\mathrm{clo}}(\rho)\right)$, a marginal of the cloning channel T_{clo}, corresponds to the evolution of the system state, obtained when tracing out the second (primed) clone. The corresponding channel of the second clone, $T_{s'}(\rho) = \mathrm{tr}_s\left(T_{\mathrm{clo}}(\rho)\right)$, provides an approximate copy to which the reference POVM E is applied. Asymmetry within the quality of the clones determines the tradeoff between the measurement error and the disturbance.

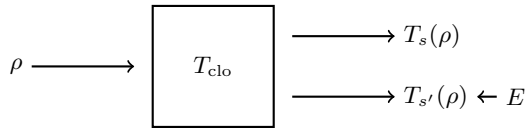


FIG. 3. Universal asymmetric quantum cloning. The initial quantum state ρ is asymmetrically, approximately cloned to the auxiliary system, initially in state $\mathbb{1}/2$. The target measurement is performed on one of the clones, while the other is compared to the initial quantum state ρ.

The optimal universal asymmetric quantum cloning channel T_{clo} for any initial quantum state ρ reads [32]

$$T_{\mathrm{clo}}(\rho) = (a_2\mathbb{1} + a_1\mathbb{F})\left(\rho \otimes \frac{\mathbb{1}}{2}\right)(a_2\mathbb{1} + a_1\mathbb{F})\,, \tag{7}$$

with $a_1^2 + a_2^2 + a_1 a_2 = 1$, $a_1, a_2 \in \mathbb{R}$, and the flip (or swap) operator $\mathbb{F} := \sum_{i,j=1}^2 |ji\rangle\langle ij|$. The parameter a_1 determines the amplitude of a swap operation between both qubits.

With our measures, the measurement-disturbance tradeoff for the asymmetric quantum cloning channel satisfies

$$\Delta = \begin{cases} \frac{1}{4}\left(\sqrt{2-3\delta} - \sqrt{\delta}\right)^2 & \text{if } \delta \leq \frac{1}{2}, \\ 0 & \text{if } \delta \geq \frac{1}{2} \end{cases} \tag{8}$$

with $\delta(a_2) = a_2^2/2$ [26].

As the cloning operation cannot be realized by a unitary two-qubit transformation, any real implementation of the protocol is embedded in a larger system. Let us thus consider an obvious analogue to the cloning operation, which can be realized by a unitary two-qubit operation. For the swapping channel T_{cs}, the system interacts with the auxiliary system via a Heisenberg Hamiltonian as

$$\begin{aligned} T_{\mathrm{cs}}(\rho) &= e^{it\mathbb{F}}\,(\rho \otimes \tilde{\rho})\,e^{-it\mathbb{F}} \\ &= (a_2\mathbb{1} + ia_1\mathbb{F})\,(\rho \otimes \tilde{\rho})\,(a_2\mathbb{1} - ia_1\mathbb{F})\,, \end{aligned} \tag{9}$$

with $t \in [0, \pi/2]$ or using a parametrization analogous to the cloning scheme with $a_1^2 + a_2^2 = 1$, $a_1, a_2 \in \mathbb{R}$. The extreme cases are no swap ($t = 0$, $a_2 = 1$) and full swap ($t = \pi/2$, $a_1 = 1$).

The δ-Δ-tradeoff for the target measurement $E = \{|j\rangle\langle j|\}_{j=1}^2$ performed on one of the outputs satisfies

$$\Delta = \frac{1}{2} - \delta, \tag{10}$$

with $\delta(t) = (1 - a_1^2)/2$, for the coherent swap [26], evidently also inferior to our optimal instruments, Eq. (6), with the tradeoff given in Eq. (5).

Experimental implementation.—For our experimental evaluation of the measurement-disturbance tradeoff we want to realize a broad range of quantum instruments including the optimal ones. For that purpose we consider

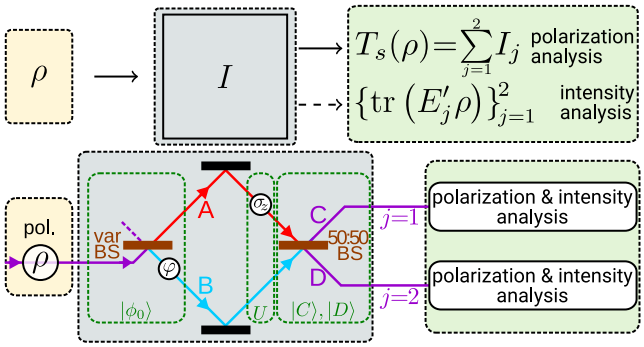


FIG. 4. Conceptual experimental setup. The state ρ is encoded in the polarization degree of freedom of a photon, which is sent to a variable beam splitter (var BS). The spatial superposition state inside of the interferometer is denoted by $|\phi_0\rangle$ and can be tuned in terms of relative intensities and phase. For the interaction U between the path and the polarization degrees of freedom we apply a σ_z operation to the polarization in one path. Projections onto the output ports $|C\rangle$ and $|D\rangle$ of a balanced 50:50 beam splitter conclude the realization of the Kraus operators as given in Eqs. (12). Polarization and intensity measurements are performed at the output ports of the interferometer. Please note that the actual experiment, while equivalent to the shown setup, is structured differently such that the polarization state ρ is created inside of the interferometer. The actual experiment is described in more detail in [26].

the polarization degree of freedom of photons to encode ρ, with $|1\rangle \leftrightarrow |H\rangle$ and $|2\rangle \leftrightarrow |V\rangle$, where $|H\rangle$ ($|V\rangle$) denotes horizontally (vertically) polarized light. The Kraus operators describing the chosen set of instruments are thus given by

$$K_{1,2} = \frac{1}{\sqrt{2}}\left[\sqrt{1\pm\gamma}\,|H\rangle\langle H| + e^{i\beta}\sqrt{1\mp\gamma}\,|V\rangle\langle V|\right] \quad (11)$$

with an arbitrary phase β. The optimal cases Eqs. (6) are achieved for $\beta = 0$.

To experimentally realize a quantum instrument and to enable analysis of the two outputs T_s and E', it is necessary to employ an additional auxiliary quantum system, which is not yet explicitly present in the instrument description of Fig. 2. For the measurement of photon polarization a natural candidate is the path degree of freedom of the photons. Since in our case a two dimensional auxiliary system is sufficient, we employ a Mach-Zehnder interferometer, which provides the two path states $|A\rangle$ and $|B\rangle$, see Fig. 4. The properties of the instrument are then determined by the initial state of this auxiliary system, $|\phi_0\rangle = \cos\alpha\,|A\rangle + e^{i\varphi}\sin\alpha\,|B\rangle$, the measurement performed on it, i.e., the detection in the output path states $|C\rangle$ and $|D\rangle$, as well as by an intermediate interaction between path and polarization. The interaction is given by a unitary evolution U, which exchanges information between the systems. We use $U = i\sigma_z\otimes|A\rangle\langle A| + \mathbb{1}\otimes|B\rangle\langle B|$, which introduces a polarization dependent phase shift in arm $|A\rangle$.

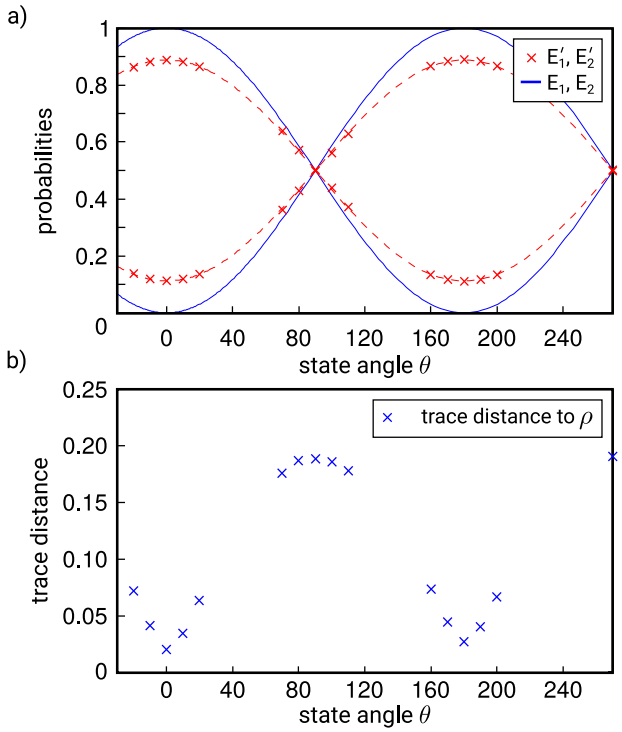


FIG. 5. Evaluating measurement error δ and disturbance Δ. a) The measurement error corresponds to the maximal distance between the outcomes of the actual measurements E'_1 and E'_2 (red crosses) to the outcomes of the ideal measurements E_1 and E_2 (blue line). b) The disturbance is obtained by taking the supremum of the trace distance between the prepared polarization states and the tomographically reconstructed states of T_s. Please note that the suprema in a) and b) are achieved for different states. Statistical error bars are negligibly small. For a detailed discussion, see [26].

For an initial path state $|\phi_0\rangle$ the Kraus operators, which act on the polarization, can then be obtained as

$$K_1 = \mathrm{tr}_{\mathrm{path}}\left[(\mathbb{1}\otimes|C\rangle\langle C|)\,U\,(\mathbb{1}\otimes|\phi_0\rangle\langle\phi_0|)\right], \quad (12a)$$
$$K_2 = \mathrm{tr}_{\mathrm{path}}\left[(\mathbb{1}\otimes|D\rangle\langle D|)\,U\,(\mathbb{1}\otimes|\phi_0\rangle\langle\phi_0|)\right]. \quad (12b)$$

Relating these expressions with Eq. (11), the parameters γ and β are given by the experimental parameters α and φ by $\gamma = \sin(2\alpha)\sin\varphi$ and $\beta = \arctan[\tan(2\alpha)\cos\varphi]$. The outcome of the measurement E' is then obtained by determining the total intensity in the output C (E'_1) and D (E'_2), respectively, the action of the quantum channel T_s by state tomography of the polarization degree of freedom.

Measurements and results.—According to Eqs. (1) and (2), the measures δ and Δ use the supremum over different input states ρ. We thus prepare for each quantum instrument different linearly polarized states ρ, which are analyzed after the interaction. The prepared polarization state $\rho = |\psi\rangle\langle\psi|$ in both arms is given by $|\psi\rangle = \cos\frac{\theta}{2}|H\rangle + \sin\frac{\theta}{2}|V\rangle$, where $|H\rangle$ and $|V\rangle$ as the eigenstates of the Pauli matrix σ_z with eigenvalues $+1$ and -1, respectively, denote horizontal and vertical po-

larization. We use 16 different values for θ, including those where extremal behavior for the disturbance or the measurement error is expected. The set of pure, linearly polarized states is sufficient as the suprema in Eqs. (1) and (2) are attained in our experimental implementation, see SM [26].

An intuitive strategy consists of setting a specific instrument and then varying the polarization state ρ, which however requires to keep the instrument parameters (α and φ) stable. It turns out to be experimentally more favorable to prepare different polarization states ρ and then vary the phase φ for fixed α and ρ. One thus associates measurements which correspond to the same state $|\phi_0\rangle$ of the auxiliary system to the same instrument.

The evaluation of the measurement error and the disturbance for one instrument of Fig. 1 is shown in Fig. 5 a) and b), respectively. The supremum over a great circle of the Bloch sphere, described by $|\psi\rangle$, has been used for the analysis. The measurement error is given by the maximal deviation of the measurement (red crosses) to the best fitting target measurement (blue solid line), see Eq. (1). While some states as eigenstates of the transformation (theoretically) do not show any disturbance, for the disturbance, the largest trace distance has to be taken into account, see Eq. (2).

The obtained values for measurement error and state disturbance are shown in Fig. 1 for the set of experimentally prepared quantum instruments. Each data point here identifies one quantum instrument, for which the supremum of the prepared quantum states in terms of measurement error and disturbance is determined. The horizontal structure is explained when considering that for a fixed α, various measurements with different φ have been taken, see Eq. (11). We could show that there exist quantum instruments, also experimentally accessible, which significantly outperform the optimal universal asymmetric cloner (red curve) and the coherent swap operation (green line) in terms of the considered distances.

Conclusion.—We applied the novel approach derived in [25] to the setting of binary qubit measurements achieving an optimal measurement-disturbance tradeoff. In this setting a reference measurement is used to quantitatively obtain the measurement error. The disturbance, on the other hand, does not depend on any reference measurement, but solely on comparing the state before and after the measurement. Our protocol is tailored for applications based on a specific measurement without restricting subsequent use of the post-measurement state.

Furthermore, we have demonstrated that the strategies of optimal universal asymmetric quantum cloning and coherent swap do not perform optimally when considering the tradeoff relation between measurement error and disturbance. Those protocols are optimal for their respective purposes such as approximate quantum cloning, but cannot compete with the optimal quantum instruments in the measurement scenario as in general they result in worse measurement-disturbance tradeoff relations. We have shown that the advantage of optimal instruments over other schemes is experimentally accessible and not only a mere theoretical improvement. In future applications our findings allow to identify these procedures which retrieve information at the physically lowest cost in terms of state disturbance.

Acknowledgments.—We thank Jonas Goeser for stimulating discussions. This research was supported in part by the National Science Foundation under Grant No. NSF PHY11-25915 and by the German excellence initiative Nanosystems Initiative Munich. LK and AKH are supported by the PhD program *Exploring Quantum Matter* of the Elite Network of Bavaria. JD acknowledges support by the International Max-Planck Research Program for Quantum Science and Technology (IMPRS-QST). JDMA is supported by an LMU research fellowship.

[1] Werner Heisenberg, *The Physical Principles of the Quantum Theory* (University of Chicago Press, 1930).

[2] Gregg Jaeger, Abner Shimony, and Lev Vaidman, "Two interferometric complementarities," Phys. Rev. A **51**, 54–67 (1995).

[3] Berthold-Georg Englert, "Fringe Visibility and Which-Way Information: An Inequality," Phys. Rev. Lett. **77**, 2154–2157 (1996).

[4] Masanao Ozawa, "Universally valid reformulation of the Heisenberg uncertainty principle on noise and disturbance in measurement," Phys. Rev. A **67**, 042105 (2003).

[5] Cyril Branciard, "Error-tradeoff and error-disturbance relations for incompatible quantum measurements," Proceedings of the National Academy of Sciences **110**, 6742–6747 (2013).

[6] Paul Busch, Pekka Lahti, and Reinhard F. Werner, "Proof of Heisenberg's Error-Disturbance Relation," Phys. Rev. Lett. **111**, 160405 (2013).

[7] Konrad Banaszek, "Fidelity Balance in Quantum Operations," Phys. Rev. Lett. **86**, 1366–1369 (2001).

[8] Christopher A. Fuchs and Asher Peres, "Quantum-state disturbance versus information gain: Uncertainty relations for quantum information," Phys. Rev. A **53**, 2038–2045 (1996).

[9] Lorenzo Maccone, "Information-disturbance tradeoff in quantum measurements," Phys. Rev. A **73**, 042307 (2006).

[10] Francesco Buscemi, Masahito Hayashi, and Michał Horodecki, "Global Information Balance in Quantum Measurements," Phys. Rev. Lett. **100**, 210504 (2008).

[11] Francesco Buscemi, Michael J. W. Hall, Masanao Ozawa, and Mark M. Wilde, "Noise and Disturbance in Quantum Measurements: An Information-Theoretic Approach," Phys. Rev. Lett. **112**, 050401 (2014).

[12] Jun Zhang, Yang Zhang, and Chang-Shui Yu, "The Measurement-Disturbance Relation and the Disturbance Trade-off Relation in Terms of Relative Entropy," *International Journal of Theoretical Physics*, **55** (2016).

[13] Giacomo M. D'Ariano, "On the Heisenberg principle, namely on the information-disturbance trade-off in a quantum measurement," Fortschritte der Physik **51**, 318–330 (2003).
[14] Andrew N. Jordan and Alexander N. Korotkov, "Uncollapsing the wavefunction by undoing quantum measurements," Contemporary Physics **51**, 125–147 (2010).
[15] Yong Wook Cheong and Seung-Woo Lee, "Balance Between Information Gain and Reversibility in Weak Measurement," Phys. Rev. Lett. **109**, 150402 (2012).
[16] Alexei Gilchrist, Nathan K. Langford, and Michael A. Nielsen, "Distance measures to compare real and ideal quantum processes," Phys. Rev. A **71**, 062310 (2005).
[17] Dennis Kretschmann, Dirk Schlingemann, and Reinhard F. Werner, "The Information-Disturbance Tradeoff and the Continuity of Stinespring's Representation," IEEE Transactions on Information Theory **54**, 1708–1717 (2008).
[18] Longfei Fan, Wenchao Ge, Hyunchul Nha, and M. S. Zubairy, "Trade-off between information gain and fidelity under weak measurements," Phys. Rev. A **92**, 022114 (2015).
[19] Tomohiro Shitara, Yui Kuramochi, and Masahito Ueda, "Trade-off relation between information and disturbance in quantum measurement," Phys. Rev. A **93**, 032134 (2016).
[20] Nicolas Gisin, Grégoire Ribordy, Wolfgang Tittel, and Hugo Zbinden, "Quantum cryptography," Reviews of Modern Physics **74**, 145–195 (2002).
[21] Jian-Wei Pan, Zeng-Bing Chen, Chao-Yang Lu, Harald Weinfurter, Anton Zeilinger, and Marek Żukowski, "Multiphoton entanglement and interferometry," Reviews of Modern Physics **84**, 777–838 (2012).
[22] Artur Ekert and Richard Jozsa, "Quantum computation and Shor's factoring algorithm," Reviews of Modern Physics **68**, 733–753 (1996).
[23] Vlatko Vedral and Martin B. Plenio, "Basics of quantum computation," Progress in Quantum Electronics **22**, 1–39 (1998).
[24] Andrew Steane, "Quantum computing," Reports on Progress in Physics **61**, 117–173 (1998).
[25] Anna-Lena K. Hashagen and Michael M. Wolf, "Universality and Optimality in the Information-Disturbance Tradeoff," ArXiv e-prints (2018), arXiv:1802.09893 [quant-ph].
[26] *Supplemental Material*.
[27] E. Brian Davies and John T. Lewis, "An operational approach to quantum probability," Communications in Mathematical Physics **17**, 239–260 (1970).
[28] John Watrous, *The Theory of Quantum Information* (Cambridge University Press, 2018).
[29] Jerzy Neyman and Egon S. Pearson, "On the Problem of the Most Efficient Tests of Statistical Hypotheses," Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences **231**, 289–337 (1933).
[30] Allowing auxiliary systems, the relevant disturbance measure is the diamond norm, $\Delta_\diamond(T_s) := \frac{1}{2}\sup_\xi \|((T_s - T_{\mathrm{id},d}) \otimes T_{\mathrm{id},d})(\xi)\|_1$, where the state ξ includes auxiliary systems. Here, for the optimal tradeoff curve, the trace norm turns out to be equal to the diamond norm distance [25].
[31] William K. Wootters and Wojciech H. Zurek, "A single quantum cannot be cloned," Nature **299**, 802–803 (1982).
[32] Anna-Lena K. Hashagen, "Universal Asymmetric Quantum Cloning Revisited," Quant. Inf. Comp. **17**, 0747–0778 (2017).

SUPPLEMENTAL MATERIAL

SM 1: OPTIMAL TRADEOFF RELATION

Theorem 1 (Total variation - trace norm tradeoff)**.** *Consider a von Neumann target measurement given by an orthonormal basis* $\{|i\rangle \in \mathbb{C}^2\}_{i=1}^2$*, and an instrument with two corresponding outcomes. Then the worst-case total variational distance* δ *and its trace-norm analogue* Δ*, defined as in Eqs. (1,2), quantifying measurement error and disturbance respectively, satisfy*

$$\Delta \geq \begin{cases} \frac{1}{2}\left(\sqrt{1-\delta}-\sqrt{\delta}\right)^2 & \text{if } \delta \leq \frac{1}{2}, \\ 0 & \text{if } \delta \geq \frac{1}{2}. \end{cases} \tag{S1}$$

The inequality is tight and equality is attained within the family of instruments defined by

$$I_j(\rho) := K_j \rho K_j, \qquad j = 1, 2, \tag{S2}$$

with

$$K_{1,2} = \frac{1}{\sqrt{2}}\left(\sqrt{1 \pm \gamma}\,|1\rangle\langle 1| + \sqrt{1 \mp \gamma}\,|2\rangle\langle 2|\right) \tag{S3}$$

with $\gamma \in [0, 1]$.

Proof. In order to derive the information-disturbance tradeoff, we need to solve the following optimization problem:
For $\gamma \in [0, 1]$

$$\begin{aligned} \text{minimize} \quad & \Delta\left(T_s = \sum_{j=1}^2 I_j\right) \\ \text{subject to} \quad & \delta\left(E' = \left\{I_j^*(\mathbb{1})\right\}_{j=1}^2\right) \leq \gamma, \\ & I_j \text{ is c.p. and} \\ & \sum_{j=1}^2 I_j^*(\mathbb{1}) = \mathbb{1}, \end{aligned} \tag{S4}$$

where the last two constraints ensure that I is an instrument. As discussed before, we assume that every element of the instrument can be expressed using a single Kraus operator. This agrees well with intuition, because more Kraus operators introduce more noise to the system. Furthermore, we assume that these Kraus operators can be chosen diagonal in the basis of the target measurement, $E = \{|j\rangle\langle j|\}_{i=1}^2$, to reflect the symmetry of the optimization problem. These assumptions simplify the optimization problem significantly. The Kraus operators given in Eq. (4) then yield the following POVM elements of the approximate measurement

$$E'_j = (1 - b_{\bar{j}}^2)\,|j\rangle\langle j| + b_j^2(\mathbb{1} - |j\rangle\langle j|), \tag{S5}$$

for $j = 1, 2$, where $\bar{j} = 2$ if $j = 1$ and $\bar{j} = 1$ if $j = 2$ with $0 \leq b_1^2, b_2^2 \leq 1$. The measurement error is thus given as

$$\begin{aligned} \delta(E') &= \sup_\rho \frac{1}{2}\sum_{j=1}^2 \left|\mathrm{tr}\left(E'_j\rho\right) - \langle j|\,\rho\,|j\rangle\right| \\ &= \sup_\rho \frac{1}{2}\sum_{j=1}^2 \left|\mathrm{tr}\left(\left(b_j^2\mathbb{1} - (b_j^2 + b_{\bar{j}}^2)\,|j\rangle\langle j|\right)\rho\right)\right| \\ &= \sup_{\|\psi\|=1} \frac{1}{2}\sum_{j=1}^2 \left|\langle\psi|\,b_j^2\mathbb{1} - (b_j^2 + b_{\bar{j}}^2)\,|j\rangle\langle j|\,|\psi\rangle\right| \\ &= \frac{1}{2}(b_1^2 + b_2^2), \end{aligned}$$

where the convexity of the l_1-norm was used. The disturbance follows from direct calculations,

$$\begin{aligned} \Delta(T_1) &= \frac{1}{2}\sup_\rho \|T_1(\rho) - \rho\|_1 \\ &= \frac{1}{2}\sup_\rho \left\|\sum_{j=1}^2 K_j\rho K_j^\dagger - \rho\right\|_1 \\ &= \frac{1}{2}\left|1 - e^{i\beta_1}b_1\sqrt{1-b_2^2} - e^{i\beta_2}b_2\sqrt{1-b_1^2}\right|. \end{aligned}$$

Without loss of generality, we may assume that $b_1, b_2 \geq 0$ in the optimization problem, such that an optimum is attained for $\beta_1 = \beta_2 = 0$. The optimization problem given in Eq. (S4) therefore simplifies:
For $\gamma \in [0, 1]$

$$\begin{aligned} \text{minimize} \quad & \frac{1}{2}\left(1 - b_1\sqrt{1-b_2^2} - b_2\sqrt{1-b_1^2}\right) \\ \text{subject to} \quad & \frac{1}{2}(b_1^2 + b_2^2) \leq \frac{1}{2}(1-\gamma), \\ & 0 \leq b_1, b_2 \leq 1. \end{aligned} \tag{S6}$$

The global minimum is achieved at

$$b_1 = b_2 = \begin{cases} \sqrt{\frac{1}{2}} & \gamma \in [-1, 0] \\ \sqrt{\frac{1}{2}}\sqrt{1-\gamma} & \gamma \in [0, 1] \end{cases}$$

and as stated in Eq. (S1). □

SM 2: TRADEOFF RELATION FOR OPTIMAL UNIVERSAL ASYMMETRIC CLONING

Theorem 2 (Total variation - trace norm tradeoff using optimal universal asymmetric cloning)**.** *Consider a von Neumann measurement given by an orthonormal basis in* $\mathbb{C}^2$ *on one of the outputs of the optimal universal* $1 \to 2$ *asymmetric quantum cloning channel. Then the worst-case total variational distance* δ *and its trace-norm analogue* Δ *satisfy*

$$\Delta = \begin{cases} \frac{1}{4}\left(\sqrt{2-3\delta}-\sqrt{\delta}\right)^2 & \text{if } \delta \leq \frac{1}{2}, \\ 0 & \text{if } \delta \geq \frac{1}{2}. \end{cases} \tag{S7}$$

Proof. The marginals of the optimal cloning channel are given by

$$T_{\mathrm{clo},i}(\rho) = a_i^2 \frac{\mathbb{1}}{2}\mathrm{tr}\,(\rho) + (1-a_i^2)\rho, \quad i = 1,2, \tag{S8}$$

with $T_{\mathrm{clo},1} = T_s$ and $T_{\mathrm{clo},2} = T_{s'}$. The marginal quantum channel T_s describes the evolution of the quantum state and its distance to the identity channel T_{id} then quantifies the disturbance. Similarly, the marginal $T_{s'}$, whose output is measured by the target measurement E, describes the measurement itself through $E'_j = T^*_{s'}(E_j)$. This is illustrated in Fig. 3. This yields for the disturbance

$$\begin{aligned}\Delta(T_s) :=& \frac{1}{2}\sup_\rho \|T_s(\rho) - \rho\|_1 \\ =& \frac{1}{2}\sup_\rho \left\| a_1^2\frac{\mathbb{1}}{2} - a_1^2\rho \right\|_1 \\ =& \frac{a_1^2}{2}.\end{aligned}$$

The measurement error turns out to be

$$\begin{aligned}\delta(E') :=& \sup_\rho \frac{1}{2}\sum_{j=1}^{2} \left|\mathrm{tr}\left(E'_j\rho\right) - \langle j|\,\rho\,|j\rangle\right| \\ =& \sup_\rho \frac{1}{2}\sum_{j=1}^{2} \left|\mathrm{tr}\left(T^*_{s'}(|j\rangle\langle j|)\rho\right) - \langle j|\,\rho\,|j\rangle\right| \\ =& \sup_\rho \frac{1}{2}\sum_{j=1}^{2} \left|\mathrm{tr}\left(|j\rangle\langle j|\,T_{s'}(\rho)\right) - \langle j|\,\rho\,|j\rangle\right| \\ =& \sup_\rho \frac{1}{2}\sum_{j=1}^{2} \left|\langle j|\,a_2^2\frac{\mathbb{1}}{2} - a_2^2\rho\,|j\rangle\right| \\ =& \frac{a_2^2}{2}.\end{aligned}$$

Substituting this into the trace-preserving condition of the optimal universal asymmetric quantum cloning channel, we obtain the theorem 2. □

SM 3: TRADEOFF RELATION FOR COHERENT SWAP

Theorem 3 (Total variation - trace norm tradeoff using the coherent swap)**.** *Consider a von Neumman measurement given by an orthonormal basis in $\mathbb{C}^2$ on one of the outputs of a coherent swap channel. Then the worst-case total variational distance δ and its trace-norm analogue Δ satisfy*

$$\Delta = \frac{1}{2} - \delta. \tag{S9}$$

Proof. Using the substitution $a_1 = a$ and $a_2 = \sqrt{1-a^2}$ with $a \in [0,1]$ yields the two marginals of the coherent swap quantum channel,

$$T_s(\rho) = a^2\tilde{\rho} + (1-a^2)\rho \tag{S10}$$

and

$$T_{s'}(\rho) = (1-a^2)\tilde{\rho} + a^2\rho. \tag{S11}$$

The disturbance is therefore

$$\begin{aligned}\Delta(T_s) :=& \frac{1}{2}\sup_\rho \|T_s(\rho) - \rho\|_1 \\ =& \frac{1}{2}a^2 \sup_\rho \|\tilde{\rho} - \rho\|_1 .\end{aligned}$$

The optimal choice for $\tilde{\rho}$ should clearly satisfy the points $(\Delta(T_s) = 0, \delta(E') = 1/2)$ and $(\Delta(T_s) = 1/2, \delta(E') = 0)$, where again $E' = T^*_{s'}(E)$. For any such choice of $\tilde{\rho}$ the disturbance thus satisfies $\Delta(T_s) \geq a^2/2$. The measurement error turns out to be

$$\begin{aligned}\delta(E') :=& \sup_\rho \frac{1}{2}\sum_{j=1}^{2} \left|\mathrm{tr}\left(E'_j\rho\right) - \langle j|\,\rho\,|j\rangle\right| \\ =& \sup_\rho \frac{1}{2}\sum_{j=1}^{2} \left|\mathrm{tr}\left(T^*_{s'}(|j\rangle\langle j|)\rho\right) - \langle j|\,\rho\,|j\rangle\right| \\ =& \sup_\rho \frac{1}{2}\sum_{j=1}^{2} \left|\mathrm{tr}\left(|j\rangle\langle j|\,T_{s'}(\rho)\right) - \langle j|\,\rho\,|j\rangle\right| \\ =& \left(1-a^2\right)\sup_\rho \frac{1}{2}\sum_{j=1}^{2} \left|\langle j|\,\tilde{\rho}\,|j\rangle - \langle j|\,\rho\,|j\rangle\right| .\end{aligned}$$

Thus, an optimal choice for $\tilde{\rho}$ that minimizes the disturbance and the measurement error is $\tilde{\rho} = \mathbb{1}/2$. A pure state with the same diagonal entries yields the same measurement error; it would, however, increase the disturbance caused to the system.

The disturbance is then

$$\Delta(T_s) = \frac{a^2}{2},$$

and the measurement error is

$$\delta(E') = \frac{1}{2}\left(1-a^2\right).$$

This gives the linear tradeoff curve given in theorem 3. □

SM 4: PROPERTIES OF DISTANCE MEASURES

The distance measures used throughout this manuscript to quantify the measurement error and the disturbance, denoted by δ and Δ, satisfy Assumption 1 and Assumption 2 of [25] respectively.

Lemma 4. *δ as defined in Eq. (1) satisfies the following properties:*

(a) $\delta(\{|i\rangle\langle i|\}_{i=1}^2) = 0$,

(b) δ is convex,

(c) δ is permutation invariant, i.e., for every permutation π and any measurement M

$$\delta\left(\{U_\pi^\dagger M_{\pi(i)} U_\pi\}_{i=1}^2\right) = \delta\left(\{M_i\}_{i=1}^2\right),$$

where U_π is the permutation matrix that acts as $U_\pi |i\rangle = |\pi(i)\rangle$, and

(d) δ is invariant under diagonal unitaries, i.e., that for every diagonal unitary D and any measurement M

$$\delta\left(\{D^\dagger M_i D\}_{i=1}^2\right) = \delta\left(\{M_i\}_{i=1}^2\right).$$

Proof. Let $\delta(M) := \sup_\rho \frac{1}{2}\sum_{i=1}^2 |\mathrm{tr}\left(M_i\rho\right) - \langle i|\,\rho\,|i\rangle|$. Then

(a) $\delta(\{|i\rangle\langle i|\}_{i=1}^2) = 0$, since

$$\delta(\{|i\rangle\langle i|\}_{i=1}^2) = \sup_\rho \frac{1}{2}\sum_{i=1}^2 |\langle i|\,\rho\,|i\rangle - \langle i|\,\rho\,|i\rangle| = 0,$$

(b) δ is convex, since for any measurements M, M' and for all $\lambda \in [0,1]$,

$$\begin{aligned}
&\delta\left(\lambda M + (1-\lambda)M'\right)\\
=&\sup_\rho \frac{1}{2}\sum_{i=1}^2 |\mathrm{tr}\left(\left(\lambda M_i + (1-\lambda)M_i'\right)\rho\right) - \langle i|\,\rho\,|i\rangle|\\
\leq&\lambda \sup_\rho \frac{1}{2}\sum_{i=1}^2 |\mathrm{tr}\left(M_i\rho\right) - \langle i|\,\rho\,|i\rangle|\\
&\quad + (1-\lambda)\sup_\rho \frac{1}{2}\sum_{i=1}^2 |(\mathrm{tr}\left(M_i'\rho\right) - \langle i|\,\rho\,|i\rangle)|\\
=&\lambda\delta(M) + (1-\lambda)\delta(M'),
\end{aligned}$$

(c) δ is permutation invariant, since for every permutation π and any measurement M

$$\begin{aligned}
&\delta\left(\{U_\pi^\dagger M_{\pi(i)} U_\pi\}_{i=1}^2\right)\\
=&\sup_\rho \frac{1}{2}\sum_{i=1}^2 \left|\mathrm{tr}\left(U_\pi^\dagger M_{\pi(i)} U_\pi\rho\right) - \langle i|\,\rho\,|i\rangle\right|\\
=&\sup_\rho \frac{1}{2}\sum_{i=1}^2 \left|\mathrm{tr}\left(M_{\pi(i)}\rho\right) - \langle \pi(i)|\,\rho\,|\pi(i)\rangle\right|\\
=&\sup_\rho \frac{1}{2}\sum_{i=1}^2 |\mathrm{tr}\left(M_i\rho\right) - \langle i|\,\rho\,|i\rangle|\\
=&\delta\left(\{M_i\}_{i=1}^2\right),
\end{aligned}$$

where U_π is the permutation matrix that acts as $U_\pi |i\rangle = |\pi(i)\rangle$, and

(d) δ is invariant under diagonal unitaries, since for every diagonal unitary D and any measurement M

$$\begin{aligned}
&\delta\left(\{D^\dagger M_i D\}_{i=1}^2\right)\\
=&\sup_\rho \frac{1}{2}\sum_{i=1}^2 \left|\mathrm{tr}\left(D^\dagger M_i D\rho\right) - \langle i|\,\rho\,|i\rangle\right|\\
=&\sup_\rho \frac{1}{2}\sum_{i=1}^2 \left|\mathrm{tr}\left(M_i\rho\right) - \langle i|\,D^\dagger\rho D\,|i\rangle\right|\\
=&\sup_\rho \frac{1}{2}\sum_{i=1}^2 |\mathrm{tr}\left(M_i\rho\right) - \langle i|\,\rho\,|i\rangle|\\
=&\delta\left(\{M_i\}_{i=1}^2\right).
\end{aligned}$$

□

Lemma 5. *Δ as defined in Eq. (2) satisfies the following properties:*

(a) $\Delta(T_{\mathrm{id}}) = 0$,

(b) Δ is convex,

(c) Δ is basis-independent, i.e., for every unitary U and every quantum channel Φ

$$\Delta\left(U\Phi\left(U^\dagger\cdot U\right)U^\dagger\right) = \Delta\left(\Phi\right).$$

Proof. Let $\Delta(\Phi) := \frac{1}{2}\sup_\rho \|\Phi(\rho) - \rho\|_1$. Then

(a) $\Delta(T_{\mathrm{id}}) = 0$, since $\Delta(T_{\mathrm{id}}) = \frac{1}{2}\sup_\rho \|\rho - \rho\|_1 = 0$,

(b) Δ is convex, since for any quantum channels Φ, Φ' and for all $\lambda \in [0,1]$,

$$\begin{aligned}
&\Delta\left(\lambda\Phi + (1-\lambda)\Phi'\right)\\
=&\frac{1}{2}\sup_\rho \|\left(\lambda\Phi + (1-\lambda)\Phi'\right)(\rho) - \rho\|_1\\
=&\frac{1}{2}\sup_\rho \|\lambda\left(\Phi(\rho) - \rho\right) + (1-\lambda)\left(\Phi'(\rho) - \rho\right)\|_1\\
\leq&\lambda\frac{1}{2}\sup_\rho \|\Phi(\rho) - \rho\|_1 + (1-\lambda)\frac{1}{2}\sup_\rho \|\Phi'(\rho) - \rho\|_1\\
=&\lambda\Delta(\Phi) + (1-\lambda)\Delta(\Phi'),
\end{aligned}$$

where we have used properties of a norm and properties of a supremum of a convex functional over a convex set,

(c) Δ is basis-independent, i.e., for every unitary U and every quantum channel Φ

$$\begin{aligned}
&\Delta\left(U\Phi\left(U^\dagger\rho U\right)U^\dagger\right)\\
=&\frac{1}{2}\sup_\rho \left\|U\Phi\left(U^\dagger\rho U\right)U^\dagger - \rho\right\|_1\\
=&\frac{1}{2}\sup_\rho \left\|U\Phi\left(\rho\right)U^\dagger - U\rho U^\dagger\right\|_1\\
=&\frac{1}{2}\sup_\rho \|\Phi\left(\rho\right) - \rho\|_1\\
=&\Delta\left(\Phi\right),
\end{aligned}$$

where we have used the fact that the trace norm is unitarily invariant.

□

SM 5: DIFFERENT MEASURES

The optimal instruments as explained in the main text and derived in Sec. result in optimal measurement-disturbance relations for all distance measures which satisfy the assumptions of [25]. For more details on the distance measure used in the main text see Sec. .

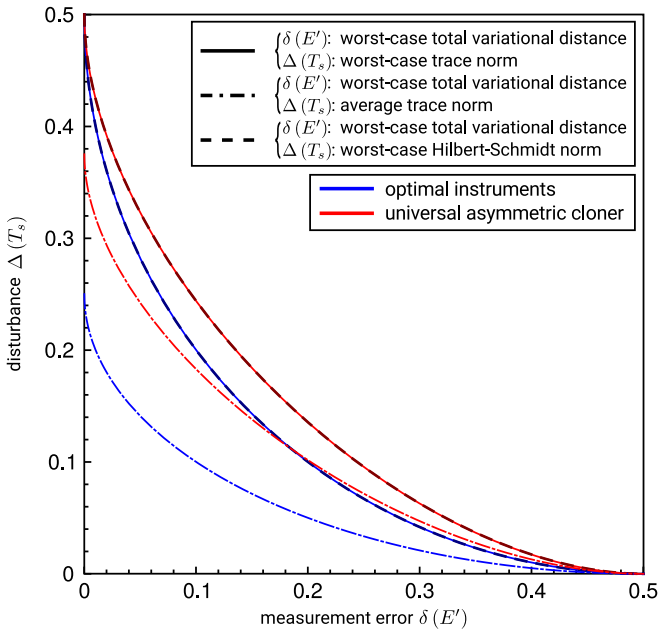


FIG. S1. Comparison of optimal quantum instruments (blue) with the optimal universal asymmetric quantum cloner (red) for different distance measures based on simulations. The tradeoff relation of the main text based on the measures of Eqs. (1) and (2) is shown (solid lines) and equivalent to a properly scaled version of the worst-case Hilbert-Schmidt norm (overlayed dashed lines) and to the worst-case infidelity (not shown). For averaging over all quantum states instead of taking the supremum of the trace norm for the disturbance, one obtains the dashdotted lines.

We here show the tradeoff relations for different choices of disturbance measures, while the measurement error is always quantified as in Eq. (1). For various meaningful measures, we observe that the optimal instruments outperform the cloner, see Fig. S1.

SM 6: EXPERIMENTAL SETUP

Due to experimental and practical limitations, the actual experimental setup has been slightly different than described in the main text. However, the actual implementation is fully equivalent to the description there. In order to be able to fully tune the attenuation in one of the interferometer arms, we use a half waveplate (HWP) sandwiched between two polarizers. Therefore, the polarization state ρ cannot be set before. Hence, we decided to first create the spatial superposition state $|\phi_0\rangle$ using waveplates and polarizers and subsequently set ρ in both interferometer arms separately. With this approach, we still achieve at this stage a separable state $\rho \otimes |\phi_0\rangle\langle\phi_0|$ within the interferometer before the interaction. As we set the polarization state directly in front of the second beam splitter of the interferometer, the reflection of beam A on the beam splitter already provides the interaction between system and auxiliary system. This reflection induces the unitary transformation U as described in the main text, enabling us to obtain the Kraus operators given in Eq. (11).

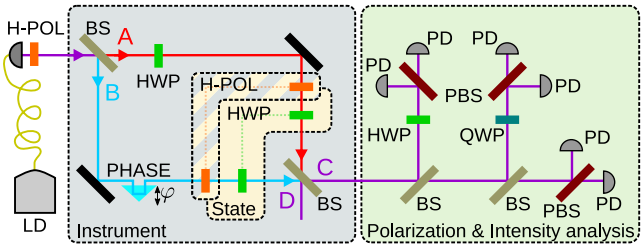


FIG. S2. Actual experimental setup. Light from a diode laser (LD) propagates through a single mode fiber and is sent through a fixed polarizer (H-POL). A beam splitter (BS) creates a spatial superposition. The attenuation of one arm can be adjusted using a half waveplate (HWP) in arm A and another H-POL. The relative phase φ can be varied using a piezo controlled prism. H-POLs together with variable HWPs ensure equal polarization in both arms as indicated by the dotted lines. As the H-POLs are used to vary the attenuation as well as to set the polarization state, they are part of both the instrument and the state preparation. The reflection from arm A on the second BS introduces a coupling between polarization and path. Polarization and intensity measurements are performed in output port C using waveplates (HWP and QWP), polarizing beam splitters (PBS) and photodiodes (PD). Output port D is not monitored, as for phase φ_0 it is redundant to the output of port C at phase $\varphi_0 + \pi$.

Since for a perfect beam splitter the output ports are interchanged for $\varphi_0 \leftrightarrow \varphi_0 + \pi$, we use only output port C to obtain data for both projections, considering the phases $\varphi_1 = \varphi_0$ and $\varphi_2 = \varphi_0 + \pi$. This way, both projections are carried out with exactly the same equipment, reducing possible experimental errors.

SM 7: CHOICE OF POLARIZATION STATES

According to the parametrization $|\psi\rangle = \cos\frac{\theta}{2}|H\rangle + \sin\frac{\theta}{2}|V\rangle$, the experimentally prepared values for θ were $\{-20°, -10°, 0°, 10°, 20°, 70°, 80°, 90°, 100°, 110°, 160°, 170°, 180°, 190°, 200°, 270°\}$. For $\theta = 0°$ and $\theta = 180°$, the prepared state corresponds to horizontal polarization $|H\rangle$ and vertical polarization $|V\rangle$, respec-

tively. Thus, the reflection in beam A only introduces a phase, as for example the state for $\theta = 0°$ is transformed according to

$$|H\rangle \otimes \left(\cos\alpha\,|A\rangle + \sin\alpha e^{i\varphi}\,|B\rangle\right) \rightarrow |H\rangle \otimes \left(i\cos\alpha\,|A\rangle + \sin\alpha e^{i\varphi}\,|B\rangle\right), \tag{S12}$$

which does not change the state of the polarization. The disturbance therefore (ideally) vanishes. In contrast, for $\theta = 90°$, we expect

$$\left(|H\rangle + |V\rangle\right) \otimes \left(\cos\alpha\,|A\rangle + \sin\alpha e^{i\varphi}\,|B\rangle\right) \rightarrow i\left(|H\rangle - |V\rangle\right) \otimes \cos\alpha\,|A\rangle + \left(|H\rangle + |V\rangle\right) \otimes \sin\alpha e^{i\varphi}\,|B\rangle, \tag{S13}$$

where normalization is omitted. For a given instrument characterized by $\{\alpha, \varphi\}$, this polarization state is expected to give the largest disturbance Δ.

For the Kraus operators given in Eq. (11), we find for $E'_j = K_j^\dagger K_j$ for $j = 1, 2$,

$$E'_{1,2} = \frac{1}{2}\begin{pmatrix} 1 \pm \sin 2\alpha\cos\varphi & 0 \\ 0 & 1 \mp \sin 2\alpha\cos\varphi \end{pmatrix}. \tag{S14}$$

Therefore, the distance of the outcome probabilities, used to obtain δ, becomes

$$\frac{1}{2}\sum_i \left|\mathrm{tr}\left(E'_i\,|\psi\rangle\langle\psi|\right) - |\langle i|\psi\rangle|^2\right| = \left|\cos\theta\left(1 - \cos\varphi\sin 2\alpha\right)\right|, \tag{S15}$$

which vanishes for $\theta = 90°$ (and $\theta = 270°$) and can be maximal for $\theta = 0°$ (and $\theta = 180°$).

SM 8: ERROR ANALYSIS OF EXPERIMENTAL DATA

The statistical error of the data shown in Fig. 1 is estimated by comparing the results obtained in redundant measurements. The standard deviation of the measurement error is estimated to be around $8.3 \cdot 10^{-5}$, whereas the 1σ-error bar for the estimated disturbance is approximately $7.0 \cdot 10^{-5}$. Those values are thus too small to be visible in Fig. 1.

Additionally to statistical errors, two different sources of systematic errors have been identified. First, the state preparation as well as the interaction are not perfectly implemented. The imperfect preparation of the initial polarization state and of the state analysis are the main reasons that the identity channel with no disturbance at all (but high measurement error) cannot be implemented perfectly, leading to a residual disturbance, which appears as an increase of the minimal disturbance Δ of the data in the plot. In any case, this type of error only reduces the quality of the prepared quantum instruments and does not lead to faulty conclusions.

However, as a second type of systematic error one has to ensure that the prepared polarization states are describing a great circle on the Bloch sphere and contain the states with extremal results sufficiently well. This error can be approximated by considering the data as shown in Fig. 5. By applying a parabolic model for the data points around the extrema of the probability graphs and the maxima of the trace distance graphs, the deviation of the extrema from the measured points can be estimated. This effect might cause a quantum instrument to look better than it actually is, i.e., less disturbing together with smaller measurement error. Yet, for the dataset shown in Fig. 5 b), the parabolic fit results in a maximum at $\theta \approx 89.95°$ with a trace distance larger by only 0.02% compared to the trace distance at $\theta = 90°$. The probabilities in Fig. 5 a) around $\theta = 0°$ and $\theta = 180°$ can nicely be described by parabolae, where the extrema coincide with our measured points. Thus, the systematic effect of underestimating the measurement error or the disturbance due to badly chosen measurement states is negligibly small.

In conclusion, the different sources of errors overall reduce the quality of the implemented quantum instruments and do not lead to an underestimation of disturbance and measurement error, respectively. We can thus show the implementation of instruments better than the optimal quantum cloner with high significance.

D Contributed core article: Article 4

A. K. Hashagen, S. T. Flammia, D. Gross and J. J. Wallman
Real randomized benchmarking
Quantum, 2:85, August 2018

Summary of article 4: Real Randomized Benchmarking [3]

Randomized benchmarking is a technique that yields quantitative estimates of the average error of the noise inherit to a physical quantum channel. In this article, we study real randomized benchmarking, where the quantum gates under consideration are taken from the real Clifford group. The real Clifford group is the normalizer of the real Pauli group, which only contains the Pauli gates with real entries. It is a subgroup of the complex Clifford group and it is not a unitary 2-design. However, in this article we show that it forms an orthogonal 2-design,

Theorem D.1 ([3, theorem 4])**.** *The real Clifford group $\mathcal{C}(n)$ is an orthogonal 2-design.*

We furthermore clarify how to sample from the real Clifford group, since this is a crucial part of the real randomized benchmarking protocol.

A very useful symmetry argument may then be employed and the twirled quantum channel, which arises due to the averaging procedure in the real randomized benchmarking protocol, is an affine combination of the Werner-Holevo channel, the ideal channel and the completely depolarizing channel. Real randomized benchmarking can thus estimate two parameters instead of one allowing to obtain more fine-grained information about the noise in the system.

Theorem D.2 ([3, theorem 7])**.** *The protocol RealRB(m, E, ρ) has an expected fidelity of the form*

$$\bar{F}(m, E, \rho) = A + b^m B + c^m C, \tag{D.1}$$

where A,B and C are functions only of (E, ρ) and b and c depend on the average noise channel.

One of these parameters is related to the average rebit fidelity, where the average is taken with respect to the orthogonal group, i.e., with respect to rebits. This figure of merit for any two quantum channels $\mathcal{C}$ and $\tilde{\mathcal{C}}$ is defined to be

$$\bar{F}^{\mathbb{R}}\left(\mathcal{C}, \tilde{\mathcal{C}}\right) = \int_{\mathbf{O}(d)} \mathrm{Tr}\left[\mathcal{C}\left(O\left|0\right\rangle\!\left\langle 0\right|O^*\right)\tilde{\mathcal{C}}\left(O\left|0\right\rangle\!\left\langle 0\right|O^*\right)\right] \mathrm{d}O, \tag{D.2}$$

and gives one of the parameters through

$$\bar{F}^{\mathbb{R}}\left(\mathcal{E}, \mathrm{id}\right) = \frac{b(d-1)+1}{d}. \tag{D.3}$$

The usual average fidelity

$$\bar{F}\left(\mathcal{C},\tilde{\mathcal{C}}\right)=\int_{\mathbf{U}(d)}\mathrm{Tr}\left[\mathcal{C}\left(U\left|0\right\rangle\!\left\langle 0\right|U^{*}\right)\tilde{\mathcal{C}}\left(U\left|0\right\rangle\!\left\langle 0\right|U^{*}\right)\right]\mathrm{d}U, \tag{D.4}$$

then allows to calibrate the second parameter through

$$\bar{F}\left(\mathcal{E},\mathrm{id}\right)=\frac{b(d^{2}+d-2)+cd(d-1)+2(d+1)}{2d(d+1)}. \tag{D.5}$$

This makes the real randomized benchmarking analysis especially interesting when considering quantum computations on rebits. It has been shown that universal quantum computing is possible using only rebits [47]. The real Clifford group then plays the role of the complex Clifford group when studying stabilizer circuits [48]. Together with the second parameter, it is then also possible to infer the usual average fidelity. The experimental data may thus be calibrated to the real randomized benchmarking model by varying the quantum gate sequence length and fitting the two parameters.

Statement of individual contribution

This project was initialized, when I, Anna-Lena Karolyn Hashagen, visited Prof. Dr. Stephen D. Bartlett and his quantum physics research group in the School of Physics at the University of Sydney, Australia, from December 2016 until April 2017. Prof. Dr. Steven T. Flammia and Dr. Joel J. Wallmann are permanent members of this quantum physics research group in Sydney. Prof. Dr. David Gross from the Institute for Theoretical Physics of the University of Cologne, Germany, was a visiting professor at that time.

Prof. Dr. Steven T. Flammia provided the idea that the randomized benchmarking protocol could be applied to other groups, not only the complex Clifford group, which was the focus of research at that time. Prof. Dr. David Gross had the idea to look at subgroups of the complex Clifford group. I then discovered that the real Clifford group is a good candidate, because it yields applicable results within quantum computing on rebits. Furthermore, I was able to prove that the real Clifford group is an orthogonal 2-design without any further input from the other authors. Therefore, the theorem [3, theorem 4], the main ingredient of real randomized benchmarking, was proven solely by me. I give two different proofs in the article, one using the structures of the commutants and another one using the frame potential of the real Clifford group. I, Anna-Lena Karolyn Hashagen, was solely responsible for describing the symmetry methods applicable to the real Clifford group.

Furthermore, I was solely responsible for deriving the real randomized benchmarking protocol (realRB protocol) and the proof of the theorem [3, theorem 7]. The protocol yields two parameters to which experimental data can be calibrated. I had the idea to relate one of those parameters to a figure of merit called the average real fidelity, which again connected our results to quantum computations on rebits.

It is important to add a section about how to sample from the real Clifford group. These subtleties, even though they are extremely important and it is highly necessary to address

them, are mostly ignored by the randomized benchmarking community. During several video conferencing sessions, I addressed this issue with the help of Prof. Dr. David Gross, who was able to point me in the right direction and suggested literature to look at. We were then able to address the problem of how to efficiently sample from the real Clifford group together. The section [3, section V] is therefore a result of joint work by Prof. Dr. David Gross and myself.

Furthermore, I was fully responsible for writing this article and finishing up the final version.

I, Anna-Lena Karolyn Hashagen, am the principal author of this article and was extensively involved in all parts of it.

Journal permission and article

Subject **Re: Permission to include article (DOI: https://doi.org/10.22331/q-2018-08-22-85) in my dissertation**

From Lidia del Rio <ldelrio@quantum-journal.org>
To Anna-Lena Hashagen <hashagen@ma.tum.de>, <info@quantum-journal.org>
Date 2018-08-28 13:47

Dear Anna-Lena,

Of course! With Quantum, the copyright of the articles stays with the authors. See more information about the licence here:

https://creativecommons.org/licenses/by/4.0/

Best of luck for your dissertation!

Cheers,

Lídia

for the executive board of Quantum

On 28/08/2018 13:10, Anna-Lena Hashagen wrote:

> Dear Sir or Madam,
>
> I am currently preparing a cumulative dissertation. I am the author of the article
>
> Hashagen, A. K. and Flammia, S. T. and Gross, D. and Wallman, J. J.
> Real randomized benchmarking
> Quantum 2, 85 (2018)
>
> and I would like to ask for permission to include this article in my dissertation.
>
> Kind regards,
> Anna-Lena Hashagen

Terms and conditions

By submitting a manuscript to Quantum, you agree with Quantum's terms and conditions. In particular, you certify that:

- you have the permission of all co-authors and other right holders to pursue publication of the work in Quantum,
- you are not infringing on anyone's copyright with the material contained in your work,
- you will be fully liable for any charges resulting from copyright infringement, and
- you will not submit this work to any other publishing venue unless it is terminally rejected by Quantum.

In addition, authors, referees and members of all boards of Quantum commit to follow the Code of Conduct laid out here.

The above summary is just for informative purposes. The binding terms and conditions follow.

Table of contents

1. Preamble and definitions

Access to Quantum is subject to the following terms and conditions, which constitute a contract between Quantum and any user. By engaging in any form of interaction with Quantum or its members, the user accepts these terms and conditions in full.

We denote by "Quantum" the Association for the Promotion of Open Access Publishing in Quantum Science, legally known as

Verein zur Förderung des Open Access Publizierens in den Quantenwissenschaften

Müllnergasse 26, 1090 Wien, Austria

(ZVR-Zahl 941922539),

the journal Quantum, the website quantum-journal.org, and the online system provided to organize submission

and peer review of works, as well as Quantum's social media accounts.

We denote by "user" any person accessing, using, exchanging, downloading, or submitting any type of content or information with, from and to Quantum and all services it provides, interacting with Quantum in any way, or holding or exercising any function within Quantum.

A "work" submitted to Quantum refers to the actual manuscript as publically available on arxiv.org ("the arXiv"), as well as all supporting material, such as, but not limited to, appendices, supplementary information, a popular summary, datasets, computer code, images, plots, videos or other recordings that are transmitted or made available to Quantum in order to assess the manuscript's suitability for publication. This refers to both the manuscript and material present in the initial submission, as well as all further versions and additions of material in later resubmissions.

The "submitter" is anyone who carries out the submission of a work to Quantum for publication, either through the provided online system or via email, and who is identified by their account in the online system or name and address used in the email. In case of multiple submitters, the definition applies in full to every single one of them.

2. Code of conduct

Quantum fosters scientific integrity and ethical conduct. Consistent with the bylaws of Quantum (such as the constitution of the Verein zur Förderung des Open Access Publizierens in den Quantenwissenschaften), its code of conduct upholds those values, detailing the ethical guidelines and expectations for participation in Quantum. Authors, referees, editors, all board members and all other users of Quantum are expected to act at all times in accordance with the principles and standards described in this code.

Unacceptable behaviour

In particular, the code of conduct specifies behaviour that Quantum deems unacceptable, both in interactions with Quantum and in professional life in general. This includes but is not limited to:

1. Plagiarism and fabrication of data and results, including misrepresentation of contributions and authorship, selective reporting, failure to promptly correct errors, or theft of data and/or other research materials, as well as misrepresentation and overstatement of results and the omission of crucial conditions and assumptions.
2. Publication (or submission for publication) of works submitted to or published in Quantum to other publishing venues, unless the work is terminally rejected by Quantum. "Other publishing venues" include other journals and conference proceedings, but exclude public pre-print servers such as the arXiv and personal or institutional websites of the authors.
3. Subversion of peer review, including failure to declare conflict of interest, failure to recuse under conflict of interest, misuse of information during review, unnecessarily delaying the peer-review process, violation of the anonymity of referees, premature solicitation of press coverage, corruption and/or bribery.
4. Impersonation of other persons or entities, as well as unrightfully claiming the ownership of scientific titles, professional positions, or affiliations.

5. Using Quantum or any other system for the dissemination of scientific works to promote hate or discriminatory speech, or to infringe on the rights of others.
6. Sharing of confidential information, such as the identity of reviewers, referee reports, and other internal correspondence to persons not involved in the peer-review process.
7. Discrimination of any kind, such as on the basis of religion, disability, age, national origin, race, ethnicity, sexual orientation, gender identity, or gender expression. Discrimination includes the use of derogatory comments or slurs.
8. Harassment, including bullying and intimidation, false accusations, threats and assault, as well as sexual harassment in public or in private.
9. The violation of public trust, including making false or misleading statements, to media, and misrepresentation to grant and/or funding agencies.

Reporting and investigation

Quantum may learn of violations through reports from witnesses or aggrieved individuals and parties, including anonymous reports filed through the dedicated online form. For the safety of all reporters, once a report has been made, Quantum editors and board members are bound to maintain the confidentiality of the report except as explicitly requested by the reporting parties. Upon receiving a report, Quantum will progress as follows:

1. The Executive Board of Quantum will name an investigator or form a small investigation body of no more than three people, each of whom must be free of conflict of interest.
2. The investigator or investigating body may solicit additional information from the reporter, with the goal of reaching a tentative conclusion over the course of two weeks.
3. The tentative conclusion of the investigating body will be delivered to the Steering Board, along with a suggested resolution action as described in the section below.
4. If the tentative conclusion and suggested resolution action are agreed upon by the Steering Board, Quantum will inform the reporter of their decision and seek agreement before proceeding.
5. The party suspected of a violation of the code of conduct will be informed of the allegations and planned resolution action and given 20 working days to respond.
6. Depending on the findings, on communication from the involved parties, and consensus of the Steering Board, the resolution action may be implemented or further investigations carried out with the aim of resolving the situation.

During the reporting and investigating process, all individuals must exercise all due diligence to prevent divulging any report details beyond those strictly necessary to enact and uphold the code of conduct. In particular, if upon receiving an initial report, it is deemed the alleged infraction would not result in a penalty more severe than a formal warning, the Executive Board may decide to directly handle the report without the aid of an investigating body, provided that no conflict of interest is introduced.

Enforcement and penalties

If a Quantum user is found through the preceding process to have committed any violation of the code of

conduct, Quantum may enforce the code of conduct in a number of different ways. An appropriate resolution is decided by the Steering Board, taking into account all factors, and having as a goal to improve the situation. Possible actions to enforce the code of conduct include but are not limited to:

1. A formal (written) warning made to the infringing party.
2. Requiring the infringing party to make a formal (written) apology.
3. Reporting the infringing party to their home institutions, employers, and/or professional societies.
4. Reporting the infringing party to the relevant authorities, in case of suspicion of criminal offences.
5. Retraction of compromised manuscripts (based on scientific reasons).
6. Refusal to consider future manuscripts from the infringing party.
7. Expulsion from the Steering, Executive or Editorial Board.

3. Submission and publication of works

Works submitted to Quantum undergo the peer-review process following the Editorial Policies of Quantum. This process ends with either the acceptance of the work for publication, or the terminal rejection of the work.

Responsibilities of the submitter

By submitting a work to Quantum, the submitter warrants all of the following points and assumes full responsibility and liability for any costs and damages resulting directly or indirectly from any of them being untrue:

1. The submitted work is an original creation of the authors listed on the manuscript, and all listed authors have made substantial contributions to the creation of the work.
2. The submitter has the permission of all authors and all other copyright and intellectual property rights holders to pursue the publication of the work in Quantum, and to grant Quantum all the rights specified in these terms and conditions.
3. The manuscript is publicly accessible on the arXiv in the section quant-ph, or at least crosslisted to quant-ph.
4. The work has not been previously published in any other journal or publishing venue, except in conference proceedings and on public pre-print servers such as the arXiv or the authors' personal or institutional websites. Works previously published in conference proceedings must substantially differ from or expand upon the conference version (for example contain previously omitted proofs) and indicate the previous publication on the first page of the manuscript.
5. The submitter has obtained permissions to grant Quantum the rights specified in these terms and conditions for all material contained in the work and has included appropriate credits and prominently marked or indicated any rights held by third parties.
6. The submitter has clearly informed Quantum at the time of submission of any parts of the work which, due to copyright or other constraints, cannot be published by Quantum under the Creative Commons Attribution 4.0 International (CC BY 4.0) licence.

7. In case of acceptance, the final published version of the work will be uploaded on the arXiv under one of the available Creative Commons licences.
8. In case of acceptance, the final published version of the work on the arXiv complies with the Crossref DOI guidelines. In particular, all references cited by the submitted work that have a DOI assigned to them contain DOI links.
9. In case of acceptance and publication in Quantum, the work will not be submitted to other publishing venues, such as journals or conference proceedings.

Rights of the submitter

The submitter is granted the following rights:

1. At any point prior to acceptance the submitter, as well as any author of the work, can withdraw a work from Quantum. A notification of withdrawal has to be submitted to the handling editor either through the online submission system or by email. Upon receiving a notification of withdrawal prior to acceptance, Quantum terminally rejects the work, thereby ending the peer-review process.
2. The submitter, as well as any author of the work, can also withdraw a work from Quantum after acceptance and publication by notifying Quantum through email. This however does not trigger a terminal rejection of the work and in particular does not invalidate the rights granted to Quantum during submission. Quantum will instead put a notification on the publication page that the work was withdrawn.

Internal correspondence

Unless explicitly agreed otherwise by all parts involved, all correspondence between editors, referees and authors during the peer review process should be treated as confidential, and may only be shared with the present and future Editorial Board members who have not declared a conflict of interest with the work.

Rights of Quantum

By submitting a work to Quantum, the submitter explicitly grants Quantum the following additional rights:

1. The right to terminally reject the work, in particular on the basis of the editors' judgement and/or referee reports.
2. The right to permanently store and share the work, referee reports, and intermediate correspondence with the referees and all current and future members of the Editorial Board who have not declared a conflict of Interest.
3. The right to share the identity of the referees with all members of the current and future Editorial Boards who have not declared a conflict of Interest.
4. The non-exclusive right to share, publish, host, distribute, print, advertise, classify, and otherwise use the manuscript, other parts of the work and all metadata associated with it under the Creative Commons Attribution 4.0 International (CC BY 4.0) licence, unless the work is terminally rejected by Quantum, expect for parts of the work that are covered by incompatible licences.
5. The right to deposit the metadata associated with the work in the Crossref system and to assign a DOI to the

work.

6. The right to publish anonymized statistics on submissions and the peer-review process.

Copyright of works published by Quantum

All manuscripts and other parts of works that were previously submitted to Quantum and then directly published by Quantum, as well as the associated meta-data, including for example a work's title, abstract, author list, figures, datasets, or popular summary, are published under the Creative Commons Attribution 4.0 International (CC BY 4.0) licence.

For material associated with a manuscript, such as that linked to from the manuscript, or a work's page on Quantum's website, especially if hosted on other platforms, other licences can apply.

Each owner of copyright on parts of a work submitted or published by Quantum retains their copyright as far as possible under the conditions stated above.

4. Data protection and privacy policy

The purpose of this data protection policy is to inform users about the type of personal data that is collected and processed by Quantum and the hosting company of this website and the extent to which this is happening.
Quantum is taking data protection very seriously and is treating your personal data according to the legal requirements.
In particular Quantum complies with the European General Data Protection Regulation (GDPR).
Please keep in mind any data transmission over the internet can be affected by security flaws.
A complete protection from unauthorized access by third parties can never be fully achieved.
Should you have any questions, please contact us through one of the channels described in the impressum.

Cookies

This website uses cookies. A cookie is a small file that is saved on the device with which you are accessing this website. Should you not want to be served a cookie when using this website, most common browsers can be configured to disallow the usage of cookies. This may affect the usability of this site.

Data for accessing this website

Quantum and the company hosting this website collect and process the following data when this website is accessed:

1. visited page
2. time of access
3. number of transmitted bytes
4. link that lead to the page being accessed
5. browser used

6. operating system used
7. ip address from which it was accessed
8. the geographic region from which it was accessed
9. text entered in forms and search boxes
10. files uploaded through forms

The data is collected in server log files and for the purpose of analyzing how this website is used by means of the WordPress plug-in WP statistics.
The data is saved exclusively on-site and not shared with any third party service such as Google Analytics.
The collected data is used exclusively for statistical analysis and improvement of the website as well as to ensure its safe and lawful operation.
The ip address is saved only in pseudo-anonymized form, either hashed or with the last block truncated.
You have the option to opt-out of the data collection with the WP statistics plug-in, by clearing all cookies for this website in your browser, refreshing the page, and then choosing “opt out” in the banner at the bottom of the page.

Sharing buttons

The sharing buttons for sharing content on social media and other platforms displayed on this website are provided by the WordPress plugin AddToAny.

Personal data

This website collects personal data only to the extent necessary and in a way that is legally permissible.
Hereby, personal data includes all information which can be used to identify your personality, such as your name, your email address, or telephone number.

Contact data

In case you contact Quantum through any of the contact forms or other means offered on this website, the data entered in such a form, as well as the time at which this contact is made, is saved and processed.
Depending on the type of contact, your data will be suitably processed and shared with the appropriate boards or persons of Quantum.
Non of this data will be shared with third parties without your explicit consent.

User rights

You have the right to request information about which of your personal data is stored by Quantum and can demand that it be corrected (in case it is incorrect) or deleted (to the extend this is compatible with applicable law). To that end please use on of the channels described in the impressum.

5. Further aspects

Disclaimer of warranty

There is no warranty for the services provided by Quantum, to the extent permitted by applicable law. Except when otherwise stated in writing, the copyright holders and/or other parties provide these services “as is” without warranty of any kind, either expressed or implied, including, but not limited to, the implied warranties of merchantability and fitness for a particular purpose or future availability. The entire risk as to the quality and performance of the services is with the user. Should the services prove defective, the user assumes the cost of all damages incurred.

Limitation of liability

In no event, unless required by applicable law or agreed to in writing, will Quantum be liable to any user for damages, including any general, special, incidental or consequential damages arising out of the use or inability to use the services provided by Quantum, even if Quantum has been advised of the possibility of such damages.

External content

Links to other works, websites, or other documents, including but not limited to hyperlinks on the website quantum-journal.org as well as in manuscripts published by Quantum, may link to content that is beyond the control of Quantum. Quantum hence does not assume any kind of responsibility or liability for the content to which such links point or transmissions that can be received through them.

Infringement notification

In case you believe that any material published by Quantum infringes on your copyright or intellectual property rights in any way, you must contact in writing, in either English or German, the

Verein zur Förderung des Open Access Publizierens in den Quantenwissenschaften
Müllnergasse 26, 1090 Wien, Austria.

Logos, images and materials created by Quantum

The term “Quantum” and the Quantum logo are a registered trademark of the Verein zur Förderung des Open Access Publizierens in den Quantenwissenschaften in the European Union (EUIPO) for the publishing of scientific papers, electronic publishing, the publishing of journals, and several other categories of goods and services.

All other company and product names, logos, trademarks, registered trademarks, and brands are property of their respective owners. All such company, product and service names used in this website are for identification purposes only. Use of these names, logos, and brands does not imply endorsement.

The logo of Quantum, as well as all images created by Quantum, are published under the Creative Commons Attribution-NonCommercial-ShareAlike 4.0 International (CC BY-NC-SA 4.0) licence. The LaTeX template quantumarticle is published on GitHub under the LaTeX Project Public licence version 1.3c.

Right to modify terms and conditions

Quantum retains the right to modify these terms and conditions following consultation with the Steering Board. All submitters of works undergoing peer-review at the time of change must be notified if affected by the changes. Submitters are always bound to the version of these terms and conditions at the time of the last (re-)submission.

These terms and conditions, as well as the Editorial Policies and the Constitution of Quantum, are published under a Creative Commons "No Rights reserved" (CC0) licence. Sharing, tweaking and reuse of these policies is allowed and encouraged. If you adapt them to other projects, we would love to hear from you.

Real Randomized Benchmarking

A. K. Hashagen[1], S. T. Flammia[2,3], D. Gross[4], and J. J. Wallman[5]

[1]Department of Mathematics, Technical University of Munich, Germany
[2]Centre for Engineered Quantum Systems, School of Physics, University of Sydney, Sydney, Australia
[3]Yale Quantum Institute, Yale University, New Haven, Connecticut 06520, USA
[4]Institute for Theoretical Physics, University of Cologne, Germany
[5]Institute for Quantum Computing and Department of Applied Mathematics, University of Waterloo, Canada
August 9, 2018

Randomized benchmarking provides a tool for obtaining precise quantitative estimates of the average error rate of a physical quantum channel. Here we define *real randomized benchmarking*, which enables a separate determination of the average error rate in the real and complex parts of the channel. This provides more fine-grained information about average error rates with approximately the same cost as the standard protocol. The protocol requires only averaging over the real Clifford group, a subgroup of the full complex Clifford group, and makes use of the fact that it forms an orthogonal 2-design. It therefore allows benchmarking of fault-tolerant gates for an encoding which does not contain the full Clifford group transversally. Furthermore, our results are especially useful when considering quantum computations on rebits (or real encodings of complex computations), in which case the real Clifford group now plays the role of the complex Clifford group when studying stabilizer circuits.

1 Introduction

The design of reliable quantum information processing devices requires the quantitative characterization of the average error rate of a physical quantum channel. Full characterization of quantum processes is possible through quantum process tomography [29]. This method is, however, infeasible in practice. Firstly, it relies upon the challenging assumption that the set of measurements and the quantum state preparation admit lower errors than the process itself. Furthermore, the number of experimental configurations required – including quantum state preparation and quantum measurements – grows exponentially with the number of qubits even when employing improvements such as compressed sensing [17, 26].

An alternative approach is randomized benchmarking (RB) and variants thereof [8, 10, 11, 14, 15, 20, 37–39]. An RB protocol gives an estimate of the average fidelity between the realized and ideal implementations of a group of quantum gates by estimating the decay rate of the survival probability over random sequences of varying lengths. The effort of implementing the RB protocol scales efficiently with the number of qubits and it is robust against measurement and state preparation errors. Due to this, RB has become a popular tool to assess the quality of quantum processes [1, 3, 4, 9, 19, 34, 42, 45, 47, 54].

In this work, we study RB protocols in which the quantum gates are taken from the *real Clifford group*, which we refer to as *real randomized benchmarking*. We define the

notion of an *orthogonal 2-design*, and show that the real Clifford group constitutes one. This property allows one to efficiently estimate the average fidelity of an experimental implementation of the real Clifford group.

There are two primary motivations for using alternative groups for randomized benchmarking. First, some gates may be significantly worse than others due to different implementations (such as fault-tolerant implementations of non-transversal gates). Including such gates in the benchmarking group would result in a rapid decay dominated by the worst gate(s), so that little information can be obtained about the majority of gates. Furthermore, some quantum codes do not allow all transversal Clifford gates. The real Clifford group might, however, be accessible. This insight was recently used to do randomized benchmarking inside the code space of the [4,2,2] code using a variant of the protocol discussed in the present manuscript [28]. Second, the average gate fidelity quantifies the error rate over the entire Hilbert space. If an experiment only involves states in a portion of Hilbert space, then the relative figure of merit should only average over the states in that portion of Hilbert space. Real randomized benchmarking allows a direct characterization of the average gate fidelity over real-valued density operators, which is directly relevant to universal quantum computation with rebits [7, 46]. Third, information about which part of the Hilbert space is afflicted by the worst errors provides more information with which to optimize the experimental implementation of a group of quantum gates.

Summary. We analyze real randomized benchmarking, where the quantum gates are taken from the real Clifford group. The real Clifford group acting on n-qubits is generated by

$$\mathcal{C}(n) := \langle Z_i, H_i, CZ_{ij} \rangle ,$$

where the subscripts indicate that the gate is acting on the ith qubit and Z is the Pauli Z-gate, H is the Hadamard gate and CZ is the controlled Z-gate, defined respectively as,

$$Z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \qquad H = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}, \qquad CZ = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix}.$$

The protocol that gives an estimate of the average fidelity between the physical and ideal implementations of these gates, denoted as $\tilde{\mathcal{C}}$ and $\mathcal{C}$ respectively, is given in protocol 2 in section 7. We assume that the error quantum channel is gate and time independent throughout. However, we note that the methods of Wallman [50] and Merkel et al. [41] can be used to prove that the gate-dependent assumption can be relaxed with negligible effect on the estimate; we leave a careful and detailed proof of this to future work. The protocol estimates the decay rate of the survival probability over random gate sequences of varying length $m + 1$ as illustrated in fig. 1.

The protocol gives an approximation to the average sequence fidelity,

$$\bar{F}(m, E, \rho) = A + b^m B + c^m C,$$

where A, B and C depend only on the state preparation and measurement, and b and c depend on the noise quantum channel. Let $S_i : \mathcal{M}_d \to \mathcal{M}_d$, $i = a, b, c$, be defined as

$$\begin{aligned} S_a(\cdot) &= \mathrm{Tr}[\cdot]\frac{\mathbb{I}}{d}, \\ S_b(\cdot) &= \frac{1}{2}(\mathrm{id} + \theta) - \mathrm{Tr}[\cdot]\frac{\mathbb{I}}{d} \qquad \text{and} \\ S_c(\cdot) &= \frac{1}{2}(\mathrm{id} - \theta), \end{aligned}$$

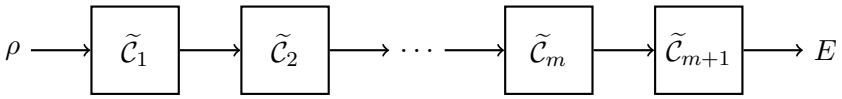


Figure 1: The main setup of real randomized benchmarking (see Protocol 2 for more details). For a fixed $m \in \mathbb{N}$, a sequence of $m+1$ real Clifford gates is applied to an initial quantum state ρ. The sequence is generated, such that in the case of its ideal implementation, it gives the identity operation. A subsequent measurement is performed given by an effect operator of a POVM, E, to measure the survival probability. Averaging over $M \in \mathbb{N}$ random realizations of sequences of length m gives the average sequence fidelity.

then

$$\begin{aligned} A &= \mathrm{Tr}[ES_a(\rho)], \\ B &= \mathrm{Tr}[ES_b(\rho)] \qquad \text{and} \\ C &= \mathrm{Tr}[ES_c(\rho)], \end{aligned}$$

as well as

$$\begin{aligned} b &= \frac{\mathrm{Tr}[T \circ S_b(\rho)]}{\mathrm{Tr}[S_b(\rho)]} \qquad \text{and} \\ c &= \frac{\mathrm{Tr}[T \circ S_c(\rho)]}{\mathrm{Tr}[S_c(\rho)]}, \end{aligned}$$

where $T : \mathcal{M}_d \to \mathcal{M}_d$ denotes the noise quantum channel.

The parameter b is linearly proportional to the average rebit fidelity, where the average is taken with respect to the orthogonal group, see eq. (35) in section 6. Together with parameter c, they give the average fidelity, see eq. (34). It is thus possible to obtain more fine-grained information about the physical implementation of real Clifford gates.

Organization of the paper. This paper starts with an introduction to the group twirl and in particular studies the twirl over the real orthogonal group. In this particular case, the exact form of a twirled quantum channel is derived. We establish the notion of an orthogonal 2-design and, after defining the real Clifford group, in section 3 we show that the real Clifford group is an orthogonal 2-design. The derived insights into the real Clifford group are then, in the following section 4, compared to the complex case. Section 5 gives a protocol on how to obtain a Haar sample from the real Clifford group in an efficient way. In section 6 we discuss the figures of merit of which real RB obtains estimates. In section 7 we give the real RB protocol that shows how to calibrate the average sequence fidelity to experimental data. Section 8 uses the results established in section 3 to derive the estimate for the average fidelity.

2 Group twirl

In this section we introduce the mathematical background necessary for real randomized benchmarking. We will first introduce the notation used throughout this paper followed by a review of the relevant representation theoretic concepts.

We consider an n-qubit system with an underlying finite-dimensional Hilbert space $\mathcal{H} \simeq \mathbb{C}^d$, $d = 2^n$. Denote by $\mathcal{M}_d(\mathbb{C})$ the set of complex-valued $d \times d$-matrices and as $\mathcal{M}_d(\mathbb{R})$ the set of real-valued $d \times d$-matrices. Every quantum state is described by a

density matrix $\rho \in \mathcal{M}_d(\mathbb{C})$, with normalization $\mathrm{Tr}[\rho] = 1$ and positivity property $\rho \geq 0$. The set of d-dimensional density matrices or quantum states is denoted as $\mathcal{D}_d := \{\rho \in \mathcal{M}_d(\mathbb{C}) | \rho \geq 0, \mathrm{Tr}[\rho] = 1\}$. A transformation of a quantum state is described by a quantum channel, which is a completely positive trace preserving linear map $T : \mathcal{M}_d(\mathbb{C}) \to \mathcal{M}_d(\mathbb{C})$. The Choi-Jamiołkowski representation [33] provides a one-to-one correspondence between linear maps $T : \mathcal{M}_d \to \mathcal{M}_{d'}$ and operators $\tau_T \in \mathcal{M}_{d'd}$ via

$$\tau_T = (\mathrm{id} \otimes T) |\Omega\rangle\langle\Omega|, \tag{1}$$

where $|\Omega\rangle\langle\Omega| = \frac{1}{d} \sum_{i,j=1}^{d} |ii\rangle\langle jj|$ is the maximally entangled state. This operator τ_T encodes every property of the linear map T and the representation shows that the set of quantum channels corresponds one-to-one to the set of bipartite quantum states which have one reduced density matrix maximally mixed [29]. This result will be used throughout this work. Denote by $\mathcal{U}(d)$ the unitary group acting on $\mathbb{C}^d$ and by $\mathcal{O}(d)$ the real orthogonal group acting on $\mathbb{C}^d$. Moreover, $\mathbb{I}$ is the identity matrix in $\mathcal{M}_d(\mathbb{C})$.

Throughout this paper, we are interested in group actions on quantum channels. For an extensive review of representations of finite and compact groups, please refer to [48]. To this end consider any finite group G with elements $g \in G$ and a unitary representation $\{U(g)\}_{g \in G}$ on $\mathbb{C}^d$. Its adjoint representation $\mathcal{U}_U : G \to \mathrm{End}(\mathcal{M}_d(\mathbb{C}))$ is defined through its action on any $X \in \mathcal{M}_d(\mathbb{C})$ as

$$\mathcal{U}_{U(g)}(X) = U(g) X U(g)^* \quad \forall g \in G, \tag{2}$$

and may be represented as a matrix $U \otimes \bar{U} \in \mathcal{M}_{d^2}$.

Indeed, let H be a general matrix group acting on some Hilbert space $\mathcal{K}$ (below, we will be interested in e.g. $H = \{U(g) \otimes \bar{U}(g) \,|\, g \in G\}$, with $\mathcal{K} = \mathbb{C}^d \otimes \mathbb{C}^d$). The Hilbert space decomposes as

$$\mathcal{K} \simeq \bigoplus_{i=1}^{k} \mathcal{K}_i \otimes \mathbb{C}^{n_i}, \tag{3}$$

where the sum is over irreducible unitary representations of H, $\mathcal{K}_i = \mathbb{C}^{d_i}$ carries the ith irreducible unitary representation, and n_i is the degeneracy of the irreducible unitary representations in $\mathcal{K}$. Every $U \in H$ is block-diagonal with respect to this decomposition, i.e. of the form

$$U \simeq \bigoplus_{i=1}^{k} U_i \otimes \mathbb{I}_{n_i \times n_i}. \tag{4}$$

The *commutant* H' of H is the algebra $H' = \{X \,|\, [X, U] = 0 \, \forall U \in H\}$ which commutes with all elements of H. By Schur's Lemma, every $X \in H'$ is of the form

$$X \simeq \bigoplus_{i=1}^{k} \mathbb{I}_{d_i \times d_i} \otimes X_i, \tag{5}$$

with $d_i = \dim \mathcal{K}_i$, and X_i acting on the n_i-dimensional space appearing on the right hand side of eq. (4). We will mainly restrict our attention to the case where all irreducible unitary representations of H on $\mathcal{K}$ are non-degenerate, i.e. $n_i = 1$ for all $i = \{1, \ldots, k\}$. In this case, eq. (5) takes the form

$$X \simeq \sum_{i=1}^{k} x_i P_i, \tag{6}$$

where the P_i are orthogonal projections onto the ith irreducible unitary representation and the $x_i \in \mathbb{C}$.

The *group twirl* associated with H is

$$\mathbb{T} : A \mapsto \int_H UAU^* \mathrm{d}U, \tag{7}$$

where the integration is w.r.t. the Haar measure on H. In particular, if H is finite, the integral is the normalized sum over the group. The group twirl is (i) idempotent, (ii) self-adjoint (w.r.t. the Hilbert-Schmidt inner product), and (iii) leaves elements $X \in H'$ of the commutant invariant. It is thus the orthogonal projection onto H'. In the non-degenerate case, one can check that this projection is given explicitly by

$$\mathbb{T}(A) = \sum_{i=1}^{k} \frac{1}{d_i} \operatorname{Tr}(AP_i)\, P_i. \tag{8}$$

Clearly, the group twirl over H only depends on the commutant H'. Thus, if a group S is such that $S' = H'$, twirling over S is equivalent to twirling over H. In practice, this freedom can be advantageous, if S has e.g. smaller cardinality than H, or simpler implementations as a quantum circuit. The notion of a *group design* captures this relation: A *unitary t-design* is any group G such that we have the equality of commutants

$$\{U^{\otimes t} \,|\, U \in G\}' = \{U^{\otimes t} \,|\, U \in \mathcal{U}(d)\}'. \tag{9}$$

General many-qubit unitaries do not have an efficient gate decomposition, while there are many-qubit unitary 2-designs and 3-designs that do. This was the original motivation for introducing the notion [11].

Phrased this way, it is natural to generalize eq. (9) to arbitrary "reference groups", beyond the now well-studied case of $\mathcal{U}(d)$. In particular, we will be concerned with the following case:

Definition 1. Let G be a matrix group acting on $\mathbb{C}^d$ for some d. Then G is an *orthogonal t-design* if we have the equality of commutants, i.e.,

$$\{U^{\otimes t} \,|\, U \in G\}' = \{O^{\otimes t} \,|\, O \in \mathcal{O}(d)\}',$$

where $\mathcal{O}(d)$ is the real orthogonal group acting on $\mathbb{C}^d$.

2.1 Group twirl over the orthogonal group

Throughout this paper, our main emphasis will be on orthogonal 2-designs. In this case, it is possible to work out the commutant easily [49]. Consider the unitary representation $O^{(2)} : g \mapsto O(g) \otimes O(g)$ of the orthogonal group $\mathcal{O}(d)$ on $\mathcal{H} = \mathbb{C}^d \otimes \mathbb{C}^d$, given as

$$G = \{O \otimes O | O \in \mathcal{O}(d)\}. \tag{10}$$

The commutant G' is spanned by three orthogonal projections [40, 49],

$$P_0 = |\Omega\rangle\langle\Omega|, \tag{11a}$$

$$P_1 = \frac{1}{2}(\mathbb{I} - \mathbb{F}) \quad \text{and} \tag{11b}$$

$$P_2 = \frac{1}{2}(\mathbb{I} + \mathbb{F}) - |\Omega\rangle\langle\Omega|, \tag{11c}$$

where $\mathbb{F} = \sum_{i,j=1}^{d} |ij\rangle\langle ji|$ is the flip (or swap) operator and $|\Omega\rangle\langle\Omega| = \frac{1}{d}\sum_{i,j=1}^{d} |ii\rangle\langle jj|$ is the maximally entangled state. For any symmetric $X \in \mathcal{M}_d(\mathbb{C})$, we have that $\mathbb{F}X =$

X, and for any antisymmetric $X \in \mathcal{M}_d(\mathbb{C})$, we get $\mathbb{F}X = -X$. The projections thus correspond to multiples of the identity, antisymmetric matrices and traceless symmetric matrices respectively. Every density operator in the commutant must thus be in the convex hull of the corresponding normalized density matrices $\rho_i = P_i/d_i$ for $i = 0, 1, 2$.

The theory discussed above therefore applies to quantum channels too. To this end, let T be a quantum channel on a d-dimensional quantum system, and let G be a matrix group on $\mathbb{C}^d$. The *twirled* channel $\tilde{T}$ over the full real orthogonal group,

$$\tilde{T}(\cdot) = \int_{\mathcal{O}(d)} OT\left(O^* \cdot O\right) O^* \, \mathrm{d}O, \tag{12}$$

can then be expressed as in eq. (7). Using the state channel duality, we see that $\rho_0 = P_0/d_0$ then corresponds to the ideal channel $T(\cdot) = \mathrm{id}$, $\rho_1 = P_1/d_1$ corresponds to the Werner-Holevo channel given by

$$T(\cdot) = \frac{\mathrm{Tr}\left[\cdot\right]\mathbb{I} - \theta}{d-1},$$

where θ denotes the usual transposition $\rho \mapsto \theta(\rho) := \rho^T$, and $\sum_i P_i/d^2$ corresponds to the completely depolarizing channel $T(\cdot) = \mathrm{Tr}\left[\cdot\right]\mathbb{I}/d$. This yields

$$\tilde{T}(\cdot) = \alpha\,\mathrm{id} + \beta\frac{\mathbb{I}}{d}\,\mathrm{Tr}\left[\cdot\right] + \gamma\frac{\mathbb{I}\,\mathrm{Tr}\left[\cdot\right] - \theta}{d-1}, \tag{13}$$

with $\alpha, \beta, \gamma \in \mathbb{R}$ satisfying $\alpha + \beta + \gamma = 1$ (which makes $\tilde{T}$ trace-preserving). As before, θ denotes the usual transposition $\rho \mapsto \theta(\rho) := \rho^T$. This immediately follows from the correspondence between a quantum channel and its Jamiołkowski state,

$$\begin{aligned}
&\int_{\mathcal{O}(d)} (O \otimes O)\, \tau_T\, (O \otimes O)^* \, \mathrm{d}O \\
&= \int_{\mathcal{O}(d)} (O \otimes O)\, (\mathrm{id} \otimes T)\, |\Omega\rangle\langle\Omega|\, (O \otimes O)^* \, \mathrm{d}O \\
&= \int_{\mathcal{O}(d)} \frac{1}{d} \sum_{i,j=1}^{d} O\,|i\rangle\langle j|\,O^* \otimes OT\left(|i\rangle\langle j|\right) O^* \, \mathrm{d}O \\
&= \int_{\mathcal{O}(d)} \frac{1}{d} \sum_{i,j=1}^{d} |i\rangle\langle j| \otimes OT\left(O^*\,|i\rangle\langle j|\,O\right) O^* \, \mathrm{d}O \\
&= \left(\mathrm{id} \otimes \int_{\mathcal{O}(d)} OT\left(O^* \cdot O\right) O^* \, \mathrm{d}O\right) |\Omega\rangle\langle\Omega| \,.
\end{aligned}$$

A twirl over a quantum channel and the corresponding twirl over its Jamiołkowski quantum state are thus equivalent descriptions. We will, however, focus on the twirling of quantum channels throughout the rest of this paper, where we will use that the theory allows us to give an explicit representation of the image of a channel under an orthogonal twirl.

3 Real Clifford group

We now define the real Clifford group and show that it is an orthogonal 2-design. Following [44], we define the real Pauli group $E(n)$ on n qubits as the n-fold tensor power

$$E(n) := \left\langle E(1)^{\otimes n} \right\rangle, \tag{14}$$

where $E(1)$ is just the real Pauli group on 1 qubit defined as

$$E(1) := \left\langle X := \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, Z := \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \right\rangle . \tag{15}$$

$E(n)$ is thus generated by tensor products of the Pauli matrices X and Z with 2×2 identity matrices $\mathbb{I}_2$.

Definition 2 (Real Clifford group)**.** The real Clifford group $\mathcal{C}(n)$ is the normalizer in $\mathcal{O}(2^n)$ of the real Pauli group $E(n)$, i.e.

$$\mathcal{C}(n) := \{O \in \mathcal{O}(2^n) | OE(n) = E(n)O\} . \tag{16}$$

In the simple case when $n = 1$ the real Clifford group is generated by

$$\mathcal{C}(1) = \left\langle Z := \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, H := \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix} \right\rangle , \tag{17}$$

and in the case when $n = 2$ the real Clifford group is

$$\mathcal{C}(2) = \left\langle \mathcal{C}(1) \otimes \mathcal{C}(1), CZ := \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix} \right\rangle , \tag{18}$$

where H is the Hadamard gate and CZ is the controlled Z-gate. See [44] for a thorough discussion of the real Clifford group.

Theorem 3. *The representation $O^{(2)} : g \mapsto O(g) \otimes O(g)$ of the real Clifford group $\mathcal{C}(n)$ decomposes into three non-degenerate irreducible unitary representations.*

Proof. See [44, theorem 6.8.1.] and proofs therein. □

We are now equipped to give the theorem that acts as the main mathematical ingredient for real randomized benchmarking.

Theorem 4. *The real Clifford group $\mathcal{C}(n)$ is an orthogonal 2-design.*

Proof. The real Clifford group $\mathcal{C}(n)$ is a subgroup of the real orthogonal group $\mathcal{O}(2^n)$. Its commutant therefore contains the commutant of the real orthogonal group. By theorem 3 these two commutants have the same dimensions, and must thus be equal. This proves the claim. □

Theorem 4 will be the main ingredient for real RB. It is, however, possible to also prove the following interesting fact about the real Clifford group.

Proposition 5. *The real Clifford group $\mathcal{C}(n)$ is an orthogonal 3-design, but it is not an orthogonal 4-design.*

Proof. See [44, notes below theorem 6.8.1] for the fact that the real Clifford group $\mathcal{C}(n)$ is an orthogonal 3-design. See [43, corollary 4.13] for the fact that it is not an orthogonal 4-design, where it is shown that in this case the commutant has an additional element to it. □

4 Complex Clifford group

The real Clifford group shares the properties observed in the last chapter with its complex counterpart. The complex Clifford group is a unitary 2-design, a unitary 3-design, but fails to be an exact unitary 4-design [27, 31, 36, 53, 55, 56]. This, however, is not a coincidence. To this end, let us first define the complex Clifford group.

The Pauli group on one qubit $\mathcal{P}(1)$ is defined as the group generated by

$$\mathcal{P}(1) := \langle X, Z, iI \rangle , \tag{19}$$

where I, X, Y, Z are the standard Pauli matrices given as

$$I := \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \quad X := \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad Y := \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} \text{ and } \quad Z := \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}. \tag{20}$$

The Pauli group on n qubits is defined to be

$$\mathcal{P}(n) := \mathcal{P}(1)^{\otimes n}. \tag{21}$$

Definition 6 (Complex Clifford group)**.** The complex Clifford group $\mathcal{X}(n)$ is the group-theoretic normalizer in the unitary group $\mathcal{U}(2^n)$ of the Pauli group $\mathcal{P}(n)$, i.e.

$$\mathcal{X}(n) := \{U \in \mathcal{U}(2^n) | U\mathcal{P}(n) = \mathcal{P}(n)U\} . \tag{22}$$

In the simple case where $n = 1$ the complex Clifford group is therefore just given as

$$\mathcal{X}(1) = \langle H, P \rangle , \tag{23}$$

where H is the Hadamard gate defined in eq. (17) and P is the $\pi/4$-phase gate given as

$$P := \begin{pmatrix} 1 & 0 \\ 0 & i \end{pmatrix}. \tag{24}$$

In the simple case when $n = 2$ the complex Clifford group is just given as

$$\mathcal{X}(2) = \langle \mathcal{X}(1) \otimes \mathcal{X}(1), CZ \rangle , \tag{25}$$

where CZ is again the controlled Z-gate defined in eq. (18).

The real Clifford group is therefore a subgroup of the complex Clifford group, $\mathcal{C}(n) \subset \mathcal{X}(n)$, and we necessarily have that for any $t \in \mathbb{N}$,

$$\{O^{\otimes t} | O \in \mathcal{C}(n)\}' \supset \{U^{\otimes t} | U \in \mathcal{X}(n)\}'.$$

Similarly, the orthogonal group is a subgroup of the unitary group, $\mathcal{O}(2^n) \subset \mathcal{U}(2^n)$, and we thus have that

$$\{O^{\otimes t} | O \in \mathcal{O}(2^n)\}' \supset \{U^{\otimes t} | U \in \mathcal{U}(2^n)\}'.$$

In the special case $t = 2$ (and in fact $t = 3$), we get the following correspondence:

$$\begin{array}{ccc} \{O^{\otimes 2} | O \in \mathcal{C}(n)\}' & \supset & \{U^{\otimes 2} | U \in \mathcal{X}(n)\}' \\ \| & & \| \\ \{O^{\otimes 2} | O \in \mathcal{O}(2^n)\}' & \supset & \{U^{\otimes 2} | U \in \mathcal{U}(2^n)\}' \end{array}$$

In [44], using the language of invariant harmonic polynomials, it was shown that any harmonic polynomial p_A that is invariant w.r.t. the complex Clifford group, must be invariant w.r.t. any of its subgroups, including the real Clifford group. If we decompose the harmonic polynomial into its real and imaginary parts, then the restrictions $p_{\mathrm{Re}(A)}$ and $p_{\mathrm{Im}(A)}$ must be invariant harmonic polynomials of the real Clifford group. Unfortunately, the resulting real polynomials may turn out to be zero. It is therefore not possible to infer the absence of harmonic invariants of $\mathcal{X}(n)$ from the absence of real harmonic invariants of $\mathcal{C}(n)$. However, it explains the observation regarding the real and complex Clifford group and their t-design properties for $t = 2, 3$.

If some family of matrix groups G acting on $\mathbb{C}^d$ fails to be a unitary t-design, it might still turn out to be useful for RB. If its commutant, $\{U^{\otimes t}|U \in G\}'$, has l additional elements to it than the commutant of the unitary group, $\{U^{\otimes t}|U \in \mathcal{U}(d)\}'$, and there are only $l = O(1)$ additional elements to the commutant, then we may term the family of matrix groups G an *algebraic almost t-design*.

An example of an algebraic almost unitary t-design is the family of Clifford groups, which form an algebraic almost 4-design with one additional generator in its commutant compared to the unitary group [30, 31, 56]. As another example, the real Clifford group is an algebraic almost unitary 2-design (and an exact orthogonal 2-design, as discussed above). The dihedral-CNOT family of groups provides yet another example [10].

The l additional factors in the commutant are useful for RB because they yield l additional decay terms in the average fidelity. A successful fit to the multi-exponential decay in a benchmarking experiment would then yield finer-grained information about the average error rate by finding the average fidelity associated to the projections onto each of the commutant algebras. In these scenarios, however, stability is an issue in the case of large l as fitting a multi-exponential decay is in general poorly conditioned [12]. For the case discussed in the most detail in this paper, the real Clifford group, there is only one extra decay term, and successful data processing methods can be employed to fit the model that we derive below with an efficient number of measurements [22, 24]. Another avenue for dealing with multi-exponential decays is to consider state preparations and measurements that optimize the contrast between various competing terms, perhaps even canceling all but one (or a constant fraction of) the exponential decay terms [8, 18]. We explore this idea in more detail in section 8.

The next section answers the question of how to obtain a Haar sample from the real Clifford group in an efficient way. This is an important ingredient for the real RB protocol.

5 Haar sample from the real Clifford group

5.1 Structure of the real Clifford group, and orthogonal transformations

Here, we summarize results and notions from [6] and [2, Chapter 7].

Consider the *phase space* $V = \mathbb{F}_2^{2n}$. Elements of phase space will often be written as $(p, q) \in \mathbb{F}_2^{2n}$, with $p, q \in \mathbb{F}_2^n$. An important piece of structure for the real Clifford group [6] is the quadratic form

$$Q\left((p, q)\right) = p \cdot q = \sum_{i=1}^{n} p_i q_i. \tag{26}$$

For $x = (p, q), x' = (p', q')$, one checks that

$$Q(x + x') - Q(x) - Q(x') = p \cdot q' - p' \cdot q = [x, x'], \tag{27}$$

where the square brackets denote the standard *symplectic form* on phase space. (While over $\mathbb{F}_2$, $-1 = +1$, we occasionally use negative signs when these would appear for analogous calculations in odd characteristic). The form Q turns V into an *orthogonal space*.

A vector $x \in V$ is *singular* if $Q(x) = 0$. A *hyperbolic pair* is a set of two singular vectors $e, f \in V$ such that $[e, f] = 1$. A two-dimensional subspace is a *hyperbolic plane* if it is spanned by a hyperbolic pair.

A space $U \subset V$ is *totally singular* if Q and $[\cdot, \cdot]$ vanish on U. Clearly, $U = \{(p, 0) \,|\, p \in \mathbb{F}_2^n\}$ is totally singular and has dimension half of V. This, by definition, means that Q has *Witt index* n, or, equivalently *sign* $+1$. A $2n$-dimensional orthogonal space in characteristic two has sign $+1$ if and only if it is isometric to the orthogonal sum of n hyperbolic planes:

$$V = \bigoplus_{i=1}^{n} \langle \{e_i, f_i\} \rangle, \qquad Q(e_i) = Q(f_i) = [e_i, e_j] = [f_i, f_j] = 0, \qquad [e_i, f_j] = \delta_{i,j}. \tag{28}$$

The set of linear transformations $\mathrm{GL}(V)$ preserving such a quadratic form of positive sign (and hence, by eq. (27), the form $[\cdot, \cdot]$) is the group $O^+(2n, 2)$. By eq. (28), a matrix S represents an element of $O^+(2n, 2)$ with respect to a hyperbolic basis $\{e_1, f_1, \ldots, e_n, f_n\}$ if and only if it is symplectic and its columns are singular – i.e. if and only if its columns form again a hyperbolic basis.

Recall that the complex Clifford group up to Pauli operators is isomorphic to the symplectic group $\mathcal{X}(n)/P(n) \simeq \mathrm{Sp}(2n, 2)$ [35]. That is true in the sense that for each $U \in \mathcal{X}(n)$, there exists a $S \in \mathrm{Sp}(2n, 2)$ such that

$$UP(x)U^* \propto P(Sx),$$

where for $x = (p, q) \in \mathbb{F}_2^{2n}$, we have defined the *Pauli operator*

$$P(x) = i^{Q(x)} \bigotimes_{i=1}^{n} X^{q_i} Z^{p_i}. \tag{29}$$

Any two Cliffords U that differ by the left- or right-action of an element of the Pauli group induce the same S.

If $U/P(n)$ contains a real-valued matrix, then $S \in O^+(2n, 2)$, i.e., in addition to being symplectic, its columns are singular. Conversely, any element of the real Clifford group is associated with such an S, which again does change under left- or right-multiplication with elements from $E(n)$.

Thus, to sample from the real Clifford group, one can proceed by 1) drawing a random element from $O^+(2n, 2)$ (see below), 2) use one of the known constructions (e.g. [13, 32]) for generating a gate sequence that implements a given symplectic matrix as a Clifford operation, and 3) multiply with a randomly chosen element of $E(n)$.

5.2 Efficient sampling from $O^+(2n, 2)$

It remains to describe an efficient protocol for drawing an element S from $O^+(2n, 2)$ uniformly at random. This will be achieved as follows:

Protocol 1: Sampling($O^+(2n, 2)$)

Initialization Choose a basis $\mathcal{B}_1$ of $\mathbb{F}_2^{2n}$.

Iterate for $i = 1$ **to** n Choose random linear combinations x of the vectors in $\mathcal{B}_i$

until x is non-zero and singular. Set $e_i = x$. Choose random linear combinations y of the vectors in $\mathcal{B}_i$ until $[e_i, y] = 1$. If y is singular, set $f_i = y$. Else, set $f_i = e_i + y$. Choose a basis $\mathcal{B}_{i+1}$ for $\langle \{e_j, f_j\}_{j=1}^{i} \rangle^{\perp}$.

Result Return matrix S, with columns given by $e_1, f_1, e_2, f_2 \ldots, e_n, f_n$.

The following statements are true for this construction:

1. By definition, the span of $\mathcal{B}_1$ is an orthogonal space of sign $+1$ and dimension $2n$.

2. Assume $V_i = \langle \mathcal{B}_i \rangle$ spans an orthogonal space of sign $+1$ and dimension $2(n - i + 1)$. Then, in expectation, one will find a non-zero singular x after no more than 4 attempts. To see this, let $\{e'_j, f'_j\}_j$ be a hyperbolic basis for V_i. Then x will be of the form

$$x = \sum_{j=1}^{n-i+1} \left(p_j e'_j + q_j f'_j \right),$$

with the p_j, q_j drawn uniformly at random. Hence

$$Q(x) = \sum_{j=1}^{n-i+1} \left(p_j Q(e'_j) + q_j Q(f'_j) + p_j q_j [e'_j, f'_j] \right) = \sum_{j=1}^{n-i+1} p_j q_j,$$

which is zero with probability at least $\frac{1}{2}$, while at least $1 - 2^{-n} \geq \frac{1}{2}$ of these cases correspond to non-zero x.

3. Under the same assumption as before, a vector y with $[e_i, y] = 1$ will be found after an expected number of 2 attempts. This is because $|\ker(y \mapsto [e_i, y])| = 2^{\dim(V_i)-1}$ and thus exactly half of all vectors in V do not lie in the kernel.

4. The vectors $\{e_i, f_i\}$ form a hyperbolic pair. Indeed, e_i is singular by construction. If y is singular, so is f_i. If $Q(y) = 1$, then

$$Q(f_i) = Q(e_i + y) = Q(e_i) + Q(y) + [e_i, y] = 0 + 1 + 1 = 0.$$

Also, $[e_i, y] = [e_i, y + e_i] = 1$.

5. The basis $\mathcal{B}_{i+1}$ of $\langle \{e_j, f_j\}_{j=1}^{i} \rangle^{\perp}$ describes the solution space of a set of liner equations over $\mathbb{F}_2$, and can thus be found efficiently. By eq. (28), it spans an orthogonal space of sign $+1$ and dimension $2(n - (i + 1) + 1)$.

Hence, by induction, the columns of S form a hyperbolic basis of $\mathbb{F}_2^{2n}$, and every such basis is equally likely to arise this way. The above procedure thus samples uniformly from $O^+(2n, 2)$.

The next chapters use the result that the real Clifford group is an orthogonal 2-design to analyze real RB using gates from the real Clifford group.

6 Figures of merit

The main goal of RB is to quantify how close a physical quantum channel $\widetilde{\mathcal{C}}$ is to the ideal quantum gate $\mathcal{C}$. In order to do so, we seek a figure of merit assessing this quality in an

efficient way. We can always write the physical quantum channel as a composition of the ideal quantum channel with an error quantum channel, $\widetilde{\mathcal{C}} = \mathcal{C} \circ \mathcal{E}$ [29]. We assess the quality of the physical quantum channel using the average fidelity

$$\bar{F}\left(\mathcal{C}, \widetilde{\mathcal{C}}\right) = \int_{\mathcal{U}(d)} \mathrm{Tr}\left[\mathcal{C}\left(U\left|0\right\rangle\!\left\langle 0\right|U^*\right)\widetilde{\mathcal{C}}\left(U\left|0\right\rangle\!\left\langle 0\right|U^*\right)\right] \mathrm{d}U, \tag{30}$$

and the average rebit fidelity

$$\bar{F}^{\mathbb{R}}\left(\mathcal{C}, \widetilde{\mathcal{C}}\right) = \int_{\mathcal{O}(d)} \mathrm{Tr}\left[\mathcal{C}\left(O\left|0\right\rangle\!\left\langle 0\right|O^*\right)\widetilde{\mathcal{C}}\left(O\left|0\right\rangle\!\left\langle 0\right|O^*\right)\right] \mathrm{d}O. \tag{31}$$

Please notice that in the case of the average rebit fidelity, the average is taken with respect to the orthogonal group. The fidelity is thus averaged over rebits. For quantum gates $\mathcal{C}(\cdot) = C \cdot C^*$ with C unitary, this simplifies to

$$\begin{aligned}\bar{F}\left(\mathcal{C}, \widetilde{\mathcal{C}}\right) &= \int_{\mathcal{U}(d)} \mathrm{Tr}\left[C\left(U\left|0\right\rangle\!\left\langle 0\right|U^*\right)C^*C\mathcal{E}\left(U\left|0\right\rangle\!\left\langle 0\right|U^*\right)C^*\right] \mathrm{d}U \\ &= \int_{\mathcal{U}(d)} \left\langle 0\right|U^*\mathcal{E}\left(U\left|0\right\rangle\!\left\langle 0\right|U^*\right)U\left|0\right\rangle \mathrm{d}U = \bar{F}\left(\mathcal{E}, \mathrm{id}\right),\end{aligned} \tag{32}$$

and similarly for the average rebit fidelity to

$$\bar{F}^{\mathbb{R}}\left(\mathcal{C}, \widetilde{\mathcal{C}}\right) = \int_{\mathcal{O}(d)} \left\langle 0\right|O^*\mathcal{E}\left(O\left|0\right\rangle\!\left\langle 0\right|O^*\right)O\left|0\right\rangle \mathrm{d}O = \bar{F}^{\mathbb{R}}\left(\mathcal{E}, \mathrm{id}\right). \tag{33}$$

These are the quantities that are related to the two parameters, which RB can estimate,

$$\bar{F}\left(\mathcal{E}, \mathrm{id}\right) = \frac{b\left(d^2 + d - 2\right) + cd\left(d-1\right) + 2\left(d+1\right)}{2d\left(d+1\right)} \tag{34}$$

as well as

$$\bar{F}^{\mathbb{R}}\left(\mathcal{E}, \mathrm{id}\right) = \frac{b\left(d-1\right) + 1}{d}. \tag{35}$$

For real density matrices the action of the twirled channel thus reduces to the action of the depolarizing channel with parameter b, and the above average gate fidelities can be interpreted as fidelities restricted to the respective commutant spaces. This makes our analysis especially interesting when considering quantum computations on rebits. It has been shown that universality holds in this case [46] and the real Clifford group now plays the role of the complex Clifford group when studying stabilizer circuits [7].

The next chapter gives the real RB protocol, which has to be executed to quantitatively describe the quality of a sequence of physical quantum gates taken from the real Clifford group.

7 Real randomized benchmarking protocol

The real randomized benchmarking protocol is given in the following. The real RB protocol considers quantum gates taken from the real Clifford group. We assume that the error quantum channel is both gate and time independent. We follow the notation of [38].

We will first describe the protocol for a given state preparation ρ, sequence length m, and final measurement E. This defines a protocol that can be repeated many times to obtain data with those labels, (m, E, ρ). Averaging these data gives a *fidelity decay curve*, which in expectation is a function only of these three data labels and the noise channel,

which is assumed to be gate and time independent. This forms the core of real RB. In the subsequent section, we will show one way to process these data to obtain accurate estimates of the parameters of the decay curve without having to fit multi-exponential decays, which is possible but in practice quite challenging. In the following, we give the real RB protocol:

Protocol 2: RealRB(m, E, ρ)

Step 1 Fix a positive integer $m \in \mathbb{N}$ that varies with every loop.

Step 2 Generate a sequence of $m+1$ quantum gates taken from the real Clifford group, i.e., $\mathcal{C}_1, \ldots, \mathcal{C}_{m+1}$, where $\mathcal{C}_j(\cdot) = C_j \cdot C_j^*$, $C_j \in \mathcal{C}(n)$, for $j = 1, \ldots, m+1$. The first m quantum gates, $\mathcal{C}_1, \ldots, \mathcal{C}_m$, are chosen independent and uniformly at random from the real Clifford group. The final quantum gate, $\mathcal{C}_{m+1}$, is chosen from the real Clifford group, such that the net sequence (if realized without errors) is the identity operation,

$$\mathcal{C}_{m+1} \circ \mathcal{C}_m \circ \ldots \circ \mathcal{C}_2 \circ \mathcal{C}_1 = \mathrm{id}, \tag{36}$$

where $\circ$ represents composition. The entire sequence is therefore given by

$$\mathcal{S}_m = \bigcirc_{j=1}^{m+1} \mathcal{C}_j \circ \mathcal{E}, \tag{37}$$

where $\mathcal{E}$ is the associated error, a completely positive trace preserving linear map.

Step 3 For each sequence, measure the survival probability given by the fidelity

$$F(m, E, \rho) = \mathrm{Tr}\left[E \mathcal{S}_m(\rho)\right], \tag{38}$$

where ρ is the initial quantum state, taking into account preparation errors, and E is an effect operator of a POVM taking into account measurement errors.

Step 4 Repeat steps 2-3 and average over M random realizations of the sequence of length m to find the averaged sequence fidelity

$$\bar{F}(m, E, \rho) = \mathrm{Tr}\left[E \bar{\mathcal{S}}_m(\rho)\right], \tag{39}$$

where

$$\bar{\mathcal{S}}_m = \frac{1}{M} \sum_m \mathcal{S}_m \tag{40}$$

is the average sequence operation.

Repeating these steps many times and averaging the results gives a good approximation to the average sequence fidelity,

$$\bar{F}(m, E, \rho) = A + b^m B + c^m C, \tag{41}$$

where A, B and C, given in eq. (58) below, depend only on the state preparation and measurement, and b and c depend on the average noise channel.

Please note that the final quantum gate $\mathcal{C}_{m+1}$ can be found efficiently by the Gottesman-Knill theorem [21].

The question of how to choose the sequence lengths m and the number of sequences at each length M is still unanswered, but is addressed in [16, 23, 30, 51]. The sequence length m should be exponentially spaced from 4, in order to avoid gate-dependent effects [41, 52], to around $1/(1-\bar{F}^{\mathbb{R}})$, for optimal information gain [23].

In the next section, we analyze RB focusing on the real Clifford group. The fact that it is an orthogonal 2-design will prove that it obeys the model given in eq. (41) when the assumptions of the noise model hold. Finally, we will discuss parameter estimation.

8 Fidelity decay and parameter estimation

The main setup of real RB is illustrated in fig. 1. A sequence of $m+1$ quantum gates $\tilde{\mathcal{C}}_j$ acts on an initial quantum state $\rho \in \mathcal{D}_d$, followed by a measurement represented by a POVM with effect operators E. Consider the physical quantum channel

$$\widetilde{\mathcal{C}}_j = \mathcal{C}_j \circ \mathcal{E}, \tag{42}$$

where $\mathcal{C}_j(\cdot) = C_j \cdot C_j^*$, $C_j \in \mathcal{C}(n)$, is a real Clifford gate and $\mathcal{E} : \mathcal{M}_d(\mathbb{C}) \to \mathcal{M}_d(\mathbb{C})$ is the associated error, a completely positive trace preserving linear map. We assume that the error quantum channel is both gate and time independent.

Let us first derive the decay curve of the expected data, eq. (41).

Theorem 7. *The protocol RealRB(m, E, ρ) has an expected fidelity of the form*

$$\bar{F}(m, E, \rho) = A + b^m B + c^m C, \tag{43}$$

where A, B and C are functions only of (E, ρ), and b and c depend on the average noise channel.

Proof. The expected fidelity of the above described sequence is given by

$$\begin{aligned} F(m, E, \rho) &= \mathrm{Tr}\left[E\left[\widetilde{\mathcal{C}}_{m+1} \circ \widetilde{\mathcal{C}}_m \circ \ldots \circ \widetilde{\mathcal{C}}_2 \circ \widetilde{\mathcal{C}}_1\right](\rho)\right] \\ &= \mathrm{Tr}\left[E\left[\mathcal{C}_{m+1} \circ \mathcal{E} \circ \mathcal{C}_m \circ \mathcal{E} \circ \ldots \circ \mathcal{C}_2 \circ \mathcal{E} \circ \mathcal{C}_1 \circ \mathcal{E}\right](\rho)\right]. \end{aligned} \tag{44}$$

Absorbing the first error quantum channel into the state as a preparation error gives

$$\begin{aligned} F(m, E, \rho) &= \mathrm{Tr}\left[E\left[\mathcal{C}_{m+1} \circ \mathcal{E} \circ \mathcal{C}_m \circ \mathcal{E} \circ \ldots \circ \mathcal{C}_2 \circ \mathcal{E} \circ \mathcal{C}_1\right](\rho)\right] \\ &= \mathrm{Tr}\left[E\left[\mathcal{S}_m\right](\rho)\right], \end{aligned} \tag{45}$$

with

$$\begin{aligned} S_m =& \mathcal{C}_{m+1} \circ \mathcal{E} \circ \mathcal{C}_m \circ \mathcal{E} \circ \ldots \circ \mathcal{C}_2 \circ \mathcal{E} \circ \mathcal{C}_1 \\ =& \overbrace{\mathcal{C}_{m+1} \circ (\mathcal{C}_m \circ \ldots \circ \mathcal{C}_1}^{=\mathbb{I}} \circ \overbrace{\mathcal{C}_1^* \circ \ldots \circ \mathcal{C}_m^*)}^{=\mathcal{D}_m^*} \circ \mathcal{E} \circ \mathcal{C}_m \circ \mathcal{E} \circ \\ & \ldots \mathcal{E} \circ \underbrace{\mathcal{C}_3 \circ (\mathcal{C}_2 \circ \mathcal{C}_1}_{=\mathcal{D}_3} \circ \underbrace{\mathcal{C}_1^* \circ \mathcal{C}_2^*)}_{=\mathcal{D}_2^*} \circ \mathcal{E} \circ \underbrace{\mathcal{C}_2 \circ (\mathcal{C}_1}_{=\mathcal{D}_2} \circ \underbrace{\mathcal{C}_1^*)}_{=D_1^*} \circ \mathcal{E} \circ \underbrace{\mathcal{C}_1}_{=\mathcal{D}_1} \\ =& \mathcal{D}_m^* \circ \mathcal{E} \circ \mathcal{D}_m \circ \ldots \circ \mathcal{D}_2^* \circ \mathcal{E} \circ \mathcal{D}_2 \circ \mathcal{D}_1^* \circ \mathcal{E} \circ \mathcal{D}_1 \\ =& \bigcirc_{j=1}^{m} \left(\mathcal{D}_j^* \circ \mathcal{E} \circ \mathcal{D}_j\right), \end{aligned} \tag{46}$$

where we have used the fact that $\mathcal{C}(n)$ is a group and defined a new quantum gate

$$\mathcal{D}_j := \bigcirc_{l=1}^{j} \mathcal{C}_l, \tag{47}$$

with $\mathcal{D}(\cdot) = D_j \cdot D_j^*$, $D_j \in \mathcal{C}(n)$. Please note that all $\mathcal{D}_j$ are independent uniformly distributed real Clifford gates.

The sequence fidelity can then be written as

$$F(m, E, \rho) = \operatorname{Tr}\left[E \bigcirc_{j=1}^{m} \left[\mathcal{D}_j^* \circ \mathcal{E} \circ \mathcal{D}_j\right](\rho)\right]. \tag{48}$$

Taking the average over the real Clifford group yields an average sequence fidelity given by

$$\begin{aligned}
\bar{F}(m, E, \rho) &= \operatorname{Tr}\left[E \frac{1}{|\mathcal{C}(n)|} \sum_{\mathcal{D}_j \in \mathcal{C}(n)} \mathcal{S}_m(\rho)\right] \\
&= \operatorname{Tr}\left[E \frac{1}{|\mathcal{C}(n)|} \sum_{\mathcal{D}_j \in \mathcal{C}(n)} \bigcirc_{j=1}^{m} \left[\mathcal{D}_j^* \circ \mathcal{E} \circ \mathcal{D}_j\right](\rho)\right] \\
&= \operatorname{Tr}\left[E \left[\frac{1}{|\mathcal{C}(n)|} \sum_{\mathcal{D}_j \in \mathcal{C}(n)} \mathcal{D}_j^* \circ \mathcal{E} \circ \mathcal{D}_j\right]^{\circ m}(\rho)\right],
\end{aligned} \tag{49}$$

where we have used the fact that all $\mathcal{D}_j$ are independent uniformly distributed Clifford gates. Because the real Clifford group is an orthogonal 2-design, see theorem 4, we have that

$$\frac{1}{|\mathcal{C}(n)|} \sum_{\mathcal{D}_j \in \mathcal{C}_m} \mathcal{D}_j^* \circ \mathcal{E} \circ \mathcal{D}_j(\cdot) = \int_{\mathcal{O}(d)} \mathcal{O}^* \circ \mathcal{E} \circ \mathcal{O}(\cdot)\, \mathrm{d}\mathcal{O}, \tag{50}$$

where $\mathcal{O}(\cdot) = O \cdot O^*$, $O \in \mathcal{O}(d)$, is the orthogonal quantum channel. By eq. (13), the averaged sequence fidelity is

$$\bar{F}(m, E, \rho) = \operatorname{Tr}\left[E \left(\alpha\, \mathrm{id} + \beta \frac{\mathbb{I}}{d} \operatorname{Tr}[\cdot] + \gamma \frac{\mathbb{I} \operatorname{Tr}[\cdot] - \theta}{d-1}\right)^{\circ m}(\rho)\right], \tag{51}$$

with $\alpha, \beta, \gamma \in \mathbb{R}$. Consider the following change of variables,

$$\alpha = \frac{1}{2}(b + c), \tag{52a}$$

$$\beta = a - b + \frac{1}{2}(b - c)d \quad \text{and} \tag{52b}$$

$$\gamma = \frac{1}{2}(c - b)(d - 1), \tag{52c}$$

such that

$$a = \alpha + \beta + \gamma, \tag{53a}$$

$$b = \alpha - \frac{\gamma}{(d-1)} \quad \text{and} \tag{53b}$$

$$c = \alpha + \frac{\gamma}{(d-1)}. \tag{53c}$$

Note that $a = \alpha + \beta + \gamma = 1$. The resulting quantum channel is

$$\tilde{T}(\cdot) = a \operatorname{Tr}[\cdot] \frac{\mathbb{I}}{d} + b\left(\frac{1}{2}(\mathrm{id} + \theta) - \operatorname{Tr}(\cdot) \frac{\mathbb{I}}{d}\right) + c \frac{1}{2}(\mathrm{id} - \theta), \tag{54}$$

with $\alpha, \beta, \gamma \in \mathbb{R}$. Its m-fold concatenation is given by

$$\tilde{T}^m(\cdot) = a^m \operatorname{Tr}[\cdot] \frac{\mathbb{I}}{d} + b^m\left(\frac{1}{2}(\mathrm{id} + \theta) - \operatorname{Tr}(\cdot) \frac{\mathbb{I}}{d}\right) + c^m \frac{1}{2}(\mathrm{id} - \theta). \tag{55}$$

Therefore, the averaged sequence fidelity is

$$\begin{aligned}
&\bar{F}(m, E, \rho) \\
&= \mathrm{Tr}\left[E\left(\alpha\,\mathrm{id} + \beta\frac{\mathbb{I}}{d}\,\mathrm{Tr}\left[\cdot\right] + \gamma\frac{\mathbb{I}\,\mathrm{Tr}\left[\cdot\right] - \theta}{d-1}\right)^{\circ m}\rho\right] \\
&= \mathrm{Tr}\left[E\left(a\,\mathrm{Tr}\left[\cdot\right]\frac{\mathbb{I}}{d} + b\left(\frac{1}{2}\left(\mathrm{id} + \theta\right) - \mathrm{Tr}\left(\cdot\right)\frac{\mathbb{I}}{d}\right) + c\frac{1}{2}\left(\mathrm{id} - \theta\right)\right)^{\circ m}\rho\right] \\
&= \mathrm{Tr}\left[E\left(a^m\,\mathrm{Tr}\left[\cdot\right]\frac{\mathbb{I}}{d} + b^m\left(\frac{1}{2}\left(\mathrm{id} + \theta\right) - \mathrm{Tr}\left(\cdot\right)\frac{\mathbb{I}}{d}\right) + c^m\frac{1}{2}\left(\mathrm{id} - \theta\right)\right)\rho\right] \\
&= a^m\,\mathrm{Tr}\left[E\frac{\mathbb{I}}{d}\right] + b^m\,\mathrm{Tr}\left[E\left(\frac{1}{2}\left(\rho + \rho^T\right) - \frac{\mathbb{I}}{d}\right)\right] + c^m\,\mathrm{Tr}\left[E\frac{1}{2}\left(\rho - \rho^T\right)\right]. \qquad (56)
\end{aligned}$$

Given that $a = 1$, the averaged sequence fidelity simplifies to

$$\bar{F}(m, E, \rho) = A + b^m B + c^m C, \qquad (57)$$

where

$$A = \mathrm{Tr}\left[E\frac{\mathbb{I}}{d}\right], \qquad (58a)$$

$$B = \mathrm{Tr}\left[E\left(\frac{1}{2}\left(\rho + \rho^T\right) - \frac{\mathbb{I}}{d}\right)\right] \quad \text{and} \qquad (58b)$$

$$C = \mathrm{Tr}\left[E\frac{1}{2}\left(\rho - \rho^T\right)\right], \qquad (58c)$$

for any fixed prepared quantum state ρ and any fixed quantum measurement represented by a POVM with effect operators E. As claimed, A, B and C are independent of the noise channel. □

The experimental data may thus be calibrated to this model using the real RB protocol discussed in section 7 by varying the parameter m and fitting the parameters b and c. It is important to notice that any quantum state preparation errors and quantum measurement errors are absorbed by A, B and C and the calibration of the model to experimental data is thus not affected by these errors. However, direct fitting is sometimes poorly conditioned in multi-exponential decays, and obtaining high accuracy can be challenging. We next describe a way to improve the contrast by choosing several different states and measurements and fitting to simpler decay curves by clever averaging. We make the simplifying assumption that the noise on the state preparations and measurements is independent of the particular state preparation, or the particular measurement. While this will not hold exactly in practice, in the regimes of interest there should still be a marked increase in statistical contrast when applying this idea [8, 18].

The standard approach in RB is to prepare eigenstates of Z on each qubit and then measure the POVM element that corresponds to $+1$ eigenstates of Z on each qubit. In the ideal case, this choice eliminates C (since ρ is symmetric). This reduces the number of parameters that need to be fit for the same data, which typically leads to a higher quality fit. We can generalize this idea so that we can accurately estimate b and c separately without coupling to all of the nuisance parameters A, B, and C at once.

The idea is to randomly compile in an extra gate that effectively prepares different Pauli eigenstates, rather than just a $+1$ eigenstate of Z on each qubit. First notice that every Pauli eigenstate is either complex symmetric ($\rho = \rho^T$) or antisymmetric ($\rho = -\rho^T$). The latter only holds if the Pauli eigenstate has an odd number of Y gates. So we should

choose from these initial states if we want to isolate the parameters b and c. Similarly, the POVM element E can be randomly chosen to be the $+1$ eigenstate or the -1 eigenstate of each given Pauli input eigenstate. For a given sequence length m, this defines four possible data sets. Letting $\beta_m(\rho, E)$ denote the data collected at sequence length m with state preparations $\rho_\pm$ that are symmetric or antisymmetric, and with ± 1 eigenstate projectors $E_\pm$, we have the four data sets

$$\beta_m(\rho_+, E_+), \quad \beta_m(\rho_+, E_-), \quad \beta_m(\rho_-, E_+), \quad \beta_m(\rho_-, E_-)\,. \tag{59}$$

Each of these obeys (in expectation) the respective form of the expected fidelity decay eq. (57). Now we can look at the symmetries of the nuisance parameters A, B, and C from eq. (58) and we see that linear combinations of the data sets can be chosen to eliminate one or more of the above parameters in the ideal case. In fact, we have that the following differences in data approach the following fidelity difference curves:

$$\beta_m(\rho_+, E_+) - \beta_m(\rho_+, E_-) \;\to\; \bar{F}(m, \rho_+, E_+) - \bar{F}(m, \rho_+, E_-) = \Delta B\, b^m \tag{60a}$$

$$\beta_m(\rho_-, E_+) - \beta_m(\rho_-, E_-) \;\to\; \bar{F}(m, \rho_-, E_+) - \bar{F}(m, \rho_-, E_-) = \Delta C\, c^m\,, \tag{60b}$$

where $\Delta B = \mathrm{Tr}[E\rho_+]$ and $\Delta C = \mathrm{Tr}[E\rho_-]$. In practice, there will not be exact cancellation, but this transformation will still greatly enhance the contrast. Therefore, collecting data at various values of m and fitting to the model $\Delta B\, b^m$ or $\Delta C\, c^m$ respectively yields a much simpler *exponential* fit model, and standard tools from regression can be applied. Each of these fits then yields a separate estimate of the parameters b and c without a strong covariance between the estimates. Estimates of the average rebit fidelity as well as the average fidelity are thus obtained by eq. (35) and eq. (34) respectively. Real RB thus provides finer-grained information about the channel.

Acknowledgments

AKH would like to give special thanks to Stephen D. Bartlett as this project was initialized when she visited his quantum physics research group at the University of Sydney. AKH would also like to thank Daniel Stilck França for many useful comments. Her work is supported by the Elite Network of Bavaria through the PhD program of excellence *Exploring Quantum Matter*. DG has received support from the Excellence Initiative of the German Federal and State Governments (Grant ZUK 81), Universities Australia and DAAD's Joint Research Co-operation Scheme (using funds provided by the German Federal Ministry of Education and Research), and the DFG (project B01 of CRC 183). Further, this work was supported by the Australian Research Council via EQuS project number CE11001013, by the US Army Research Office grant numbers W911NF-14-1-0098 and W911NF-14-1-0103, by the Australia-Germany Joint Research Co-operation Scheme, and by an Australian Research Council Future Fellowship FT130101744. This research was undertaken thanks in part to funding from TQT, CIFAR, the Government of Ontario, and the Government of Canada through CFREF, NSERC and Industry Canada.

After this paper was completed, we learned of closely related independent work by Brown and Eastin [5] that derives randomized benchmarking protocols for certain subgroups of the Clifford group.

References

[1] S. Asaad, C. Dickel, N. K. Langford, S. Poletto, A. Bruno, M. A. Rol, D. Deurloo, and L. DiCarlo. Independent, extensible control of same-frequency superconducting qubits

by selective broadcasting. *npj Quantum Inf.*, 2:16029, Aug 2016. DOI: 10.1038/npjqi.2016.29.

[2] M. Aschbacher. *Finite group theory*, volume 10. Cambridge University Press, 2000. DOI: 10.1017/CBO9781139175319.

[3] R. Barends, J. Kelly, A. Megrant, A. Veitia, D. Sank, E. Jeffrey, T. C. White, J. Mutus, A. G. Fowler, B. Campbell, Y. Chen, Z. Chen, B. Chiaro, A. Dunsworth, C. Neill, P. O'Malley, P. Roushan, A. Vainsencher, J. Wenner, A. N. Korotkov, A. N. Cleland, and J. M. Martinis. Superconducting quantum circuits at the surface code threshold for fault tolerance. *Nature*, 508:500–503, Apr 2014. DOI: 10.1038/nature13171.

[4] K. R. Brown, A. C. Wilson, Y. Colombe, C. Ospelkaus, A. M. Meier, E. Knill, D. Leibfried, and D. J. Wineland. Single-qubit-gate error below $\mathbf{10^{-4}}$ in a trapped ion. *Phys. Rev. A*, 84:030303, Sep 2011. DOI: 10.1103/PhysRevA.84.030303.

[5] W. G. Brown and B. Eastin. Randomized benchmarking with restricted gate sets. *Phys. Rev. A*, 97:062323, 2018. DOI: 10.1103/PhysRevA.97.062323.

[6] A. R. Calderbank, P. J. Cameron, W. M. Kantor, and J. J. Seidel. Z4-Kerdock codes, orthogonal spreads, and extremal Euclidean line-sets. In *Proceedings of the London Mathematical Society*, volume 75, pages 436–480. Cambridge University Press, 1997. DOI: 10.1112/S0024611597000403.

[7] A. R. Calderbank, E. M. Rains, P. W. Shor, and N. J. A. Sloane. Quantum error correction and orthogonal geometry. *Phys. Rev. Lett.*, 78:405–408, Jan 1997. DOI: 10.1103/PhysRevLett.78.405.

[8] A. Carignan-Dugas, J. J. Wallman, and J. Emerson. Characterizing universal gate sets via dihedral benchmarking. *Phys. Rev. A*, 92:060302, Dec 2015. DOI: 10.1103/PhysRevA.92.060302.

[9] J. M. Chow, J. M. Gambetta, L. Tornberg, J. Koch, L. S. Bishop, A. A. Houck, B. R. Johnson, L. Frunzio, S. M. Girvin, and R. J. Schoelkopf. Randomized benchmarking and process tomography for gate errors in a solid-state qubit. *Phys. Rev. Lett.*, 102: 090502, Mar 2009. DOI: 10.1103/PhysRevLett.102.090502.

[10] A. W. Cross, E. Magesan, L. S. Bishop, J. A. Smolin, and J. M. Gambetta. Scalable randomized benchmarking of non-Clifford gates. *npj Quantum Inf.*, 2:16012, Apr 2016. DOI: 10.1038/npjqi.2016.12.

[11] C. Dankert, R. Cleve, J. Emerson, and E. Livine. Exact and approximate unitary 2-designs and their application to fidelity estimation. *Phys. Rev. A*, 80:012304, Jul 2009. DOI: 10.1103/PhysRevA.80.012304.

[12] P. De Groen and B. De Moor. The fit of a sum of exponentials to noisy data. *J. Comput. Appl. Math.*, 20:175–187, 1987. DOI: 10.1016/0377-0427(87)90135-X.

[13] J. Dehaene and B. De Moor. Clifford group, stabilizer states, and linear and quadratic operations over GF(2). *Phys. Rev. A*, 68:042318, Oct 2003. DOI: 10.1103/PhysRevA.68.042318.

[14] J. Emerson, R. Alicki, and K. Zyczkowski. Scalable noise estimation with random unitary operators. *J. Opt. B*, 7(10):S347, 2005. DOI: 10.1088/1464-4266/7/10/021.

[15] J. Emerson, M. Silva, O. Moussa, C. Ryan, M. Laforest, J. Baugh, D. G. Cory, and R. Laflamme. Symmetrized characterization of noisy quantum processes. *Science*, 317 (5846):1893–1896, 2007. DOI: 10.1126/science.1145699.

[16] J. M. Epstein, A. W. Cross, E. Magesan, and J. M. Gambetta. Investigating the limits of randomized benchmarking protocols. *Phys. Rev. A*, 89(6):062321, Jun 2014. DOI: 10.1103/PhysRevA.89.062321.

[17] S. T. Flammia, D. Gross, Y. Liu, and J. Eisert. Quantum tomography via compressed

sensing: error bounds, sample complexity and efficient estimators. *New J. Phys.*, 14 (9):095022, 2012. DOI: 10.1088/1367-2630/14/9/095022.

[18] M. A. Fogarty, M. Veldhorst, R. Harper, C. H. Yang, S. D. Bartlett, S. T. Flammia, and A. S. Dzurak. Nonexponential fidelity decay in randomized benchmarking with low-frequency noise. *Phys. Rev. A*, 92:022326, Aug 2015. DOI: 10.1103/PhysRevA.92.022326.

[19] J. P. Gaebler, A. M. Meier, T. R. Tan, R. Bowler, Y. Lin, D. Hanneke, J. D. Jost, J. P. Home, E. Knill, D. Leibfried, and D. J. Wineland. Randomized benchmarking of multiqubit gates. *Phys. Rev. Lett.*, 108:260503, Jun 2012. DOI: 10.1103/PhysRevLett.108.260503.

[20] J. M. Gambetta, A. D. Córcoles, S. T. Merkel, B. R. Johnson, J. A. Smolin, J. M. Chow, C. A. Ryan, C. Rigetti, S. Poletto, T. A. Ohki, M. B. Ketchen, and M. Steffen. Characterization of addressability by simultaneous randomized benchmarking. *Phys. Rev. Lett.*, 109:240504, Dec 2012. DOI: 10.1103/PhysRevLett.109.240504.

[21] D. Gottesman. The Heisenberg representation of quantum computers. In S. P. Corney, R. Delbourgo, and P. D. Jarvis, editors, *Proceedings of the XXII International Colloquium on Group theoretical methods in physics*, pages 32–43. Cambridge, MA, International Press, 1999.

[22] C. Granade. Learning multiexponential models with QInfer. `http://www.cgranade.com/blog/2016/10/07/rb-multiexponential.html`, Oct 2016.

[23] C. Granade, C. Ferrie, and D. G. Cory. Accelerated randomized benchmarking. *New J. Phys.*, 17(1):013042, Jan 2015. DOI: 10.1088/1367-2630/17/1/013042.

[24] C. Granade, C. Ferrie, I. Hincks, S. Casagrande, T. Alexander, J. Gross, M. Kononenko, and Y. Sanders. QInfer: Statistical inference software for quantum applications. *Quantum*, 1:5, Apr 2017. ISSN 2521-327X. DOI: 10.22331/q-2017-04-25-5.

[25] D. Gross, K. Audenaert, and J. Eisert. Evenly distributed unitaries: On the structure of unitary designs. *J. Math. Phys.*, 48(5):052104, 2007. DOI: 10.1063/1.2716992.

[26] D. Gross, Y. Liu, S. T. Flammia, S. Becker, and J. Eisert. Quantum state tomography via compressed sensing. *Phys. Rev. Lett.*, 105:150401, Oct 2010. DOI: 10.1103/PhysRevLett.105.150401.

[27] D. Gross, S. Nezami, and M. Walter. Schur-Weyl duality for the Clifford group with applications: Property testing, a robust Hudson Theorem, and de Finetti representations. *ArXiv e-prints: arXiv:1712.08628 [quant-ph]*, 2017.

[28] R. Harper and S. Flammia. Fault tolerance in the IBM Q Experience. *ArXiv e-prints: arXiv:1806.02359 [quant-ph]*, 2018.

[29] T. Heinosaari and M. Ziman. *The Mathematical Language of Quantum Theory: From Uncertainty to Entanglement*. Cambridge University Press, 2012. DOI: 10.1017/CBO9781139031103.

[30] J. Helsen, J. J. Wallman, S. T. Flammia, and S. Wehner. Multi-qubit randomized benchmarking using few samples. *ArXiv e-prints: arXiv:1701.04299 [quant-ph]*, Jan 2017.

[31] J. Helsen, J. J. Wallman, and S. Wehner. Representations of the multi-qubit Clifford group. *J. Math. Phys.*, 59, 2018. DOI: 10.1063/1.4997688.

[32] E. Hostens, J. Dehaene, and B. De Moor. Stabilizer states and Clifford operations for systems of arbitrary dimensions and modular arithmetic. *Phys. Rev. A*, 71:042315, Apr 2005. DOI: 10.1103/PhysRevA.71.042315.

[33] A. Jamiołkowski. Linear transformations which preserve trace and positive semidefiniteness of operators. *Rep. Math. Phys.*, 3:275–278, Dec 1972. DOI: 10.1016/0034-4877(72)90011-0.

[34] E. Knill, D. Leibfried, R. Reichle, J. Britton, R. B. Blakestad, J. D. Jost, C. Langer, R. Ozeri, S. Seidelin, and D. J. Wineland. Randomized benchmarking of quantum gates. *Phys. Rev. A*, 77:012307, Jan 2008. DOI: 10.1103/PhysRevA.77.012307.

[35] R. Koenig and J. A. Smolin. How to efficiently select an arbitrary Clifford group element. *J. Math. Phys.*, 55(12):122202, 2014. DOI: 10.1063/1.4903507.

[36] R. Kueng and D. Gross. Qubit stabilizer states are complex projective 3-designs. *ArXiv e-prints: arXiv:1510.02767 [quant-ph]*, 2015.

[37] B. Lévi, C. C. López, J. Emerson, and D. G. Cory. Efficient error characterization in quantum information processing. *Phys. Rev. A*, 75:022314, Feb 2007. DOI: 10.1103/PhysRevA.75.022314.

[38] E. Magesan, J. M. Gambetta, and J. Emerson. Robust randomized benchmarking of quantum processes. *Phys. Rev. Lett.*, 106:180504, 2011. DOI: 10.1103/PhysRevLett.106.180504.

[39] E. Magesan, J. M. Gambetta, and J. Emerson. Characterizing quantum gates via randomized benchmarking. *Phys. Rev. A*, 85:042311, Apr 2012. DOI: 10.1103/PhysRevA.85.042311.

[40] C. B. Mendl and M. M. Wolf. Unital quantum channels – convex structure and revivals of Birkhoff's theorem. *Commun. Math. Phys.*, 289(3):1057–1086, 2009. DOI: 10.1007/s00220-009-0824-2.

[41] S. T. Merkel, E. J. Pritchett, and B. H. Fong. Randomized benchmarking as convolution: Fourier analysis of gate dependent errors. *ArXiv e-prints: arXiv:1804.05951 [quant-ph]*, 2018.

[42] J. T. Muhonen, A. Laucht, S. Simmons, J. P. Dehollain, R. Kalra, F. E. Hudson, S. Freer, K. M. Itoh, D. N. Jamieson, J. C. McCallum, A. S. Dzurak, and A. Morello. Quantifying the quantum gate fidelity of single-atom spin qubits in silicon by randomized benchmarking. *J. Phys. Condens. Matter*, 27(15):154205, 2015. DOI: 10.1088/0953-8984/27/15/154205.

[43] G. Nebe, E. M. Rains, and N. J. A. Sloane. The invariants of the Clifford group. *Des. Codes Cryptogr.*, 24(1):99–122, Sep 2001. DOI: 10.1023/A:1011233615437.

[44] G. Nebe, E. M. Rains, and N. J. A. Sloane. *Self-Dual Codes and Invariant Theory*. Algorithms and Computation in Mathematics. Springer Berlin Heidelberg, 2006. DOI: 10.1007/3-540-30731-1.

[45] S. Olmschenk, R. Chicireanu, K. D. Nelson, and J. V. Porto. Randomized benchmarking of atomic qubits in an optical lattice. *New J. Phys.*, 12(11):113007, 2010. DOI: 10.1088/1367-2630/12/11/113007.

[46] T. Rudolph and L. Grover. A 2 rebit gate universal for quantum computing. *ArXiv e-prints: arXiv:quant-ph/0210187*, Oct 2002.

[47] C. A. Ryan, M. Laforest, and R. Laflamme. Randomized benchmarking of single- and multi-qubit control in liquid-state NMR quantum information processing. *New J. Phys.*, 11(1):013034, 2009. DOI: 10.1088/1367-2630/11/1/013034.

[48] B. Simon. *Representations of finite and compact groups*, volume 10 of *Graduate studies in mathematics*. American Mathematical Society, 1996. DOI: 10.1090/gsm/010.

[49] K. G. H. Vollbrecht and R. F. Werner. Entanglement measures under symmetry. *Phys. Rev. A*, 64:062307, Nov 2001. DOI: 10.1103/PhysRevA.64.062307.

[50] J. J. Wallman. Randomized benchmarking with gate-dependent noise. *Quantum*, 2: 47, Jan 2018. DOI: 10.22331/q-2018-01-29-47.

[51] J. J. Wallman and S. T. Flammia. Randomized benchmarking with confidence. *New J. Phys.*, 16(10):103032, 2014. DOI: 10.1088/1367-2630/16/10/103032.

[52] Joel Wallman. Randomized benchmarking with gate-dependent noise. *Quantum*, 2: 47, 2018. DOI: 10.22331/q-2018-01-29-47.

[53] Z. Webb. The Clifford group forms a unitary 3-design. *Quantum Inf. Comput.*, 16: 1379–1400, 2016. DOI: 10.26421/QIC16.15-16.

[54] T. Xia, M. Lichtman, K. Maller, A. W. Carr, M. J. Piotrowicz, L. Isenhower, and M. Saffman. Randomized benchmarking of single-qubit gates in a 2D array of neutral-atom qubits. *Phys. Rev. Lett.*, 114:100503, Mar 2015. DOI: 10.1103/PhysRevLett.114.100503.

[55] H. Zhu. Multiqubit Clifford groups are unitary 3-designs. *Phys. Rev. A*, 96(6):062336, 2017. DOI: 10.1103/PhysRevA.96.062336.

[56] H. Zhu, R. Kueng, M. Grassl, and D. Gross. The Clifford group fails gracefully to be a unitary 4-design. *ArXiv e-prints: arXiv:1609.08172 [quant-ph]*, Sep 2016.

Appendix

Following [25] it is also possible to use the frame potential and establish that the real Clifford group is an orthogonal 2-design.

Theorem 8 (Orthogonal frame potential)**.** *Let $D = \{O_k \in \mathcal{O}(d)\}_{k=1,\dots,K}$ be a set of orthogonal real matrices. Define the frame potential of D to be*

$$\mathcal{P}(D) = \frac{1}{K^2} \sum_{O_k, O_{k'} \in D} \left| \mathrm{Tr}\left[O_k^* O_{k'}\right] \right|^4 . \tag{61}$$

The set D is an orthogonal 2-design if and only if $\mathcal{P}(D) = 3$.

Proof. Following the idea and notation of [25], the group D is an orthogonal 2-design if and only if $\Delta := \tau_D - \tau_{\text{twirl}} = 0$, where

$$\begin{aligned}
\tau_D &= (\mathrm{id}_{d^2} \otimes T_D) \left|\Omega\right\rangle\!\left\langle\Omega\right| \\
&= \left(\mathrm{id}_{d^2} \otimes \frac{1}{K} \sum_{O_k \in D} (O_k \otimes O_k) \cdot (O_k \otimes O_k)^* \right) \left|\Omega\right\rangle\!\left\langle\Omega\right| , \\
\tau_{\text{twirl}} &= (\mathrm{id}_{d^2} \otimes T_{\text{twirl}}) \left|\Omega\right\rangle\!\left\langle\Omega\right| \\
&= \left(\mathrm{id}_{d^2} \otimes \int_{\mathcal{O}(d)} (O \otimes O) \cdot (O \otimes O)^* \, \mathrm{d}O \right) \left|\Omega\right\rangle\!\left\langle\Omega\right| .
\end{aligned}$$

In order to see what this means in terms of the frame potential, let us introduce a basis with regards to the minimal projections given in eq. (11). Within the subspace of P_0, we introduce an orthonormal basis $\left\{ \left|i_0\right\rangle_j \right\}_{j=1}^{d_0}$ with dimension $d_0 = \dim P_0$. Similarly, we introduce an orthonormal basis $\left\{ \left|i_1\right\rangle_j \right\}_{j=1}^{d_1}$ within the subspace of P_1 with dimension $d_1 = \dim P_1$ as well as an orthonormal basis $\left\{ \left|i_2\right\rangle_j \right\}_{j=1}^{d_2}$ within the subspace of P_2 with dimension $d_2 = \dim P_2$. These then form an orthonormal basis $\left\{ \left|i\right\rangle_j \right\}_{j=1}^{d^2}$ in $\mathbb{C}^d \otimes \mathbb{C}^d$. The maximally entangled state is then given by $\left|\Omega\right\rangle\!\left\langle\Omega\right|$ with

$$\left|\Omega\right\rangle = \frac{1}{d} \sum_{m=0}^{2} \sum_{i_m=1}^{d_m} \left|i_m\right\rangle \otimes \left|i_m\right\rangle .$$

Using this decomposition and using eq. (8), we see that

$$\begin{aligned}
\tau_{\text{twirl}} &= \left(\mathrm{id}_{d^2} \otimes \sum_{m=0}^{2} \frac{\mathrm{Tr}\left[\,\cdot\, P_m\right]}{\mathrm{Tr}\left[P_m\right]} P_m \right) \left|\Omega\right\rangle\!\left\langle\Omega\right| \\
&= \frac{1}{d^2} \sum_{m=0}^{2} \sum_{i_m, j_m=1}^{d_m} \left|i_m\right\rangle\!\left\langle j_m\right| \otimes \frac{\mathrm{Tr}\left[\left|i_m\right\rangle\!\left\langle j_m\right| P_m\right]}{\mathrm{Tr}\left[P_m\right]} P_m \\
&= \frac{1}{d^2} \sum_{m=0}^{2} \frac{1}{\mathrm{Tr}\left[P_m\right]} P_m \otimes P_m .
\end{aligned}$$

Furthermore,

$$\tau_D = \frac{1}{d^2} \sum_{i,j=1}^{d^2} \left|i\right\rangle\!\left\langle j\right| \otimes \frac{1}{K} \sum_{O_k \in D} (O_k \otimes O_k) \left|i\right\rangle\!\left\langle j\right| (O_k \otimes O_k)^* .$$

We will now show that $\|\Delta\|_2^2 := \mathrm{Tr}\left[|\Delta|^2\right] = 0$, from which the claim follows. We thus want to compute

$$\mathrm{Tr}\left[\Delta^*\Delta\right] = \mathrm{Tr}\left[\tau_{\mathrm{twirl}}^*\tau_{\mathrm{twirl}} - \tau_{\mathrm{twirl}}^*\tau_D - \tau_D^*\tau_{\mathrm{twirl}} + \tau_D^*\tau_D\right].$$

Its first term is easily calculated to give

$$\mathrm{Tr}\left[\tau_{\mathrm{twirl}}^*\tau_{\mathrm{twirl}}\right] = \frac{1}{d^4}\,\mathrm{Tr}\left[\sum_{m=0}^{2}\frac{1}{\mathrm{Tr}\left[P_m\right]^2}\left[P_m\otimes P_m\right]\right] = \frac{3}{d^4}.$$

The second and third term yield

$$\begin{aligned}\mathrm{Tr}\left[\tau_{\mathrm{twirl}}^*\tau_D\right] =& \frac{1}{d^4}\,\mathrm{Tr}\left[\left(\sum_{m=0}^{2}\frac{1}{\mathrm{Tr}\left[P_m\right]}P_m\otimes P_m\right)\right.\\ &\left.\left(\sum_{i,j=1}^{d^2}|i\rangle\langle j|\otimes\frac{1}{K}\sum_{O_k\in D}\left(O_k\otimes O_k\right)|i\rangle\langle j|\left(O_k\otimes O_k\right)^*\right)\right]\\ =& \frac{1}{d^4K}\sum_{m=0}^{2}\sum_{i,j=1}^{d^2}\frac{1}{\mathrm{Tr}\left[P_m\right]}\mathrm{Tr}\left[P_m|i\rangle\langle j|\right]\mathrm{Tr}\left[P_m\sum_{O_k\in D}\left(O_k\otimes O_k\right)|i\rangle\langle j|\left(O_k\otimes O_k\right)^*\right],\end{aligned}$$

where we have used the fact that P_m commutes with $(O_k\otimes O_k)$. We then have that

$$\begin{aligned}\mathrm{Tr}\left[\tau_{\mathrm{twirl}}^*\tau_D\right] =& \frac{1}{d^4K}\sum_{m=0}^{2}\sum_{i_m=1}^{d_m}\frac{1}{\mathrm{Tr}\left[P_m\right]}\mathrm{Tr}\left[\sum_{O_k\in D}\left(O_k\otimes O_k\right)P_m|i_m\rangle\langle i_m|\left(O_k\otimes O_k\right)^*\right]\\ =& \frac{1}{d^4}\sum_{m=0}^{2}\sum_{i_m=1}^{d_m}\frac{\mathrm{Tr}\left[P_m|i_m\rangle\langle i_m|\right]}{\mathrm{Tr}\left[P_m\right]}\\ =& \frac{3}{d^4}.\end{aligned}$$

The last term is

$$\begin{aligned}\mathrm{Tr}\left[\tau_D^*\tau_D\right] =& \frac{1}{d^4K^2}\sum_{i,j=1}^{d^2}\sum_{v,w=1}^{d^2}\mathrm{Tr}\left[|j\rangle\langle i|v\rangle\langle w|\right]\\ &\mathrm{Tr}\left[\sum_{O_k\in D}\sum_{O_{k'}\in D}\left(O_k\otimes O_k\right)|j\rangle\langle i|\left(O_k\otimes O_k\right)^*\left(O_{k'}\otimes O_{k'}\right)|v\rangle\langle w|\left(O_{k'}\otimes O_{k'}\right)^*\right]\\ =& \frac{1}{d^4K^2}\sum_{i,j=1}^{d^2}\sum_{O_k,O_{k'}\in D}\mathrm{Tr}\left[\left(O_k\otimes O_k\right)|j\rangle\langle i|\left(O_k\otimes O_k\right)^*\left(O_{k'}\otimes O_{k'}\right)|i\rangle\langle j|\left(O_{k'}\otimes O_{k'}\right)^*\right]\\ =& \frac{1}{d^4K^2}\sum_{O_k,O_{k'}\in D}\left|\mathrm{Tr}\,O_k^*O_{k'}\right|^4\\ =& \frac{\mathcal{P}(D)}{d^4}.\end{aligned}$$

The group D is an orthogonal 2-design if and only if $\Delta = 0$, which gives $\mathcal{P}(D) = 3$ and thus the claim follows. □

Theorem 9. *Let $D = \{O_k \in \mathcal{O}(d)\}_{k=1,\ldots,K}$ be a set of real orthogonal matrices with the symmetry $G = \{O_k\otimes O_k | O_k \in D\}$ affording the character ζ_G. Then the following are equivalent:*

1. *The set D is an orthogonal 2-design.*

2. *The symmetry has no more than three irreducible representations.*

3. *It holds that $\langle \zeta_G, \zeta_G \rangle = 3$.*

Before we can prove this theorem, let us recall the notion of a character, which plays a major role in the analysis of unitary representations. We mainly follow [48].

Definition 10 (Character)**.** Given any unitary representation $U : g \mapsto U(g)$ of a group G, we define its character by

$$\xi(g) := \mathrm{Tr}\left[U(g)\right]. \tag{62}$$

A character is called irreducible, if the unitary representation under consideration is irreducible.

Denote by $\left\{V^{(i)}\right\}_i$ the irreducible unitary representations of G and the corresponding irreducible characters by $\{\chi_i\}_i$. The irreducible characters are orthonormal in the group algebra [48] with the inner product given by

$$\langle \chi_i, \chi_j \rangle := \frac{1}{|G|} \sum_{g \in G} \overline{\chi_i(g)} \chi_j(g) = \delta_{ij}. \tag{63}$$

If a representation is the direct sum of subrepresentations, then the corresponding character is the sum of the characters of those subrepresentations. This holds true especially for the decomposition into irreducible representations.

Theorem 11 ([48, theorem III.2.4.])**.** *Every character ξ is of the form*

$$\xi = \sum_i n_i \chi_i \tag{64}$$

for nonnegative integers n_i and every such sum is the character of some representation.

Proof. The proof is given in [48], but is reproduced here for the convenience of the reader. It is a standard result in representation theory that any finite-dimensional unitary representation can be written as a direct sum of finite irreducible unitary representations [48, theorem II.2.3.],

$$U = \oplus_i n_i V^{(i)}, \tag{65}$$

for some n_i. But then we necessarily have that

$$\xi = \sum_i n_i \chi_i, \tag{66}$$

which yields the claim. □

Corollary 12. *If the character ξ has a decomposition as in theorem 11 given by eq. (64), then*

$$\langle \xi, \xi \rangle = \sum_i n_i^2. \tag{67}$$

Proof. Substituting the decomposition into the inner product gives

$$\begin{aligned}\langle \xi, \xi \rangle &= \left\langle \sum_i n_i \chi_i, \sum_j n_j \chi_j \right\rangle \\ &= \sum_{i,j} n_i n_j \langle \chi_i, \chi_j \rangle \\ &= \sum_{i,j} n_i n_j \delta_{i,j} \\ &= \sum_i n_i^2, \end{aligned} \tag{68}$$

where we have used the orthogonality relation of irreducible characters given in eq. (63). □

With these definitions and results at hand, we may now prove theorem 9, following the idea of [25].

Proof of theorem 9. Consider the unitary representation associated to the symmetry G as given in the theorem 9 and its afforded character denoted by ζ_G. The frame potential can be related to the character ζ_G by

$$\mathcal{P}(D) = \langle \zeta_G, \zeta_G \rangle . \tag{69}$$

Due to corollary 12 this must equal $\langle \zeta_G, \zeta_G \rangle = 3$ if and only if G has exactly three irreducible components. This in turn is equivalent to D being an orthogonal 2-design by theorem 8. The claim thus follows. □

Theorem 13. *The real Clifford group $\mathcal{C}(n)$ is an orthogonal 2-design.*

Proof. Given that the real Clifford group has three irreducible representations by theorem 3, it must be an orthogonal 2-design by theorem 9. □

E Contributed further article: Article 5

D. S. França and A. K. Hashagen
Approximate randomized benchmarking for finite groups
Journal of Physics A: Mathematical and Theoretical, 51(39):395302, August 2018

Summary of article 5: Approximate Randomized Benchmarking for Arbitrary Finite Groups [5]

In this paper we analyze randomized benchmarking for quantum gates that form a representation of an arbitrary finite group. The same symmetry idea can be applied. The protocol requires an averaging procedure giving rise to a twirl of a quantum channel. The twirled quantum channel is covariant with respect to that finite group and is thus an element of the commutant of the adjoint representation. The decomposition of the underlying Hilbert space into irreducible unitary representations therefore determines its structure. The sought after average fidelity is then just a simple function of the trace of the quantum channel, again yielding an exponential decay of the average fidelity.

Corollary E.1 ([5, corollary 14])**.** *Suppose we perform randomized benchmarking for a unitary representation U of a finite group $\mathbf{G}$ s.t. $U \otimes \bar{U} = \oplus_{\alpha \in \hat{G}} \left(\mathbb{C}^{d_\alpha} \otimes \mathbb{C}^{m_\alpha}\right)$ and a quantum channel T. Then there exist $\lambda_1, \ldots, \lambda_k \in \overline{B_1(0)}$ and $a_0, a_1, \ldots, a_k \in \mathbb{C}$ s.t.*

$$F(m, E, \rho) = a_0 + \sum_{i=1}^{k} a_k \lambda_i^m. \tag{E.1}$$

for $m \geq \max m_\alpha$. Moreover, $k \leq \sum_\alpha m_\alpha$ corresponds to the number of distinct eigenvalues of $\mathbb{T}(T)$ and λ_i are its eigenvalues.

Calibrating this exponential decay to experimental data thus yields estimates on the eigenvalues and therefore the spectrum of the twirled quantum channel. This allows us to obtain information about the noise in the system.

The randomized benchmarking protocol assumes that we are able to efficiently obtain Haar samples of the finite group. This might not be possible in practice for an arbitrary finite group. It is, however, possible to obtain approximate samples using Markov chain Monte Carlo methods. We therefore show how to apply these results to randomized benchmarking and we obtain a stability result for approximate randomized benchmarking.

Corollary E.2 ([5, Corollary 18])**.** *Let μ be the Haar measure on $\mathbf{G}$ and $\nu_1, \ldots, \nu_m$ probability measures on $\mathbf{G}$ s.t.*

$$\|\mu - \nu_k\|_1 \leq \epsilon_k, \tag{E.2}$$

for all $1 \leq k \leq m$ *and* $\epsilon_k \geq 0$. *Denote by* $\tilde{\nu}$ *the distribution of* $(\mathcal{D}_1, \ldots, \mathcal{D}_m)$ *if we pick the* $\mathcal{U}_k$ *independently from* ν_k. *Then*

$$|\operatorname{Tr}\left[\mathbb{T}_{\tilde{\nu},m}(T)(\rho)E\right] - F(m,E,\rho)| \leq 4\sqrt{\frac{\log(|\mathbf{G}|)}{1-|\mathbf{G}|^{-1}}\sum_{k=1}^{m}\epsilon_k}. \tag{E.3}$$

Furthermore, we consider randomized benchmarking with quantum gate sequences solely chosen from a set of generators closed under inversion and one additional arbitrary quantum gate. This increases the practicality of randomized benchmarking as it is unrealistic to assume that all quantum gates from a group may be implemented. Usually gates have to be broken down into generators and it is therefore a natural question how randomized benchmarking with generators performs. In this setting we make the additional assumption that the error quantum channel describing the noise in the system is almost covariant with respect to the group under consideration. We show that randomized benchmarking in this setting is possible if the set of generators is rapidly mixing.

Corollary E.3 ([5, corollary 21])**.** *Let* $\mathcal{S}_{b,m+b+1}$ *be the average gate sequence and* $\lambda \in [0,1)$. *Then for any POVM element* E *and state* $\rho \in \mathcal{M}_d$*:*

$$|\operatorname{Tr}\left[\mathbb{E}\left(S_{b,m+1}\right)(\rho)E\right] - F(m,E,\rho)| \leq \epsilon + \mathcal{O}\left(\delta^2\frac{\lambda}{1-\lambda}m\right). \tag{E.4}$$

We apply these results to two finite groups. The first one we consider is the subgroup of monomial unitary matrices, which nonzero entries are the n-th roots of unity only. The group of monomial unitary matrices is particularly interesting, because it contains the T-gate, which is necessary for universal quantum computing together with the Clifford gates. The algorithm is simulated for this subgroup with $n = 8$ and the error quantum channel is assumed to depolarize to a random state with probability $(1-p)$. The error analysis shows a good estimation of the average fidelity using this method.

The second example we consider is the full complex Clifford group, as it is the standard group when considering randomized benchmarking. It has the advantage that we only need to estimate one parameter in this case. The error quantum channel describing the noise is taken to be a convex combination of the identity and a random quantum channel such that it is approximately covariant. The numerical results clearly show that both methods, the approximate randomized benchmarking as well as the generator randomized benchmarking, are effective to estimate the average fidelity. The latter method is only applicable in a high fidelity regime, which is not restrictive, since this is the regime of interest in general.

Statement of individual contribution

This project originated from the previous paper on *Real Randomized Benchmarking* by Anna-Lena K. Hashagen, Prof. Dr. Steven T. Flammia, Prof. Dr. David Gross and Dr. Joel J. Wallman [3]. The idea was that instead of choosing a specific group and analyzing the randomized benchmarking protocol on a case by case basis, it would be of interest to give a general randomized

benchmarking protocol. Daniel Stilck França proof-read my article and this is how the idea of a joint project started with the aim to generalize the ideas from [3].

I was responsible for the whole section about the symmetry considerations and the analysis of the structure of the groups' commutant. The lemma [5, lemma 9] and corollary [5, corollary 10] are straightforward generalizations of principles used in my previous paper [3]. Daniel Stilck França and I quickly agreed on their formulation and presentation during discussions. I derived lemma [5, lemma 11] in the case of irreducibly covariant representations. Daniel Stilck França then generalized the proof to arbitrary representations of finite groups. The general randomized benchmarking protocol found in section [5, section 4] was formulated during a discussion and is of the spirit to the one presented in [3]. Theorem [5, theorem 13] and corollary [5, corollary 14] were proven during discussions at the blackboard in a joint effort. The section about approximate twirls [5, section 5] was proven solely by Daniel Stilck França. In order to overcome one of the practical issues of randomized benchmarking, we showed that randomized benchmarking is possible if only access to the generators and one more element of the group under consideration is given. The given protocol is the same as the one given earlier; it just accommodates the more complicated notation for the gate sequences. Daniel Stilck França was mainly responsible for this.

In this article we study two examples. I initially suggested the group of monomial unitary matrices as a possible application of our protocol and Daniel Stilck França refined this to the case of monomial matrices with entries that are n-th roots of unity, which, as a finite subgroup, fits our framework perfectly. We both suggested to apply our results to the complex Clifford group, as this is a natural choice within randomized benchmarking. Daniel Stilck França was solely responsible for all the numerics in this article. After the first draft of this article was written, we had many discussions on how to improve the presentation of this article and we both made minor and major changes to all parts of the article.

Daniel Stilck França is the principal author of this article. I, Anna-Lena Karolyn Hashagen, was extensively involved in all parts of this article, except the section about approximate twirls and the numerics.

Journal permission and article

Subject **Re: Permission to include article (DOI: https://doi.org/10.1088/1751-8121/aad6fa) in my dissertation**
From Permissions <permissions@iop.org>
To Anna-Lena Hashagen <hashagen@ma.tum.de>
Date 2018-08-30 18:41

Dear Dr Hashagen,

Thank you for your email and for taking the time to seek this permission.

When you transferred the copyright in your article to IOP, we granted back to you certain rights, including the right to include all or part of the Final Published Version of the article within any thesis or dissertation. Please note you may need to obtain separate permission for any third party content you included within your article.

Please include citation details, "© IOP Publishing. Reproduced with permission. All rights reserved" and for online use, a link to the Version of Record.

The only restriction is that if, at a later date, you wanted your thesis/dissertation to be published commercially, further permission would be required.

I wish you the best of luck with the completion of your dissertation.

Kind regards,

Isabella Formisano

Editorial Assistant

Copyright & Permissions Team
Gemma Alaway – Senior Rights & Permissions Adviser
Christina Colwell - Rights & Permissions Assistant

Contact Details
E-mail: permissions@iop.org<mailto:permissions@iop.org>
For further information about copyright and how to request permission: <https://emea01.safelinks.protection.outlook.com/?url=http%3A%2F%2Fiopscience.iop.org%2Fpage%2Fcopyright&data=02%7C01%7Cchristina.colwell%40iop.org%7Cc490136a905a4fd5d84f08d4f46893da%7Cf9ee42e6bad04e639115f704f9https://publishingsupport.iopscience.iop.org/copyright-journals/

<https://publishingsupport.iopscience.iop.org/copyright-journals/>
<https://publishingsupport.iopscience.iop.org/copyright-journals/>See also: https://publishingsupport.iopscience.iop.org/

Please see our Author Rights Policy https://publishingsupport.iopscience.iop.org/author-rights-policies/<https://emea01.safelinks.protection.outlook.com/?url=http%3A%2F%2Fioppublishing.org%2Fauthor-rights%2F&data=02%7C01%7Cchristina.colwell%40iop.org%7Cc490136a905a4fd5d84f08d4f46893da%7Cf9ee42e6bad04e639115f704f9ccceed%7C1%7C0%7C636402177379623514&sdata=h66
Please note: We do not provide signed permission forms as a separate attachment. Please print this email and provide it to your publisher as proof of permission.

Please note: Any statements made by IOP Publishing to the effect that authors do not need to get permission to use any content where IOP Publishing is not the publisher is not intended to constitute any sort of legal advice. Authors must make their own decisions as to the suitability of the content they are using and whether they require permission for it to be published within their article.

_____
From: Anna-Lena Hashagen <hashagen@ma.tum.de>
Sent: Tuesday, August 28, 2018 12:12
To: Permissions
Cc: hashagen@ma.tum.de
Subject: Permission to include article (DOI: https://doi.org/10.1088/1751-8121/aad6fa) in my dissertation

Dear Sir or Madam,

I am currently preparing a cumulative dissertation. I am the author of
the article

D. S. Franca and A. K. Hashagen
Approximate randomized benchmarking for finite groups
J. Phys. A: Math. Theor. 51 (2018) 395302

and I would like to ask for permission to include this article in my
dissertation.

Kind regards,
Anna-Lena Hashagen

--
Anna-Lena Hashagen

Mathematical Physics
Department of Mathematics
Technische Universität München
Boltzmannstraße 3
85748 Garching bei München
+49 89 289 17036

Quick check guide for our current author rights policy

As a Named Author of an article published with IOP Publishing, you are granted back certain reuse & depositing rights under the copyright form and Author Rights Policy.

This table is intended as a quick reference guide for our current Author Rights policy. You must read the full Author Rights Policy, to ensure you are adhering to the full terms and conditions of the policy, before posting online or reusing any content you published with IOP Publishing.

This Author Rights Policy only applies to some of our journals (those listed at the end of the Policy).

All other journals published by IOP Publishing have different author rights policies, please check our Author Rights Policy webpage for information. If this doesn't cover the journal you published your article in, please see this page which sets out our partners who have different policies which they handle themselves.

As an author, which version of my article may I post and when? See quick check guide below.

Author Rights	Preprint	Accepted Manuscript	Final Published Version	Further Info
Posting on Personal Website	Yes – anytime	Yes – no embargo	No	See full conditions in Policy.
Posting on employer's or institution's website	Yes – at anytime	Yes – 12 month embargo	No	See full conditions in Policy.
Posting on non-commercial institutional or subject repository	Yes – at anytime	Yes – 12 month embargo	No	See full conditions in Policy. See exceptions for HEFCE post-2014 REFpolicy requirements in polic
Posting on non-commercial Scientific Social Network	Yes – at anytime	Yes – 12 month embargo	No	See full conditions in Policy.
Posting on commercial Scientific Social Network	No (unless endorsed STM voluntary principles for article sharing)	No	No	See full conditions in Policy. Examples of commercial SSNs are ResearchGate, Mendeley*, Academia.edu
Author Rights	**Preprint**	**Accepted Manuscript**	**Final Published Version**	**Further Info**

***Posting in own private library on Mendeley**	Yes – no embargo	Yes – no embargo	No	See full conditions in Policy.
Posting on arXiv	Yes – anytime	Yes – 12 month embargo (NB. a few of our journals allow immediate posting)	No	See full conditions in Policy (including which journals allow immediate posting
Posting on bioRxiv	Yes – anytime	Yes – 12 month embargo	No	See full conditions in Policy.
Use all or part of the article without modification in personal compilations of own works (provided not created by third party publisher)	Yes	Yes	No	See full conditions in article copyright form
Making copies for teaching purposes	Yes	Yes	Yes	See full conditions in article copyright form
Include in a research thesis or dissertation (provided not published commercially)	Yes – at anytime	Yes	Yes	See full conditions in article copyright form
Make oral presentation of article & include summary/highlights of it in papers distributed at presentations	Yes	Yes	Yes	See full conditions in article copyright form
Include summary/highlights of article in conference proceedings	Yes	Yes	Yes	See full conditions in article copyright form
Use original figures allowed under the quota of the STM Permissions Guidelines if publishing an article with another STM signatory publisher	Yes – 3 figures only may be used	Yes – 3 figures only may be used	Yes – 3 figures only may be used	See full conditions in article copyright form & in STM Permissions Guidelines section.
Use original text allowed under the quota of the STM Permissions Guidelines if publishing an article with another STM signatory publisher	Yes – short text extracts (single text extracts of less than 400 words may be used (with a maximum of 800 words from a journal issue))	Yes – short text extracts (single text extracts of less than 400 words may be used (with a maximum of 800 words from a journal issue))	Yes – short text extracts (single text extracts of less than 400 words may be used (with a maximum of 800 words from a journal issue))	See full conditions in article copyright form & in STM Permissions Guidelines section

This page last updated June 2018.

IOP Publishing Journal of Physics A: Mathematical and Theoretical

J. Phys. A: Math. Theor. **51** (2018) 395302 (31pp) https://doi.org/10.1088/1751-8121/aad6fa

Approximate randomized benchmarking for finite groups

D S França and A K Hashagen

Department of Mathematics, Technical University of Munich, Germany

E-mail: dsfranca@mytum.de and hashagen@ma.tum.de

Received 22 March 2018, revised 24 July 2018
Accepted for publication 31 July 2018
Published 28 August 2018

Abstract

We investigate randomized benchmarking (RB) in a general setting with quantum gates that form a representation, not necessarily an irreducible one, of a finite group. We derive an estimate for the average fidelity, to which experimental data may then be calibrated. Furthermore, we establish that RB can be achieved by the sole implementation of quantum gates that generate the group as well as one additional arbitrary group element. In this case, we need to assume that the noise is close to being covariant. This yields a more practical approach to RB. Moreover, we show that RB is stable with respect to approximate Haar sampling for the sequences of gates. This opens up the possibility of using Markov chain Monte Carlo methods to obtain the random sequences of gates more efficiently. We demonstrate these results numerically using the well-studied example of the Clifford group as well as the group of monomial unitary matrices. For the latter, we focus on the subgroup with nonzero entries consisting of nth roots of unity, which contains T gates.

Keywords: randomized benchmarking, quantum gates, Clifford gates, monomial unitary, random walks on groups, fidelity estimation

(Some figures may appear in colour only in the online journal)

1. Introduction

One of the main obstacles to build reliable quantum computers is the need to implement quantum gates with high fidelity. Therefore, it is key to develop techniques to estimate the quality of quantum gates and thus certify the quality of a quantum computer. To this end, one could perform tomography for the underlying noise in the implementation and in principle obtain a complete description of it [1, 2]. However, in general, the number of measurements necessary to estimate for a complete tomography of the noise scales exponentially with the system size and is not a practical solution to the problem. Thus, it is vital to develop

techniques to estimate the level of noise in systems more efficiently, even if we only obtain partial information.

Randomized benchmarking (RB) is a protocol to estimate the average fidelity of a set of quantum gates forming a representation of a group [3–6]. The very important case of Clifford gates has already been widely studied and some rigorous results that show its efficiency under some noise scenarios are available [7, 8], such as when the noise is independent of the gate and time. Besides its efficiency, another highlight of the protocol is that it is robust against state preparation and measurement errors. This makes it very attractive from an experimental point of view and its applicability was demonstrated successfully [9–17].

In this work, we show how to extend these protocols to gates that are representations of a finite group[1]; these must not necessarily be irreducible or form a 2-design. Although other works, such as [18–21], already extended the protocol to other specific groups of interest, we focus on showing how to estimate the average fidelity based on properties of the particular representation at hand for arbitrary finite groups. To this end, we investigate the structure of quantum channels that are covariant under a unitary representation of a group and derive formulas for their average fidelity in terms of their spectra. We then show that one can use RB to estimate the average fidelity of these gates under the assumption that they are subject to time and gate independent noise.

In order for this procedure to be efficient, it is necessary that we may multiply, invert and sample uniformly distributed elements of the group efficiently and that the given representation does not decompose into too many irreducible unitary representations, as we will discuss in more detail later. This is the case for the well-studied case of Cliffords.

The usual RB protocol assumes that we can implement sequences of gates that are sampled from the Haar distribution of the group [3–6]. We further generalize the RB protocol by showing that it is possible to implement sequence gates that are approximately Haar distributed instead. Therefore, it is possible to use Markov chain Monte Carlo methods to obtain the samples, potentially more efficiently. This result is of independent interest to the RB literature, as it shows that the protocol is stable against small errors in the sampling.

Moreover, we show how one can perform RB by just implementing gates that generate the group and one additional random element from the group at each round of the protocol. Thus, this last gate will generally not be an element of the generators. Mostly considering generators provides a more natural framework to the protocol, as often one is only able to implement a certain number of gates that generate the group and must, therefore, decompose the gates into generators. However, this protocol works under the assumption that the noise affecting the gates is already close to being covariant with respect to (w.r.t.) the group and not for arbitrary quantum channels, as in the usual setting. Moreover, we still need the ability to implement one gate which might not be contained in the set of generators and still assume that the same quantum channel that describes the noise on the generators also describes the noise on this gate. To illustrate our techniques, we apply them to subgroups of the monomial unitary matrices, i.e. products of $d-$dimensional permutation and diagonal unitary matrices. These can be seen as a generalization of stabilizer groups [22]. We focus on the subgroup of monomial unitary matrices whose nonzero entries are roots of unity. We show that we only need to estimate two parameters and multiplying and inverting elements of it can be done in time $\mathrm{O}(d)$. Moreover, they include the T-gate, which is known to form a universal set for quantum computation together with the Clifford gates [23]. Therefore, one can use the protocol described

[1] Most of the results in this work can easily be extended to compact groups. However, as it is not clear that implementing the RB protocol for compact groups is relevant for applications and given that this would make some proofs less accessible, we restrict to finite groups here.

here to estimate the noise from T-gates more efficiently. We make numerical simulations for our protocol and these subgroups and show that it is able to reliably estimate the average gate fidelity. Moreover, we numerically compare our techniques based on approximate Haar samples and implementation of generators to the usual protocol for Cliffords and show that the three yield indistinguishable results in the high fidelity regime.

This paper is structured as follows: we start by fixing our notation and reviewing basic results on Markov chains and covariant quantum channels; needed in section 2. In section 3 we derive the average fidelity of quantum channels in terms of their spectra and we give basic results on the decay of the probability of measurement outcomes under covariant quantum channels. These form the basis for the RB protocol for general groups, which we discuss and analyze in section 4. In section 5 we prove that it is also possible to implement the protocol using approximate samples. We then discuss the generalized RB protocol based on implementing random sequences of gates that generate the group in section 6. In this section, we also discuss the conditions under which this protocol applies. Finally, in section 7, we apply our techniques to the subgroup of monomial unitary matrices and perform numerical experiments for it. In the same section, we also compare numerically the RB protocols developed here with the usual one in the case of the Clifford group.

2. Notation and preliminaries

We will be interested in finite dimensional quantum systems. Denote by $\mathcal{M}_d$ the space of $d \times d$ complex matrices. We will denote by $\mathcal{D}_d$ the set of d-dimensional quantum states, i.e. positive semi-definite matrices $\rho \in \mathcal{M}_d$ with trace 1. We will call a linear map $T : \mathcal{M}_d \to \mathcal{M}_{d'}$ a quantum channel if it is trace preserving and completely positive. We will denote the adjoint of a quantum channel T with respect to the Hilbert–Schmidt scalar product by T^*. We will call a collection of positive semidefinite matrices $\{E_i\}_{i=1}^l$ a positive operator valued measure (POVM) if the POVM elements E_i, called effect operators, sum up to the identity. Throughout this paper, we will use the channel-state duality that provides a one-to-one correspondence between a quantum channel $T : \mathcal{M}_d \to \mathcal{M}_d$ and its Choi–Jamiolkowski state $\tau_T \in \mathcal{M}_{d^2}$ obtained by letting T act on half of a maximally entangled state, i.e.

$$\tau_T := (T \otimes \mathrm{id}_d)\left(|\Omega\rangle\langle\Omega|\right), \tag{1}$$

where $|\Omega\rangle\langle\Omega| \in \mathcal{M}_{d^2}$ is a maximally entangled state, that is,

$$|\Omega\rangle\langle\Omega| = \frac{1}{d}\sum_{i,j=1}^{d} |ii\rangle\langle jj|, \tag{2}$$

where $\{|i\rangle\}_{i=1}^d$ is an orthonormal basis in $\mathbb{C}^d$. Please refer to [24] for more on these concepts. To measure the distance between two states we will use the Schatten $1-$norm for $A \in \mathcal{M}_d$, denoted by $\|\cdot\|_1$ and given by

$$\|A\|_1 := \mathrm{Tr}\left((A^\dagger A)^{\frac{1}{2}}\right), \tag{3}$$

where $\dagger$ denotes the adjoint. Then, given two states $\rho, \sigma \in \mathcal{D}_d$, their trace distance is given by $\|\rho - \sigma\|_1/2$. This norm on $\mathcal{M}_d$ induces a norm on linear operators $\Phi : \mathcal{M}_d \to \mathcal{M}_d$ through

$$\|\Phi\|_{1\to 1} := \sup_{X\in\mathcal{M}_d, X\neq 0} \frac{\|\Phi(X)\|_1}{\|X\|_1}. \tag{4}$$

Given a random quantum channel $T : \mathcal{M}_d \to \mathcal{M}_d$, we will denote its expectation value by $\mathbb{E}(T)$.

We will also need some basic facts from the representation theory of finite groups. We refer to e.g. [25] for more on this and the proofs of the statements we use here. We will be particularly interested in the commutant of the algebra generated by the group. To this end we introduce:

Definition 1 (Commutant). Let $\mathcal{A}$ be an algebra of operators on a Hilbert space $\mathcal{H}$. Then the commutant $\mathcal{A}'$ of $\mathcal{A}$ is defined by

$$\mathcal{A}' := \{B | BA = AB \text{ for all } A \in \mathcal{A}\}. \tag{5}$$

Recall that a function $U : G \to \mathcal{M}_d$ is called a unitary representation of a finite group G on a finite-dimensional Hilbert space $\mathcal{H} \simeq \mathbb{C}^d$ if we have for all $g_1, g_2 \in G$ that $U_{g_1} U_{g_2} = U_{g_1 g_2}$. We will denote the unitary corresponding to g by U_g. From basic results of representation theory, we know that there exists distinct $\alpha_1, \ldots, \alpha_k \in \hat{G}$, where $\hat{G}$ denotes the set of equivalence classes of irreducible unitary representations (irreps), such that the unitary representation can be written as a direct sum of irreps, i.e. $U \cong \oplus U^{\alpha_i} \otimes \mathbb{I}_{m_\alpha}$ with $m_\alpha > 0$ denoting the degeneracy of the α_ith irrep. The structure of the commutant is then described in the following theorem.

Theorem 2 ([25, theorem IX.11.2]). *Let U be a unitary representation of a finite group G on $\mathcal{H}$. Write $\mathcal{H} = \oplus_{\alpha \in \hat{G}} \left(\mathbb{C}^{d_\alpha} \otimes \mathbb{C}^{m_\alpha}\right)$ so that $U_g = \oplus_{i=1}^k U_g^{\alpha_i} \otimes \mathbb{I}_{m_\alpha}$ with $\{\alpha_i\}_{i=1}^k$ distinct elements in $\hat{G}$. Let $\mathcal{A}(U)$ be the algebra of operators generated by the $\{U_g\}_{g \in G}$, and $\mathcal{A}(U)'$ its commutant. Then*

$$\mathcal{A}(U) = \left\{ \oplus_{i=1}^k A_i \otimes \mathbb{I}_{m_\alpha} \middle| A_i \in \mathcal{M}_{d_{\alpha_i}} \right\}, \tag{6a}$$

$$\mathcal{A}(U)' = \left\{ \oplus_{i=1}^k \mathbb{I}_{d_{\alpha_i}} \otimes B_i \middle| B_i \in \mathcal{M}_{m_\alpha} \right\}. \tag{6b}$$

Given a finite group G, we will call the uniform probability distribution on it its Haar measure. For a proof of its existence and basic properties, we refer to [25, section VII.3]. Given some unitary representation $U : G \to \mathcal{M}_d$, we call the function $\chi : G \to \mathbb{C}$ given by $g \mapsto \mathrm{Tr}(U_g)$ the character of the representation. We will denote the character of an irreducible representation $\alpha \in \hat{G}$ by χ^α and remark that one can find the decomposition in theorem 2 through characters [25, section III.2].

2.1. Covariant quantum channels and twirls

The definition of covariance of quantum channels is central to the study of their symmetries and will be one of the building blocks of the generalized RB protocol:

Definition 3 (Covariant quantum channel [26]). A quantum channel $T : \mathcal{M}_d \to \mathcal{M}_d$ is covariant w.r.t. a unitary representation $U : G \to \mathcal{M}_d$ of a finite group G, if for all $g \in G$

$$T\left(U_g \cdot U_g^\dagger\right) = U_g T(\cdot) U_g^\dagger. \tag{7}$$

In general, one allows different unitary representations of the group in the input and output of the channel in the definition of covariance, but here we will restrict to the case when we have the same unitary representation. There are many different and equivalent characterizations of covariance. Here we mention that covariance is equivalent to the Choi–Jamiolkowski

state τ_T commuting with $U_g \otimes \bar{U}_g$ for all $g \in G$. To see this, note that given a unitary representation U of G we may define its adjoint representation $\mathcal{U} : G \to \mathrm{End}(\mathcal{M}_d)$ through its action on any $X \in \mathcal{M}_d$ by conjugation,

$$\mathcal{U}_g(X) = U_g X U_g^\dagger. \tag{8}$$

Through the Choi–Jamiolkowski isomorphism, it is easy to see that the adjoint representation is equivalent to the unitary representation $U_g \otimes \bar{U}_g \in \mathcal{M}_{d^2}$. As we can rephrase (7) as T commuting with the adjoint representation, this translates to the Choi–Jamiolkowski state commuting with $U_g \otimes \bar{U}_g$. This means in particular that we may use structural theorems, like theorem 2, to investigate covariant channels, as covariance implies that the channel is in the commutant of the adjoint representation.

Theorem 4. *Let $T : \mathcal{M}_d \to \mathcal{M}_d$ be a quantum channel that is covariant w.r.t. a unitary representation U of a finite group G and let $\oplus_{\alpha\in\hat{G}} \left(\mathbb{C}^{d_\alpha} \otimes \mathbb{C}^{m_\alpha}\right)$ be the decomposition of the underlying Hilbert space into irreps α of G with multiplicity m_α for the unitary representation $U \otimes \bar{U}$. Then:*

$$T = \oplus_{\alpha\in\hat{G}} \mathbb{I}_{d_\alpha} \otimes B_\alpha \tag{9}$$

with $B_\alpha \in \mathcal{M}_{m_\alpha}$.

Proof. As T is covariant, it must be an element of the commutant of the adjoint representation, i.e. $T \in \mathcal{A}(\mathcal{U})'$. The decomposition then follows from theorem 2. □

This decomposition further simplifies when no multiplicities in the decomposition of the unitary representations into its irreducible components are present. We call such channels irreducibly covariant. Here we briefly mention some of the results of [27], where the structure of such channels is investigated.

Theorem 5 ([27, theorem 40]). *A quantum channel $T : \mathcal{M}_d \to \mathcal{M}_d$ is irreducibly covariant w.r.t. an irrep $U : G \to \mathcal{M}_d$ of a finite group G if and only if it has a decomposition of the following form:*

$$T = l_{\mathrm{id}} P^{\mathrm{id}} + \sum_{\alpha\in\hat{G},\alpha\neq\mathrm{id}} l_\alpha P^\alpha, \tag{10}$$

with $l_{\mathrm{id}} = 1$, $l_\alpha \in \mathbb{C}$ and where $P^{\mathrm{id}}, P^\alpha : \mathcal{M}_d \to \mathcal{M}_d$ are projectors defined as

$$P^\alpha(\cdot) = \frac{\chi^\alpha(e)}{|G|} \sum_{g\in G} \chi^\alpha\left(g^{-1}\right) U_g \cdot U_g^\dagger, \tag{11}$$

with $\alpha \in \hat{G}$ and $e \in G$ the identity of the group. They have the following properties:

$$P^\alpha P^\beta = \delta_{\alpha\beta} P^\alpha, \qquad (P^\alpha)^* = P^\alpha \text{ and } \qquad \sum_{\alpha\in\hat{G}} P^\alpha = \mathrm{id}_d, \tag{12}$$

where $\mathrm{id}_d : \mathcal{M}_d \to \mathcal{M}_d$ is the identity map and the coefficients l_α are the eigenvalues of the quantum channel T.

That is, in the case of an irreducibly covariant channel we can also write down the projections onto different eigenspaces and diagonalize the channel.

One of the most important concepts in this paper is that of the twirl of a channel.

Definition 6 (Twirl). Let $T : \mathcal{M}_d \to \mathcal{M}_d$ be a quantum channel, G a finite group with Haar measure μ and $U : U \to \mathcal{M}_d$ a unitary representation of G. We define the twirl of T w.r.t. G, denoted by $\mathcal{T}(T) : \mathcal{M}_d \to \mathcal{M}_d$, as

$$\mathcal{T}(T)(\cdot) = \int_G \mathcal{U}_g^* \circ T \circ \mathcal{U}_g(\cdot) \mathrm{d}\mu. \tag{13}$$

Strictly speaking the twirled channel, of course, depends on the particular group and unitary representation at hand. However, we will omit this in the notation, as the group in question should always be clear from context. It is then easy to show that $\mathcal{T}(T)$ is a quantum channel that is covariant w.r.t. this representation.

2.2. Random walks on groups

We will need some basic tools from the field of random walks on groups to motivate and explain our protocol to perform RB with generators or with approximate samples. Therefore, we review these basic concepts here and refer to e.g. [28, chapter 2.6] for more details and proofs. Given a finite group G, we denote the set of probability measures on G by $\mathcal{P}(G)$. If X, Y are two independent random variables on G with distributions $\mu, \nu \in \mathcal{P}(G)$, respectively, we denote their joint distribution on $G \times G$ by $\mu \otimes \nu$. Analogously, we will denote the joint distribution of $Y_1, \ldots, Y_n$ i.i.d. variables with distribution ν by $\nu^{\otimes n}$ and the m-fold Cartesian product of G with itself by G^m. The random walk on G with increment distribution ν is defined as follows: it is a Markov chain with state space G. Given that the current state X_n of the chain is g, the next state X_{n+1} is given by multiplying the current state on the left by a random element of G selected according to ν. That is, we have

$$P(X_{n+1} = g_2 | X_n = g_1) = \nu\left(g_2 g_1^{-1}\right). \tag{14}$$

Another way of tracking the transition probabilities for these chains is through the transition matrix of the chain, π. For $g_1, g_2 \in G$, this matrix is defined as

$$\pi(g_1, g_2) = \nu\left(g_2 g_1^{-1}\right). \tag{15}$$

If X_0 is distributed according to $\mu \in \mathcal{P}(G)$, we have that the distribution of X_n is given by $\pi^n \mu$, where we just expressed μ as a vector in $\mathbb{R}^{|G|}$. We recall the following fundamental result about random walks on groups:

Theorem 7. *Let G be a finite group and A be a set of generators of G that is closed under inversion. Moreover, let ν be the uniform distribution on A and $X_1, X_2, \ldots$ be a random walk with increment distribution ν. Then the distribution of X_n converges to the Haar distribution on G as $n \to \infty$.*

Proof. We refer to e.g. [28, section 2.6.1] for a proof and more details on this. □

Given a generating subset A of G that is closed under inverses and ν the uniform distribution on A, we will refer to the random walk with increment ν as the random walk generated by A. This result provides us with an easy way of obtaining samples which are approximately Haar distributed if we have a set of generators by simulating this random walk for long enough. The speed of this convergence is usually quantified in the total variation distance. Given two probability measure μ, ν on G, we define their total variation distance to be given by:

$$\|\mu - \nu\|_1 := \frac{1}{2}\sum_{g\in G} |\mu(g) - \nu(g)|. \tag{16}$$

We then define the mixing time of the random walk as follows:

Definition 8 (Mixing time of random walk). Let G be a finite group and A a set of generators closed under inverses and μ be the Haar measure on the group. For $\epsilon > 0$, the mixing time of the chain generated by A, $t_1(\epsilon)$, is defined as

$$t_1(\epsilon) := \inf\{n \in \mathbb{N} | \forall \nu \in \mathcal{P}(G) : \|\pi^n \nu - \mu\|_1 \leqslant \epsilon\}. \tag{17}$$

We set t_{mix} to be given by $t_1(4^{-1})$, as this is a standard choice in literature [28, section 4.5]. One can then show that $t_1(\epsilon) \leqslant \lceil \log_2 \left(\epsilon^{-1}\right) \rceil t_{\text{mix}}$ (see [28, section 4.5] for a proof). There is a huge literature devoted to determining the mixing time of random walks on groups and we refer to [29] and references therein for more details. For our purposes it will be enough to note that in most cases we have that $t_1(\epsilon)$ scales logarithmically with ϵ^{-1} and $|G|$. Another distance measure which is quite useful in the study of convergence of random variables is the relative entropy D. For two probabilities measures μ, ν on $\{1, \ldots, d\}$ we define their relative entropy to be

$$D(\mu||\nu) := \begin{cases} \sum_{i=1}^{d} \mu(i) \log\left(\frac{\mu(i)}{\nu(i)}\right), & \text{if} \mu(i) = 0 \text{forall} i \text{s.t.} \nu(i) = 0, \\ +\infty, & \text{else.} \end{cases} \tag{18}$$

One of its main properties is that for $\mu, \nu \in \mathcal{P}(G)$ we have [30]

$$D(\mu^{\otimes n}||\nu^{\otimes n}) = nD(\mu||\nu). \tag{19}$$

3. Fidelities

Given a quantum channel $T : \mathcal{M}_d \to \mathcal{M}_d$ and a unitary channel $\mathcal{U} : \mathcal{M}_d \to \mathcal{M}_d$, the average fidelity between them is defined as

$$F(T, \mathcal{U}) = \int \text{Tr}\left(T(|\psi\rangle\langle\psi|)\mathcal{U}(|\psi\rangle\langle\psi|)\right) \text{d}\psi, \tag{20}$$

where we are integrating over the Haar measure on quantum states. In case $\mathcal{U}$ is just the identity, we refer to this quantity as being the average fidelity of the channel and denote it by $F(T)$. As shown in [31], the average fidelity of a channel is a simple function of its entanglement fidelity, given by

$$F_e(T) = \text{Tr}\left(T \otimes \text{id}\left(|\Omega\rangle\langle\Omega|\right)|\Omega\rangle\langle\Omega|\right), \tag{21}$$

with $|\Omega\rangle\langle\Omega|$ the maximally entangled state. One can then show that

$$F(T) = \frac{dF_e(T) + 1}{d + 1}. \tag{22}$$

Thus, we focus on estimating the entanglement fidelity instead of estimating the average fidelity. This can be seen to be just a function of the trace of the channel and the dimension, as we now show.

Lemma 9. *Let $T : \mathcal{M}_d \to \mathcal{M}_d$ be a quantum channel. Then $F_e(T) = d^{-2}\text{Tr}(T)$. Here we mean the trace of T as a linear operator between the vector spaces $\mathcal{M}_d$.*

Proof. The entanglement fidelity is

$$\begin{aligned} F_e(T) &= \mathrm{Tr}\left(T \otimes \mathrm{id}\left(|\Omega\rangle\langle\Omega|\right)|\Omega\rangle\langle\Omega|\right) \\ &= \frac{1}{d^2}\sum_{i,j,k,l=1}^{d} \mathrm{Tr}\left(\left[T\left(|i\rangle\langle j|\right)\otimes|i\rangle\langle j|\right]|l\rangle\langle k|\otimes|l\rangle\langle k|\right) \\ &= \frac{1}{d^2}\sum_{i,j=1}^{d} \mathrm{Tr}\left(T\left(|i\rangle\langle j|\right)\left(|i\rangle\langle j|\right)^{\dagger}\right). \end{aligned}$$

Note that $\{|i\rangle\langle j|\}_{i,j=1}^{d}$ is an orthonormal basis of $\mathcal{M}_d$ and $\mathrm{Tr}\left(T\left(|i\rangle\langle j|\right)\left(|i\rangle\langle j|\right)^{\dagger}\right)$ corresponds to the Hilbert–Schmidt scalar product between $T\left(|i\rangle\langle j|\right)$ and $|i\rangle\langle j|$. Therefore, we have that

$$\sum_{i,j=1}^{d} \mathrm{Tr}\left(T\left(|i\rangle\langle j|\right)\left(|i\rangle\langle j|\right)^{\dagger}\right) = \mathrm{Tr}\left(T\right),$$

where again $\mathrm{Tr}\left(T\right)$ is meant as the trace of T as a linear operator. □

That is, if we know the eigenvalues or the diagonal elements of T w.r.t. some basis, we may determine its entanglement and average fidelity. The RB protocol explores the fact that twirling a channel does not change its trace and that the trace of covariant channels has a much simpler structure, as made clear in the next corollary.

Corollary 10. *Let $T : \mathcal{M}_d \to \mathcal{M}_d$ be a quantum channel that is covariant w.r.t. a unitary representation $U : G \to \mathbb{C}^d$ of a finite group G and let $\oplus_{\alpha\in\hat{G}}\left(\mathbb{C}^{d_\alpha}\otimes\mathbb{C}^{m_\alpha}\right)$ be the decomposition of $\mathbb{C}^d\otimes\mathbb{C}^d$ into irreps α of G with multiplicity m_α for the unitary representation $U\otimes\bar{U}$. Choose a basis s.t.*

$$T = \oplus_{\alpha\in\hat{G}}\mathbb{I}_{d_\alpha}\otimes B_{m_\alpha} \tag{23}$$

with $B_\alpha \in \mathcal{M}_{m_\alpha}$. Then

$$F_e(T) = d^{-2}\sum_{\alpha\in\hat{G}} d_\alpha \mathrm{Tr}\left(B_\alpha\right). \tag{24}$$

Proof. The claim follows immediately after we combine theorem 4 and lemma 9. □

This shows that the spectrum of quantum channels that are covariant w.r.t. a unitary representation of a finite group has much more structure and is simpler than that of general quantum channels. In particular, if the unitary representation $U\otimes\bar{U}$ is such that $\sum_\alpha m_\alpha \ll d^2$, then we know that the spectrum of the quantum channel is highly degenerate and we only need to know a few points of it to estimate the trace. We will explore this fact later in the implementation of the RB protocol.

We will now show in lemma 11 that the probability of measurement outcomes has a very simple form for covariant channels and their powers.

Lemma 11. *Let $T : \mathcal{M}_d \to \mathcal{M}_d$ be a quantum channel that is covariant w.r.t. a unitary representation $U : G \to \mathcal{M}_d$ of a finite group G and let $\oplus_{\alpha\in\hat{G}}\left(\mathbb{C}^{d_\alpha}\otimes\mathbb{C}^{m_\alpha}\right)$ be the decomposition of $\mathbb{C}^d\otimes\mathbb{C}^d$ into irreps α of G with multiplicity m_α for the unitary representation $U\otimes\bar{U}$. Moreover, let $\rho\in\mathcal{D}_d$, $E\in\mathcal{M}_d$ be a POVM element and $m \geqslant \max m_\alpha$. Then there exist $\lambda_1,\ldots,\lambda_k \in \overline{B_1(0)}$, the unit ball in the complex plane, and $a_0, a_1, \ldots, a_k \in \mathbb{C}$ s.t.*

$$\mathrm{Tr}\left(T^m(\rho)E\right) = a_0 + \sum_{i=1}^{k} a_k \lambda_i^m. \tag{25}$$

Moreover,

$$k \leqslant \sum_{\alpha\in\hat{G}} m_\alpha - 1 \tag{26}$$

corresponds to the number of distinct eigenvalues of T and λ_i are its eigenvalues.

Proof. As T is a linear map from $\mathcal{M}_d$ to $\mathcal{M}_d$ it has a Jordan decomposition [32]. That is, there exists an invertible linear operator $X : \mathcal{M}_d \to \mathcal{M}_d$ such that

$$X^{-1} \circ T \circ X = D + N, \quad [D, N] = 0.$$

Here $D : \mathcal{M}_d \to \mathcal{M}_d$ is diagonal in the standard basis $\{|i\rangle\langle j|\}_{i,j=1}^d$ of $\mathcal{M}_d$ with diagonal entries given by the eigenvalues of T and $N : \mathcal{M}_d \to \mathcal{M}_d$ nilpotent. As we have that T is covariant, it follows from the decomposition in theorem 4 that the eigenvalues can be at most $\max m_\alpha = m_0$−fold degenerate and $N^{m_0} = 0$. Thus, it follows that T^m is diagonalizable, as $m \geqslant \max m_\alpha$. We then have

$$X^{-1} \circ T^m \circ X = D^m.$$

We can then rewrite the scalar product

$$\mathrm{Tr}\left(T^m(\rho)E\right) = \mathrm{Tr}\left(X \circ D^m \circ X^{-1}(\rho)E\right) = \mathrm{Tr}\left(D^m(X^{-1}(\rho))X^*(E)\right).$$

Let $b_{i,j}$ and $c_{i,j}$ be the matrix coefficient of $X^*(E)$ and $X^{-1}(\rho)$, respectively, in the standard basis. That is

$$X^\dagger(E) = \sum_{i,j=1}^{d} b_{i,j} |i\rangle\langle j|, \qquad X^{-1}(\rho) = \sum_{i,j=1}^{d} c_{i,j} |i\rangle\langle j|.$$

Exploring the fact that D is diagonal in this basis we obtain

$$\mathrm{Tr}\left(T^m(\rho)E\right) = \sum_{i,j=1}^{d} b_{i,j}c_{i,j}d_{i,j}^m,$$

where $d_{i,j}$ are just the eigenvalues of T, including multiplicities. To arrive at the curve in (25), we group together all terms corresponding to the same eigenvalue λ_i. Moreover, note that quantum channels always have 1 in their spectrum, which gives the a_0 term that does not depend on m. The fact that $\lambda_i \in \overline{B_1(0)}$ follows from the fact that they are given by the eigenvalues of the channel and these are always contained in the unit circle of the complex plane [33]. □

Finally, we show that twirling does not change the entanglement fidelity and thus does not change the average fidelity, as observed in [31] and elsewhere in the literature. Thus, when we want to estimate the average fidelity of a channel $T : \mathcal{M}_d \to \mathcal{M}_d$ we may instead work with the twirled channel $\mathcal{T}(T)$ and explore its rich structure.

Theorem 12. *Let $T : \mathcal{M}_d \to \mathcal{M}_d$ be a quantum channel, G be a finite group and $U : G \to \mathcal{M}_d$ be a unitary representation. Then*

$$F_e(T) = F_e(\mathcal{T}(T)). \tag{27}$$

Proof. We present a slightly different proof of this fact here. Note that $\mathcal{U}_g^* \circ T \circ \mathcal{U}_g$ is just a similarity transformation of T and thus $\mathrm{Tr}\left(\mathcal{U}_g^* \circ T \circ \mathcal{U}_g\right) = \mathrm{Tr}\,(T)$, where again we mean the trace of these channels as linear operators. Thus, integrating over all $\mathcal{U}_g$ does not change the entanglement fidelity, as $F_e(T) = d^{-2}\mathrm{Tr}\,(T)$. □

4. Randomized benchmarking protocol

The RB protocol, as discussed in [3–6, 8, 34–39] is a protocol to estimate the average fidelity of the implementation of gates coming from some group G. Its usual setting is the Clifford group, but we discuss it for general groups here. Other papers have investigated the protocol for gates beyond Cliffords, such as [18, 19, 21]. But all of these have restricted their analysis to some other specific group. As we will see later, we can analyze the protocol for arbitrary groups by just investigating properties of the given unitary representation. We mostly follow the notation of [38]. We assume that the error quantum channel is gate and time independent. That is, whenever we want to implement a certain gate $\mathcal{U}_g$, where $\mathcal{U}_g(\cdot) = U_g \cdot U_g^\dagger$ with $U_g \in U(d)$, we actually implement $\mathcal{U}_g \circ T$ for some quantum channel $T : \mathcal{M}_d \to \mathcal{M}_d$. We assume that we are able to multiply and invert elements of G and draw samples from the Haar measure on G efficiently to implement this protocol, but will later relax this sampling condition. The protocol is as follows:

Step 1 Fix a positive integer $m \in \mathbb{N}$ that varies with every loop.

Step 2 Generate a sequence of $m+1$ quantum gates. The first m quantum gates $\mathcal{U}_{g_1}, \ldots, \mathcal{U}_{g_m}$ are independent and Haar distributed. The final quantum gate, $\mathcal{U}_{g_{m+1}}$ is chosen such that in the absence of errors the net sequence is just the identity operation,

$$\mathcal{U}_{g_{m+1}} \circ \mathcal{U}_{g_m} \circ \ldots \circ \mathcal{U}_{g_2} \circ \mathcal{U}_{g_1} = \mathrm{id}, \tag{28}$$

where $\circ$ represents composition. Thus, the whole quantum gate sequence is

$$\mathcal{S}_m = \bigcirc_{j=1}^{m+1} \mathcal{U}_{g_j} \circ T, \tag{29}$$

where T is the associated error quantum channel.

Step 3 For every sequence, measure the sequence fidelity

$$\mathrm{Tr}\left(\mathcal{S}_m(\rho)E\right), \tag{30}$$

where ρ is the initial quantum state, including preparation errors, and E is an effect operator of some POVM including measurement errors.

Step 4 Repeat steps 2–3 and average over M random realizations of the sequence of length m to find the averaged sequence fidelity given by

$$\bar{F}(m, E, \rho) = \frac{1}{M} \sum_m \mathrm{Tr}\left(\mathcal{S}_m(\rho)E\right). \tag{31}$$

Step 5 Repeat steps 1–4 for different values of m and obtain an estimate of the expected value of the sequence fidelity

$$F(m,E,\rho) = \mathrm{Tr}\left(\mathbb{E}(\mathcal{S}_m)(\rho)E\right). \tag{32}$$

4.1. Analysis of the protocol

We will now show how we can estimate the average fidelity from the data produced by the protocol, that is, an estimate on the curve $F(m,E,\rho) = \mathrm{Tr}\left(\mathbb{E}(\mathcal{S}_m)(\rho)E\right)$.

Theorem 13. *Let $T : \mathcal{M}_d \to \mathcal{M}_d$ be a quantum channel and G a group with a unitary representation $U : G \to \mathcal{M}_d$. If we perform the RB protocol for G we have*

$$\mathbb{E}(\mathcal{S}_m) = \mathcal{T}(T)^m. \tag{33}$$

Proof. Although the proof is identical to the case in which G is given by the Clifford group, we will cover it here for completeness. Given some sequence $\{\mathcal{U}_{g_1},\ldots,\mathcal{U}_{g_{m+1}}\}$ of unitary gates from G, define the unitary operators

$$\mathcal{D}_i = \bigcirc_{j=1}^{i}\mathcal{U}_{g_i}. \tag{34}$$

Note that we have

$$\begin{aligned}
\mathcal{S}_m =& \mathcal{U}_{g_{m+1}} \circ T \circ \mathcal{U}_{g_m} \circ T \circ \ldots \circ \mathcal{U}_{g_2} \circ T \circ \mathcal{U}_{g_1} \\
=& \overbrace{\mathcal{U}_{g_{m+1}} \circ (\mathcal{U}_{g_m} \circ \ldots \circ \mathcal{U}_{g_1})}^{=\mathbb{I}} \circ \overbrace{\mathcal{U}^*_{g_1} \circ \ldots \circ \mathcal{U}^*_{g_m})}^{=\mathcal{D}_m^*} \circ T \circ \mathcal{U}_{g_m} \circ T \circ \\
& \ldots T \circ \underbrace{\mathcal{U}_{g_3} \circ (\mathcal{U}_{g_2} \circ \mathcal{U}_{g_1})}_{=\mathcal{D}_3} \circ \underbrace{\mathcal{U}^*_{g_1} \circ \mathcal{U}^*_{g_2})}_{=\mathcal{D}_2^*} \circ T \circ \underbrace{\mathcal{U}_{g_2} \circ (\mathcal{U}_{g_1})}_{=\mathcal{D}_2} \circ \underbrace{\mathcal{U}^*_{g_1}}_{=\mathcal{D}_1^*} \circ T \circ \underbrace{\mathcal{U}_{g_1}}_{=\mathcal{D}_1} \\
=& \mathcal{D}_m^* \circ T \circ \mathcal{D}_m \circ \ldots \circ \mathcal{D}_2^* \circ T \circ \mathcal{D}_2 \circ \mathcal{D}_1^* \circ T \circ \mathcal{D}_1 \\
=& \bigcirc_{j=1}^{m} \left(\mathcal{D}_j^* \circ T \circ \mathcal{D}_j\right).
\end{aligned} \tag{35}$$

Here we have absorbed the first channel T as SPAM error. As we have that each of the $\mathcal{U}_{g_i}$ is independent and Haar-distributed, it follows that the $\mathcal{D}_i$ are independent and Haar distributed as well. It then follows from (35) that

$$\mathbb{E}\left(\mathcal{S}_m\right) = \mathbb{E}\left(\bigcirc_{j=1}^{m}\left(\mathcal{D}_j^* \circ T \circ \mathcal{D}_j\right)\right) = \bigcirc_{j=1}^{m}\mathbb{E}\left(\mathcal{D}_j^* \circ T \circ \mathcal{D}_j\right) = \mathcal{T}(T)^m.$$

□

We can then use our structural results on covariant quantum channels to obtain a more explicit form for the curve $F(m,E,\rho)$.

Corollary 14. *Suppose we perform RB for a unitary representation $U : G \to \mathcal{M}_d$ of a finite group G s.t. $U \otimes \bar{U} = \oplus_{\alpha\in\hat{G}}\left(\mathbb{C}^{d_\alpha} \otimes \mathbb{C}^{m_\alpha}\right)$ and a channel T. Then there exist $\lambda_1,\ldots,\lambda_k \in \overline{B_1(0)}$ and $a_0, a_1, \ldots, a_k \in \mathbb{C}$ s.t.*

$$F(m,E,\rho) = a_0 + \sum_{i=1}^{k} a_k \lambda_i^m. \tag{36}$$

for $m \geqslant \max m_\alpha$. Moreover, $k \leqslant \sum_\alpha m_\alpha$ corresponds to the number of distinct eigenvalues of $\mathcal{T}(T)$ and λ_i are its eigenvalues.

Proof. The claim follows immediately from theorem 13 and lemma 11. □

That is, by fitting the curve to experimental data we may obtain estimates on the λ_i and thus on the spectrum of $\mathcal{T}(T)$. If we know the multiplicity of each eigenvalue, then we can estimate the trace as well and thus the average fidelity. However, in the case in which we have more than one parameter to estimate, it is not clear which eigenvalue corresponds to which irrep and we therefore cannot simply apply the formula in corollary 10. Suppose we are given an estimate $\{\hat{\lambda}_1, \ldots, \hat{\lambda}_k\}$ of the parameters sorted in decreasing order and let $d_\alpha^\uparrow$ be the dimensions of the irreps sorted in ascending and $d_\alpha^\downarrow$ in descending order. We define the minimal fidelity, $F_{\min}$, to be given by

$$F_{\min} = \frac{1}{d^2} \sum d_\alpha^\downarrow \hat{\lambda}_i \tag{37}$$

and the maximum fidelity, $F_{\max}$, to be given by

$$F_{\max} = \frac{1}{d^2} \sum d_\alpha^\uparrow \hat{\lambda}_i. \tag{38}$$

That is, we look at the pairings of d_α and $\hat{\lambda}_i$ that produces the largest and the smallest estimate for the fidelity. These then give the most pessimistic and most optimistic estimate, respectively. The fact that we cannot associate a λ_i to each irrep causes some problems in this approach from the numerical point of view and we comment on them in appendix A. We also note that since the first version of this work, a modified version of the protocol we describe here was given in [40]. Their protocol provides a way of isolating the individual parameters in the case of irreducibly covariant channels.

5. Approximate twirls

In the description of our RB protocol, we assume that we are able to obtain samples from the Haar measure of the group G. It is not possible or efficient to obtain samples of the Haar measure for most groups, but a lot of research has been done on how to obtain approximate samples efficiently using Markov chain Monte Carlo methods, as discussed in section 2.2. Here we discuss how to use samples which are approximately Haar distributed for RB. Note that these results may also be interpreted as a stability result w.r.t. not sampling exactly from the Haar measure of G. We will assume we are able to pick the $\mathcal{U}_{g_k}$ independently and that they are distributed according to a measure ν_k s.t.

$$\|\nu_k - \mu\|_1 \leqslant \epsilon_k, \tag{39}$$

for $\epsilon_k \geqslant 0$. Our goal is to show that under these assumptions we may still implement the RB protocol discussed before and obtain measurement statistics that are close to the ones obtained using Haar samples.

Motivated by this, we define the $\tilde{\nu}$-twirl of a channel.

Definition 15 ($\tilde{\nu}$-twirl to the power m). Let $\tilde{\nu}$ be a probability measure on G^m, $T : \mathcal{M}_d \to \mathcal{M}_d$ a quantum channel and $U : G \to \mathcal{M}_d$ a $d-$dimensional unitary representation of G. We define the $\tilde{\nu}$-twirl to the power m to be given by

$$\mathcal{T}_{\tilde{\nu},m}(T) = \sum_{i_1,\ldots,i_m=1}^{|G|} \tilde{\nu}\left(g_{i_1}, \ldots, g_{i_m}\right) \bigcirc_{k=1}^m \mathcal{U}_{g_{i_k}} \circ T \circ \mathcal{U}_{g_{i_k}}^*. \tag{40}$$

This definition boils down to the regular twirl for $\tilde{\nu} = \mu^{\otimes m}$, μ the Haar measure on G. We will now show that by sampling $\mathcal{U}_{g_k}$ close to Haar we have that the $\tilde{\nu}$-twirl of a channel is also close to the usual twirl.

Lemma 16 (Approximate twirl). *Let $T : \mathcal{M}_d \to \mathcal{M}_d$ be a quantum channel, G a finite group with a $d-$dimensional unitary representation $U : G \to \mathcal{M}_d$ and $\tilde{\nu}$ a probability measure on G^m. Let $\mathcal{T}_{\tilde{\nu},m}(T)$ be the $\tilde{\nu}$-twirl to the power m and $\mathcal{T}(T)$ be the twirl w.r.t. the Haar measure on G given by μ. Moreover, let $\|\cdot\|$ be a norm s.t. $\|T\| \leqslant 1$ for all quantum channels. Then*

$$\|\mathcal{T}_{\tilde{\nu},m}(T) - \mathcal{T}(T)^m\| \leqslant 2\|\tilde{\nu} - \mu^{\otimes m}\|_1. \tag{41}$$

Proof. Observe that we may write

$$\|\mathcal{T}_{\tilde{\nu},m}(T) - \mathcal{T}(T)^m\| = \left\| \sum_{i_1,\ldots,i_m=1}^{|G|} \left(\tilde{\nu}(g_{i_1},\ldots,g_{i_m}) - \frac{1}{|G|^m} \right) \bigcirc_{k=1}^m \mathcal{U}_{g_{i_k}} \circ T \circ \mathcal{U}_{g_{i_k}}^* \right\|.$$

The claim then follows from (16), as

$$\sum_{i_1,\ldots,i_m=1}^{|G|} \left| \tilde{\nu}(g_{i_1},\ldots,g_{i_m}) - \frac{1}{|G|^m} \right| = 2\|\tilde{\nu} - \mu^{\otimes m}\|_1,$$

the triangle inequality and the fact that $\| \bigcirc_{k=1}^m \mathcal{U}_{g_{i_k}} \circ T \circ \mathcal{U}_{g_{i_k}}^* \| \leqslant 1$. □

Thus, in order to bound $\|\mathcal{T}_{\tilde{\nu},m}(T) - \mathcal{T}(T)^m\|$ in any norm in which quantum channels are contractions, it suffices to bound $\|\tilde{\nu} - \mu^{\otimes m}\|_1$. Examples of such norms are the $1 \to 1$ norm and the diamond norm [41, theorem 2.1]. We remark that other notions of approximate twirling were considered in the literature [39, 42], but these works were mostly concerned with the case of the unitary group and not arbitrary finite groups. Although it would be straightforward to adapt their definitions to arbitrary finite groups, it is not clear at first sight that their notions of approximate twirls behave well when taking powers of channels that have been twirled approximately. This is key for RB. Given random unitaries $\{U_i\}_{i=1}^m$ from G, let $\mathcal{D}_k = \bigcirc_{i=1}^k \mathcal{U}_i$, as before.

Theorem 17. *Let μ be the Haar measure on G and $\nu_1, \ldots, \nu_m$ probability measures on G s.t.*

$$\|\mu - \nu_k\|_1 \leqslant \epsilon_k, \tag{42}$$

for all $1 \leqslant k \leqslant m$ and $\epsilon_k \geqslant 0$. Denote by $\tilde{\nu}$ the distribution of $(\mathcal{D}_1, \ldots, \mathcal{D}_m)$ if we pick the $\mathcal{U}_k$ independently from ν_k . Then

$$\|\mathcal{T}_{\tilde{\nu},m}(T) - \mathcal{T}(T)^m\|_{1\to 1} \leqslant 4\sqrt{\frac{\log(|G|)}{1 - |G|^{-1}} \sum_{k=1}^m \epsilon_k}. \tag{43}$$

Proof. We refer to appendix C for a proof and only sketch the main steps here. We start by applying lemma 16 to reduce the problem of estimating this norm to estimating the total variation distance between $\tilde{\nu}$ and μ. We then show that the total variation distance between $\tilde{\nu}$ and μ and $\otimes_{k=1}^m \mu_k$ coincide. We bound this total variation distance by the relative entropy using

Pinsker's inequality, explore tensorization properties of the relative entropy, and then use a reverse Pinsker inequality. We then obtain the final bound from (43). □

Note that the same result holds for any norm that contracts under quantum channels, such as the diamond norm.

Corollary 18. *Let μ be the Haar measure on G and $\nu_1, \ldots, \nu_m$ probability measures on G s.t.*

$$\|\mu - \nu_k\|_1 \leqslant \epsilon_k, \tag{44}$$

for all $1 \leqslant k \leqslant m$ and $\epsilon_k \geqslant 0$. Denote by $\tilde{\nu}$ the distribution of $(\mathcal{D}_1, \ldots, \mathcal{D}_m)$ if we pick the $\mathcal{U}_k$ independently from ν_k. Then

$$|\mathrm{Tr}\left(\mathcal{T}_{\tilde{\nu},m}(T)(\rho)E\right) - F(m,E,\rho)| \leqslant 4\sqrt{\frac{\log(|G|)}{1-|G|^{-1}}\sum_{k=1}^{m}\epsilon_k}. \tag{45}$$

Proof. It follows from Hölder's inequality that

$$\begin{aligned}|\mathrm{Tr}\left(\mathcal{T}_{\tilde{\nu},m}(T)(\rho)E\right) - F(m,E,\rho)| &= |\mathrm{Tr}\left(E\left(\mathcal{T}_{\tilde{\nu},m}(T)(\rho) - \mathcal{T}(T)^m(\rho)\right)\right)| \\ &\leqslant \|E\|_\infty \|\mathcal{T}_{\tilde{\nu},m}(T) - \mathcal{T}(T)^m\|_{1\to 1},\end{aligned}$$

where we have used the submultiplicativity of the $1 \to 1$-norm. As E is the element of a POVM, we have $\|E\|_\infty \leqslant 1$ and the claim then follows from theorem 17. □

This shows that we may use approximate twirls instead of exact ones and obtain expectation values that are close to the perfect twirl. Given that we want to assure that the statistics we obtain for some $m \in \mathbb{N}$ are $\delta > 0$ close to our target distribution, we would have to sample the U_{g_k} such that

$$\|\mu - \nu_k\|_1 \leqslant \frac{\delta^2(1-|G|^{-1})}{16\log(|G|)m}, \tag{46}$$

as can be seen by plugging in this bound in the result of corollary 18. If we use a random walk on a group to sample from the Haar distribution we have to run each chain for $t_1\left(\frac{\delta^2(1-|G|^{-1})}{16\log(|G|)m}\right)$ steps, which gives a total runtime of $\mathrm{O}\left(t_{\mathrm{mix}}\log\left(\frac{16\log(|G|)m}{\delta^2(1-|G|^{-1})}\right)\right)$. For a fixed δ, this will be efficient if the chain mixes rapidly, that is, t_{mix} is small, and we choose m to be at most of the order of the dimension.

6. Randomized benchmarking with generators

One of the downsides of the usual RB protocol [3–6, 34–39] is that we assume that we may implement any gate of the group. Usually, gates have to be broken down into generators, as discussed in [43, section 1.2.3 and chapter 8]. Therefore, it would be desirable both from the point of view of justifying the noise model and the implementation level of the protocol to mostly need to implement gates from a set of generators. We describe here a protocol to perform RB by just implementing gates from a set of generators closed under inversion and one arbitrary gate. We also make the additional assumption that the quantum channel that describes the noise is already approximately covariant in a sense we will make precise soon.

This protocol is inspired by results of the last section that suggest a way of performing RB by just implementing gates coming from a set A that generates the group G and is closed under inversion and one additional arbitrary gate from G at each round of the protocol. From the basic results of random walks discussed in section 2.2, we know that if we pick gates $U_{g_1}, U_{g_2}, \ldots$ uniformly at random from A, it follows that $U_{g_b} U_{g_{b-1}} \ldots U_{g_1}$ will be approximately distributed like the Haar measure on G for $b \simeq t_{\text{mix}}$. However, one should note that in this setting the $\mathcal{D}_i$, defined in (34), will not be independent of each other. To see this, note that given $\mathcal{D}_i = \mathcal{U}_g$, we know that the distribution of the D_{i+1} is restricted to elements $h \in G$ of the form $h = ag$ with $a \in A$, which clearly show that they are not independent in general. However, if we look at $\mathcal{D}_{i+l}$ for $l \sim t_{\text{mix}}$, then their joint distribution will be close to Haar. That is, looking at $\mathcal{D}_i$ and $\mathcal{D}_j$ which are far enough apart from each other, we may again assume that they are both almost Haar distributed and if we look at each $\mathcal{D}_i$ individually we may assume that they are almost Haar distributed. One way to explore this observation for RB protocols only having to implement the generators is to look at the following class of quantum channels:

Definition 19 (δ-covariant quantum channel). A quantum channel $T : \mathcal{M}_d \to \mathcal{M}_d$ is called δ-covariant w.r.t. a unitary representation $U : G \to \mathcal{M}_d$ of a group G, if there exist quantum channels $T_c, T_n : \mathcal{M}_d \to \mathcal{M}_d$ such that

$$T = (1 - \delta) T_c + \delta T_n, \tag{47}$$

and T_c is covariant w.r.t. U.

That is, T is almost covariant w.r.t. the group. Similar notions of approximate covariance were also introduced in [44]. Their notion of an approximate covariant channel is arguably more natural than ours, as they quantify how close a channel is to being covariant using the minimal distance to a covariant channel in the diamond norm. Unfortunately, we also need information on how close the powers of the channel are to being covariant and it is not clear how to derive such bounds using their definition but it is straightforward using ours. Another issue related to our definition is the fact that it does not cover quantum channels that are unitary, unless they are the identity. That is because unitary channel are extremal points of the set of quantum channels [26] and thus cannot be written in any nontrivial convex combination with another quantum channel. Thus, it remains an open problem how to generalize these methods to unitary noise, as preliminary numerical evidence suggests that the protocol also gives good estimates in this case. The standard example of quantum channels that satisfy our definition are quantum channels that are close to the identity channel, i.e. we have δ small and T_c the identity channel.

We will need to fix some notation before we describe the protocol. For a given sequence of unitaries $s_i = (U_{g_1}, U_{g_2}, \ldots)$ we let $\mathcal{S}_{s_i,c,d} = \bigcirc_{j=c}^{d} \mathcal{U}_{g_j} \circ T$ for $c, d \in \mathbb{N}$ and the gates chosen according to the sequence.

Thus, if we apply random generators b times as an initialization procedure and only start fitting the curve after this initialization procedure we may also estimate the average fidelity.

This yields the following protocol.

Step 1 Fix a positive integer $m \in \mathbb{N}$ that varies with every loop and another integer $b \in \mathbb{N}$.

Step 2 Generate a sequence of $b + m + 1$ quantum gates, s_i. The first $b + m$ quantum gates $\mathcal{U}_{g_1}, \ldots, \mathcal{U}_{g_{b+m}}$ are chosen independently and uniformly at random from A. The final quantum gate, $\mathcal{U}_{g_{b+m+1}}$ is chosen as

$$\mathcal{U}_{g_{b+m+1}} = (\mathcal{U}_{g_{b+m}} \circ \ldots \circ \mathcal{U}_{g_2} \circ \mathcal{U}_{g_1})^{-1}. \tag{48}$$

Step 3 For each sequence s_i, measure the sequence fidelity

$$\mathrm{Tr}\left(\mathcal{S}_{s_i,b+1,b+m+1}(\mathcal{S}_{s_i,1,b}(\rho))E\right), \tag{49}$$

where ρ is the initial quantum state and E is an effect operator of a POVM.

Step 4 Repeat steps 2–3 and average over M random realizations of the sequence of length m to find the averaged sequence fidelity

$$\bar{F}(m,E,\rho) = \frac{1}{M}\sum_{i=1}^{M}\mathrm{Tr}\left(\mathcal{S}_{s_i,b+1,b+m+1}(\mathcal{S}_{s_i,1,b}(\rho))E\right). \tag{50}$$

Step 5 Repeat steps 1–4 for different values of m to obtain an estimate of the expected value of the average survival probability

$$F(m,E,\rho) = \mathbb{E}\left(\mathrm{Tr}\left(\mathcal{S}_{s_i,b+1,b+m+1}(\mathcal{S}_{s_i,1,b}(\rho))E\right)\right). \tag{51}$$

We will now prove that this procedure gives rise to the same statistics as if we were using samples from the Haar distribution up to $\mathrm{O}(\delta^2)$.

Theorem 20. *Let T be $\delta-$covariant w.r.t. a unitary representation $U: G \to \mathcal{M}_d$ of a finite group G, A a subset of G that generates G and is closed under inversion and $\delta > 0$. Suppose we run the protocol above with $b = t_1(m^{-1}\epsilon)$ for some ε and $m \geqslant b$. Then*

$$\|\mathcal{T}(T)^m - \mathbb{E}(\mathcal{S}_{b,b+m+1})\|_{1\to 1} \leqslant \epsilon + \mathrm{O}\left(\delta^2 bm\right). \tag{52}$$

Proof. We refer to appendix D for a proof. □

Corollary 21. *Let $\mathcal{S}_{b,m+b+1}$ and b be as in theorem 20. Then for any POVM element E and state $\rho \in \mathcal{M}_d$:*

$$\left|\mathrm{Tr}\left(\mathbb{E}\left(S_{b,m+1}\right)(\rho)E\right) - F(m,E,\rho)\right| \leqslant \epsilon + \mathrm{O}\left(\delta^2 bm\right). \tag{53}$$

Proof. The proof is essentially the same as that of corollary 18. □

This shows that performing RB by only implementing the generators is feasible as long as we have a $\delta-$covariant channel with δ small and a rapidly mixing set of generators, that is, $\delta^2 bm \ll 1$. Recall that a Markov chain is said to be rapidly mixing if the mixing time scales polylogarithmically with the size of the state space [45]. In our case, the size of the state space is given by the size of the group we are benchmarking. Thus, for groups whose size scales polynomially with the dimension of the system or equivalently exponentially in the number of qubits, this translates to b scaling like $\mathcal{O}(n^k \log^k(n\epsilon^{-1}m))$, for n the number of qubits and k a natural number. This scaling renders the protocol reliable if δ is roughly smaller than the inverse of a polynomial on the number of qubits.

7. Numerics and examples

Here we show how to apply our methods to groups that might be of special interest and discuss some numerical examples. Many relevant questions for the practical application of our work are still left open and have two different flavors: the numerical and statistical side. From the numerical point of view, it is not clear at first how to fit the data gathered by a RB protocol to an exponential curve if we have several parameters. We refer to appendix A for a discussion

of these issues and some proposals of how to overcome them. From a statistical point of view, it is not clear how to derive confidence intervals for the parameters and how large we should choose the different parameters of the protocol, such as m and M. We refer to appendix B for a discussion of these issues and preliminary results in this direction.

7.1. Monomial unitary matrices

We consider how to apply our methods of generalized RB to some subgroups of the monomial unitary matrices $MU(d)$.

Definition 22. Let $\{|i\rangle\}_{i=1}^{d}$ be an orthonormal basis of $\mathbb{C}^d$. We define the group of monomial unitary matrices, $MU(d)$ to be given by $U \in U(d)$ of the form $U = DP$ with $D, P \in U(d)$ and D diagonal w.r.t. $\{|i\rangle\}_{i=1}^{d}$ and P a permutation matrix.

Subgroups of this group can be used to describe many-body states in a formalism that is broader than the stabilizer formalism of Paulis and have other applications to quantum computation (see [22]). As the group above is not finite and it is unreasonable to assume that we may implement diagonal gates with phases of an arbitrary precision, we focus on the following subgroups:

Definition 23. We define $MU(d, n)$ to be the subgroup of the monomial unitary matrices of dimension d whose nonzero entries consist only of nth roots of unity.

Another motivation to consider these subgroups is that they contain the T-gate [23],

$$T = |0\rangle\langle 0| + \mathrm{e}^{\mathrm{i}\frac{\pi}{4}}\,|1\rangle\langle 1| \tag{54}$$

in case $n \geqslant 8$. Thus these gates, together with Cliffords, constitute a universal set of quantum gates [23]. Also note that the group considered here contains the group considered in [20]. There they also consider the group generated by diagonal matrices containing nth roots of unity, CNOTs and Pauli X gates. Although the latter two are permutations, they do not generate the whole group of permutations and the groups do not coincide. We now show that we have to estimate two parameters for them.

Lemma 24 (Structure of channels covariant w.r.t. monomial unitaries). *Let $MU(d, n)$ be such that $n \geqslant 3$ and $T : \mathcal{M}_d \to \mathcal{M}_d$ a quantum channel. Then the following are equivalent:*

(i) $T(\rho) = UT\big(U^\dagger \rho U\big)U^\dagger \quad \forall U \in MU(d, n), \rho \in \mathcal{S}_d$.
(ii) There are $\alpha, \beta \in \mathbb{R}$ *so that*

$$T(\cdot) = \mathrm{Tr}\,(\cdot)\,\frac{\mathbb{I}}{d} + \alpha\,\left(\mathrm{id} - \sum_{i=1}^{d} |i\rangle\langle i|\,\langle i|\cdot|i\rangle\right) + \beta\,\left(\sum_{i=1}^{d} |i\rangle\langle i|\,\langle i|\cdot|i\rangle - \mathrm{Tr}\,(\cdot)\,\frac{\mathbb{I}}{d}\right). \tag{55}$$

Moreover, the terms in the r.h.s. of (55) are projections of rank 1, $d^2 - d$ and $d - 1$, respectively.

Proof. We refer to appendix E for a proof. □

This result shows that we only need to estimate two parameters when performing RB with these subgroups. They are therefore a natural candidate to apply our methods to and we investigate this possibility further. We begin by analyzing the complexity of multiplying and

inverting elements of $MU(d,n)$. We show this more generally for $MU(d)$, as it clearly gives an upper bound for its subgroups as well. We may multiply and invert elements of $MU(d)$ in time $\mathrm{O}(d)$. To multiply elements in $MU(d)$ we need to multiply two permutations of d elements, which can be done in time $\mathrm{O}(d)$, multiply a vector $u \in \mathbb{C}^d$ with a permutation matrix, which can be done in time $\mathrm{O}(d)$, and multiply d elements of $U(1)$ with each other, which again can be done in time $\mathrm{O}(d)$. This shows that multiplying elements of this group takes $\mathrm{O}(d)$ operations. To invert an element of $MU(d)$ we need to invert a permutation, which again takes $\mathrm{O}(d)$, invert d elements of $U(1)$ and apply a permutation to the resulting vector. This also takes $\mathrm{O}(d)$ operations. Moreover, one can generate a random permutation and an element of $U(1)^d$ in time $\mathrm{O}(d)$, giving $\mathrm{O}(Mmd)$ complexity for the classical part of the RB procedure. Although this scaling is not efficient in the number of qubits as in the case of Clifford gates [3], the fact that it is linear in the dimension and not superquadratic as in the general case still allows for our method to be applied to high dimensions.

To exemplify our methods, we simulate our algorithm for some dimensions and number of sequences M. We run the simulations for $MU(d,8)$, as it is the smallest one that contains the T-gate. We consider the case of a quantum channel T that depolarizes to a random state $\sigma \in \mathcal{D}_d$ with probability $(1-p)$, that is

$$T(\rho) = p\rho + (1-p)\sigma, \tag{56}$$

where $\sigma \in \mathcal{D}_d$ is chosen uniformly at random from the set of states. Although the state σ is chosen at random each time we run the protocol, note that it is also fixed for each run. This implies that this quantum channel will in general not be covariant, as $\sigma \neq \mathbb{I}/d$ almost surely. It is not difficult to see that for this class of channels the entanglement fidelity is $F_e(T) = (p(d^2-1)+1)/d^2$ and we, therefore, measure our error in terms of the parameter p. The results are summarized in table 1.

We also obtain numerical results for unitary noise models. Here we consider quantum channels that are given by a conjugation with a unitary U of the form

$$U = \otimes_{j=1}^{n} \mathrm{e}^{\mathrm{i}\theta_j \sigma_{X,j}}$$

for systems of n qubits, $\sigma_{X,j}$ the Pauli X matrix acting on the jth qubit and $\theta_j \in [0, 2\pi)$. We sampled channels of this form by picking the θ_j independently and uniformly at random from some interval $(0,a)$. The magnitude of a is a proxy for 'how noisy' this unitary will be on average. Moreover, we use the methods described in appendix A.2 to isolate the relevant parameters. The results are summarized in table 2. These numerical results of tables 1 and 2 clearly show that we may estimate the fidelity to a good degree with our procedure.

7.2. Clifford group

As mentioned before, the Clifford group is the usual setup of RB, as we only have to estimate one parameter and it is one of the main building blocks of quantum computing [23]. Thus, we apply our protocols based on approximate samples of the Haar distribution and generator based protocols to Clifford gates. It is known that the Clifford group on n qubits, $\mathcal{C}(n)$, is generated by the Hadamard gate H, the π−gate and the *CNOT* gate between different qubits, defined as

$$H = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}, \qquad \pi = \begin{pmatrix} 1 & 0 \\ 0 & \mathrm{i} \end{pmatrix} \qquad \text{and} \qquad \mathit{CNOT} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{pmatrix}, \tag{57}$$

Table 1. Error analysis of the RB protocol described in section 4 to the group $MU(d, 8)$ and depolarizing noise as defined in (56). We take the initial state to be $|0\rangle\langle 0|$, the POVM element to be $|0\rangle\langle 0|$, $p = 0.9$ and we always choose $m = 40$. Moreover, we generate 100 different channels for each combination of dimension and number of sequences. The table shows the resulting mean and median error as well as the standard deviation for different values of d and M. Here, we define the error to be given by $|F - \hat{F}|$, where F is the true average fidelity of the channel and $\hat{F}$ the estimate we obtain from our protocol. These results indicate that the protocol performs well with this range of parameters for several different dimensions, as we observed small errors for all combinations of dimension and M. Note that increasing the number of random sequences M by one order of magnitude reduced the error, although this certainly requires more experimental effort.

d	M	Mean error ($\times 10^{-3}$)	Median error ($\times 10^{-3}$)	Standard deviation ($\times 10^{-3}$)
64	1000	9.17	2.14	3.93
128	100	6.08	1.48	2.14
128	1000	5.17	1.01	1.13
1024	100	9.17	2.14	3.93
1024	1000	4.55	1.13	1.77

Table 2. Error analysis of the RB protocol described in section 4 to the group $MU(d, 8)$ and unitary noise. We generate 100 different channels for each value of a and always perform the protocol for 10 qubits. For each run of the protocol we generate 1000 sequences of gates and choose $m = 20$. The table shows the resulting mean and median error as well as the standard deviation for different values of a. Here, we define the error to be given by $|F - \hat{F}|$, where F is the true average fidelity of the channel and $\hat{F}$ the estimate we obtain from our protocol. These results indicate that the protocol performs well with this range of parameters and unitary noise.

a	Mean error ($\times 10^{-4}$)	Median error ($\times 10^{-4}$)	Standard deviation ($\times 10^{-4}$)
0.1	3.90	2.00	0.45
0.2	2.63	1.80	2.10
0.3	3.19	1.9	3.05
0.4	4.11	2.05	4.04
0.5	4.71	2.11	4.01

respectively. We refer to e.g. [46, section 5.8] for a proof of this claim. We need a set of generators that is closed under taking inverses for our purposes. All but the $\pi-$gate are their own inverse, so we add the inverse of the $\pi-$gate to our set of generators to assure that the random walk converges to the Haar measure on the Clifford group. That is, we will consider the set A of generators of the Clifford group $\mathcal{C}(n)$ consisting of Hadamard gates, $\pi-$gates and its inverse on each individual qubit and *CNOT* between any two qubits,

$$A = \{\pi_i, \pi_i^{-1}, H_i, CNOT_{i,j}\}. \tag{58}$$

To the best of our knowledge, there is no rigorous estimate available for the mixing time of the random walk generated by A and it would certainly be interesting to investigate this question further. However, based on our numerical results and the results of [42], we conjecture that it is rapidly mixing, i.e. $t_{\text{mix}} = \mathrm{O}(n^2 \log(n))$. This would be more efficient than the algorithm proposed in [47], which takes $\mathrm{O}(n^3)$ operations. To again test our methods we perform similar numerics as in the case of the monomial unitaries.

Table 3. For each combination of p, M and b we generate 20 different random quantum channels and perform generator RB for the Clifford group on five qubits. In all these cases we pick $m = 20$. The average error is defined as the average of the absolute value between the exact fidelity and the one estimated using our protocol. The table shows the average error and its standard deviation in terms of different choices of b, M and p.

p	b	M	Average error ($\times 10^{-3}$)	Standard deviation of error ($\times 10^{-4}$)
0.98	10	10	5.49	1.38
0.95	10	100	1.44	3.92
0.95	5	100	1.52	7.94
0.95	5	20	1.56	7.44
0.90	10	20	3.20	1.58
0.80	10	50	8.63	6.01

Table 4. For each combination of p, M and b we generate 20 different random quantum channels and perform generator RB for the Clifford group on five qubits. In all these cases we pick $m = 20$. The average error is defined as the average of the absolute value between the exact fidelity and the one estimated using our protocol. The table shows the average error and its standard deviation in terms of different choices of b, M and p.

p	b	M	Average error ($\times 10^{-2}$)	Standard deviation of error ($\times 10^{-3}$)
0.7	5	100	2.07	1.15
0.65	5	100	2.29	1.95
0.60	5	100	27.1	52.30
0.55	5	100	44.5	67.30

Table 5. For each combination of p, M and b we generate 20 different convex combinations of the identity and a random unitary and perform generator RB for the Clifford group on five qubits. In all these cases we pick $m = 20$. The average error is defined as the average of the absolute value between the exact fidelity and the one estimated using our protocol. The table shows the average error and its standard deviation in terms of different choices of b, M and p.

p	b	M	Average error ($\times 10^{-3}$)	Standard deviation of error ($\times 10^{-4}$)
0.98	10	100	2.30	9.44
0.95	10	100	1.15	9.19
0.90	10	100	3.62	2.22
0.85	10	100	6.67	39.4
0.80	10	100	83.4	55.9

We simulate the following noise model: we first pick a random isometry $V : \left(\mathbb{C}^2\right)^{\otimes n} \to \left(\mathbb{C}^2\right)^{\otimes n} \otimes \left(\mathbb{C}^2\right)^{\otimes n}$ and generate the quantum channel

$$T(\rho) = p\rho + (1-p)\mathrm{tr}_2\left(V\rho V^\dagger\right), \tag{59}$$

where tr_2 denotes the partial trace over the second tensor factor. That is, T is just the convex combination of the identity and a random channel and is δ-covariant w.r.t. a group with $\delta = p$. This sampling procedure ensures that the channel T will not have any further symmetries. From the discussion in section 6 we expect this to work best for p close to one. The results for

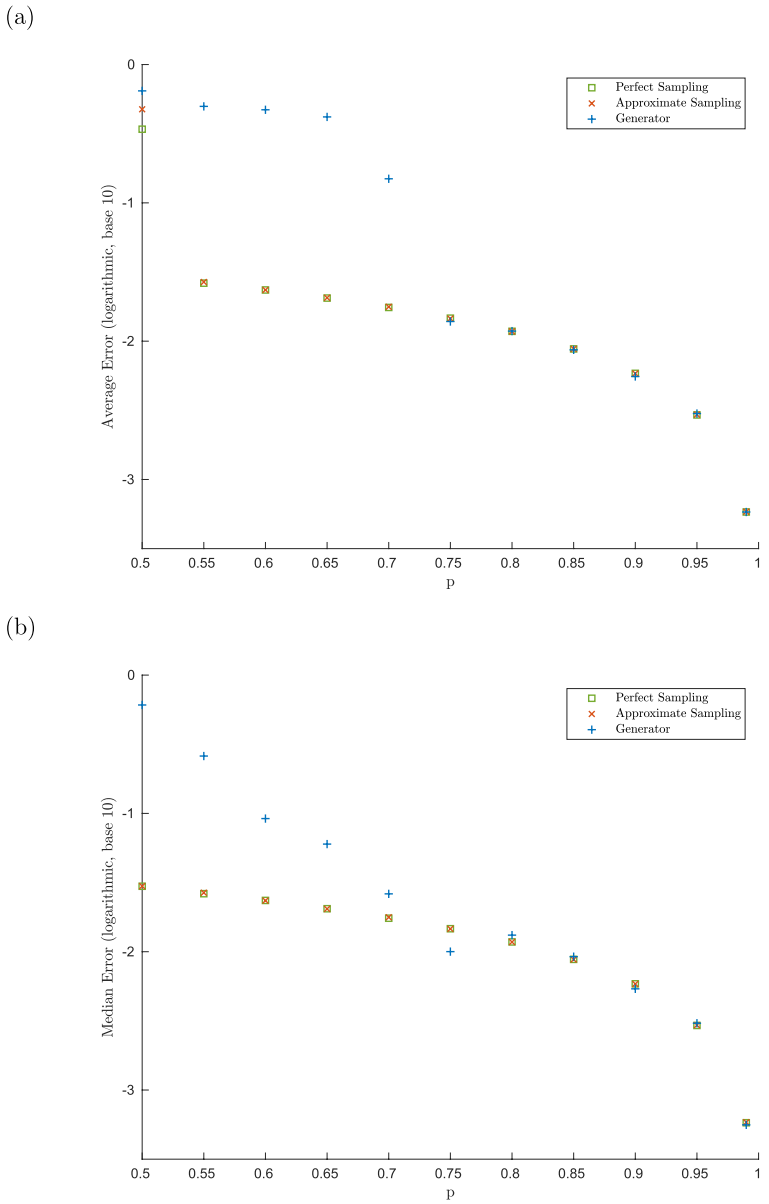


Figure 1. Plot of the average error (a) and mean error (b) as a function of p for different versions of the RB protocol for the Clifford group and the random quantum channel noise model as defined in (59). For each value of p we generated 20 instances of the random channel with $M = 100$ and $m = 20$. For the generator RB we chose $b = 5$ and to obtain the approximate samples we ran the chain for 20 steps.

p close to one are summarized in table 3. The average error increases as the channel becomes noisier, but generally speaking we are able to obtain an estimate which is 10^{-3} close to the true value with M around 20 and $m = 20$.

We also performed some numerical experiments for p significantly away from one, which are summarized in table 4.

The noise model above favors quantum channels with a high Kraus rank. Here we also consider the case of quantum channels of the form

$$T(\rho) = p\rho + (1-p)U\rho U^{\dagger},$$

where U is a randomly chosen (Haar) unitary. These channels have Kraus rank 2 and are δ-covariant with $\delta = p$. The numerical results can be found in table 5.

These results show that these methods are effective to estimate the average fidelity under less restrictive assumptions on the gates we may implement using RB if we have a high fidelity, as indicated in tables 3 and 5. However, in case we do not have a high fidelity, these methods are not reliable, as can be seen in table 4. Note that our numerical results seem to indicate that the cut-off of the range of average fidelities we can reliably detect occurs at larger values of the fidelity in the case of channels with lower Kraus rank, as can be see in table 5. This should not severely restrict the applicability of these methods, as one is usually interested in the high fidelity regime when performing RB.

Finally, in figure 1 we compare the three different RB protocols discussed in this paper. We compare the usual RB protocol, which we call the perfect sampling protocol, to the one with approximate samples and the generator RB for the random quantum channel noise model. The curve makes clear that using approximate and exact samples leads to virtually indistinguishable estimates and that all protocols have similar performance for p close to one.

8. Conclusion and open problems

We have generalized the RB protocol to estimate the average gate fidelity of unitary representations of arbitrary finite groups. Our protocol is efficient when multiplying, inverting and sampling elements from the group can be done efficiently and we have shown some potential applications that go beyond the usual Clifford one. Moreover, we showed that using approximate samples instead of perfect ones from the Haar measure on the group does not lead to great errors. This can be seen as a stability result for RB protocols w.r.t. sampling which was not available in the literature and is also relevant in the Clifford case. We hope that this result can be useful in practice when one is not given a full description of the group but rather a set of generators. Moreover, we have shown how to perform RB by just implementing a set of generators and one arbitrary gate under some noise models. This protocol could potentially be more feasible for applications, as the set of gates we need to implement is on average simpler.

However, some questions remain open and require further work. It is straightforward to generalize the technique of interleaved RB to this more general scenario and this would also be a relevant development. It would be important to derive confidence intervals for the estimates as was done for the Clifford case in [7, 8]. Moreover, it would be relevant to estimate not only the mean fidelity but also the variance of this quantity. The assumption that the noisy channel is the same for all gates is not realistic in many scenarios and should be seen as a 0-order approximation, as in [37]. It would be desirable to generalize our results to the case in which the channel depends weakly on the gate.

Acknowledgments

DSF acknowledges support from the graduate program TopMath of the Elite Network of Bavaria, the TopMath Graduate Center of TUM Graduate School at Technische Universität München and by the Technische Universität München Institute for Advanced Study, funded

by the German Excellence Initiative and the European Union Seventh Framework Programme under grant agreement no. 291763.

AKH's work is supported by the Elite Network of Bavaria through the PhD programme of excellence *Exploring Quantum Matter*.

Appendix A. Numerical considerations

Here we gather some comments on the numerical issues associated with the RB procedure when estimating more than one parameter.

A.1. Fitting the data to several parameters

In order to be able to estimate the average fidelity following the protocols discussed so far, it is necessary to fit noisy data points $\{x_i\}_{i=1}^m \subset \mathbb{R}$ to a curve $f : \mathbb{R} \to \mathbb{R}$ of the form

$$f(x) = a_0 + \sum_{k=1}^{n} a_k \mathrm{e}^{-b_k x}, \tag{A.1}$$

with $a_0, a_1, \ldots, a_n, b_1, \ldots, b_n \in \mathbb{C}$. Although this may look like an innocent problem at first sight, fitting noisy data to exponential curves is a difficult problem from the numerical point of view for large n. It suffers from many stability issues, as thoroughly discussed in [48]. Here we are going to briefly comment on some of the issues and challenges faced when trying to fit the data, although we admittedly only scratch the surface. For a more thorough analysis of some methods and issues, we refer to [48, 49].

We assume that we know the maximum number of different parameters, $2n + 1$, which we are fitting. This is given by the structure of the unitary representation at hand, as discussed in lemma 11. Luckily, significant progress has been made in the recent years to develop algorithms to overcome the issues faced in this setting and it is now possible to fit curves to data with a moderate number of parameters. It is also noteworthy that for $n = 2$ there exist stable algorithms based on geometric sums [49] which works for equispaced data, as is our case. For estimating more than two parameters one can use the algorithms proposed in [48], available at [50]. It should be said that the reliability and convergence of most algorithms found in the literature depends strongly on the choice of a good initial point. This tends not to be a problem, as we might have some assumptions where our fidelity approximately lies and choose the initial b_k accordingly. What could be another source of numerical instabilities is the fact that we have to input the model with a number of parameters, n. In case the eigenvalues of T are very close for different irreps, then this will lead to numerical instabilities. This is the case if the noise is described by a depolarizing channel, for example. Furthermore, it might be the case that the initial state in our protocol does not intersect with all eigenspaces of the channel. This may lead to some parameters a_k being zero and we are not able to estimate some of the b_k from them.

Moreover, it is in principle not possible to tell which parameter corresponds to which irrep given the decomposition in lemma 11, which is again necessary to estimate the trace of the channel. So even in the case in which we have a small number of parameters, it is important that the different irreps associated to our parameters have a similar dimension or to assume that the spectrum of the twirled channel contains eigenvalues that are very close to each other. In this way, the most pessimistic estimate on the fidelity, as defined in (37), is not very far from the most optimistic, defined in (38). This is one of the reasons we focus on examples that

only have a small number of parameters, say one or two, and irreps of a similar dimension to avoid having numerical instabilities or estimates that range over an interval that is too large.

It is therefore important to develop better schemes to fit the data in the context of RB for more than one or two parameters. This is important from a statistical point of view, as it would be desirable to obtain confidence intervals for the parameters from the RB data. We will further develop this issue in appendix B. It would be worthwhile pursuing a Bayesian approach to this problem, as was done in [51] for the usual RB protocol.

A.2. Isolating the parameters

One way to possibly deal with this issue is to isolate each parameter, that is, by preparing states that only have support on one of the irreps that are not the trivial one. In the case of non-degenerate unitary representations, discussed in theorem 5, we have the following:

Theorem A.1 (Isolating parameters). *Let $U : G \to \mathcal{M}_d$ be a simply covariant irrep of a finite group G and $T : \mathcal{M}_d \to \mathcal{M}_d$ a channel which is covariant w.r.t. U. Then, for all eigenvalues λ_α there is a quantum state $\rho_\alpha = \frac{\mathbb{I}}{d} + X$, where $X = X^\dagger$ and $\mathrm{Tr}\,(X) = 0$, such that*

$$T^m(\rho_\alpha) = \frac{\mathbb{I}}{d} + \lambda_\alpha^m X. \tag{A.2}$$

Proof. Consider the projections to the irreducible subspaces P_α defined in (11). For a self-adjoint operator $X \in \mathcal{M}_d$ we have that

$$P_\alpha(X)^\dagger = \frac{\chi^\alpha(e)}{|G|} \sum_{g\in G} \chi^\alpha(g)\, U_g^\dagger X U_g = P_\alpha(X),$$

as we are summing over the whole group and $\overline{\chi^\alpha(g^{-1})} = \chi^\alpha(g)$. Therefore, we have that the P_α are hermiticity preserving. As the image of P_α is the eigenspace corresponding to the irreps, we thus only have to show that there exists a self-adjoint X such that $P_\alpha(X) \neq 0$. But the existence of such an X is clear, as we may choose a basis of $\mathcal{M}_d$ that consists of self-adjoint operators. Moreover, as for α not the trivial representation all eigenvectors are orthogonal to $\mathbb{I}$, it follows that $\mathrm{Tr}\,(X) = 0$ and that for $\epsilon > 0$ small enough $\frac{\mathbb{I}}{d} + \epsilon X$ is positive semidefinite. To finish the proof, note that simply irreducible channels always satisfy

$$T(\mathbb{I}) = \mathbb{I}.$$

□

Note that this also proves that the spectrum of irreducibly covariant channels is always real. That is, if we can prepare a state such as in (A.2), then we can perform the RB with this as an initial state and estimate the eigenvalue corresponding to each irrep. This would bypass the problems discussed in appendix A.1. The proof of theorem A.1 already hints a way of determining how to isolate the parameter: just apply the projector P_α to some states ρ_i. If the output is not zero, then we can in principle write down a state that 'isolates' the parameter as in the proof of theorem A.1. This idea was explored in [40] to obtain a way of isolating the parameters for irreducibly covariant channels and general representation under additional assumptions. However, in the case of the monomial unitary matrices discussed in section 7.1, we can examine the projections and see how to isolate the parameters. To isolate the parameter α in (55), we can prepare initial states $\rho \in \mathcal{D}_d$ that have $1/d$ as their diagonal elements and at least one nonzero off-diagonal element, as then the projector corresponding to β vanishes on ρ and does not vanish on the one corresponding to α. To isolate the parameter β, one can prepare

states ρ that are diagonal in the computational basis but are not the maximally mixed state, as can be seen by direct inspection.

Appendix B. Statistical considerations

One of the main open questions left in our work is how to derive good confidence intervals for the average fidelity. For the case of the Clifford group, discussed in section 7.2, one can directly apply the results of [7, 8], but it is not clear how one should pick m and M for arbitrary finite groups. Especially in the case in which we are not working with Cliffords, it is not clear how many sequences per point, M, we should gather and how big m should be, as it depends on the choice of the algorithm picked for fitting the curve. As noted in appendix A.1, this is not a trivial problem from a numerical point of view. However, it is possible to obtain estimates on how much the observed survival probability deviates from its expectation value by just using Hoeffding's inequality:

Theorem B.1. *Let $\bar{F}(m, E, \rho)$ be the observed average fidelity with M sequences and $F(m, E, \rho)$ the average fidelity for any of the protocols discussed before and $\epsilon > 0$. Then:*

$$\mathbb{P}(|F(m, E, \rho) - \bar{F}(m, E, \rho)| \geqslant \epsilon) \leqslant \mathrm{e}^{-2M\epsilon^2}. \tag{B.1}$$

Proof. This is just a straightforward application of Hoeffding's inequality [52], as $\bar{F}(m, E, \rho)$ is just the empirical average of a random variable whose value is contained in $[0, 1]$ and whose expectation value is $F(m, E, \rho)$. □

This bound is extremely general, as we did not even have to use any property of the random variables or of the group at hand. One should not expect it to perform well for specific cases and the scaling it gives is still undesirable for applications. Indeed, to assure we are 10^{-4} close to the expectation value with probability of 0.95, we need around 6×10^8 sequences, which is not feasible. Thus, it is necessary to derive more refined bounds for specific groups.

Appendix C. Proof of theorem 17

Theorem C.1. *Let μ be the Haar measure on G and $\nu_1, \ldots, \nu_m$ probability measures on G s.t.*

$$\|\mu - \nu_k\|_1 \leqslant \epsilon_k, \tag{C.1}$$

for all $1 \leqslant k \leqslant m$ and $\epsilon_k \geqslant 0$. Denote by $\tilde{\nu}$ the distribution of $(\mathcal{D}_1, \ldots, \mathcal{D}_m)$ if we pick the $\mathcal{U}_k$ independently from ν_k . Then

$$\|\mathcal{T}_{\tilde{\nu},m}(T) - \mathcal{T}(T)^m\|_{1\to 1} \leqslant 4\sqrt{\frac{\log(|G|)}{1 - |G|^{-1}} \sum_{k=1}^{m} \epsilon_k}. \tag{C.2}$$

Proof. From lemma 16 it suffices to show

$$\|\tilde{\nu} - \mu^{\otimes m}\|_1 \leqslant 2\sqrt{\frac{\log(|G|)}{1 - |G|^{-1}} \sum_{k=1}^{m} \epsilon_k},$$

as the $1 \to 1$ norm contracts under quantum channels [41].

We will first show that

$$\|\tilde{\nu} - \mu^{\otimes m}\|_1 = \| \otimes_{k=1}^m \nu_k - \mu^{\otimes m}\|_1.$$

We may rewrite the distribution $\tilde{\nu}$ in terms of the ν_k as follows:

$$\begin{aligned}\mathbb{P}(\mathcal{D}_1 = g_1, \mathcal{D}_2 = g_2, \ldots, \mathcal{D}_m = g_m) &= \mathbb{P}(U_1 = g_1, U_2 = g_2 g_1^{-1}, \ldots, U_m = g_m g_{m-1}^{-1}) \\ &= \nu_1(g_1)\nu_2(g_2 g_1^{-1}) \ldots \nu_m(g_m g_{m-1}^{-1}),\end{aligned}$$

as the U_{g_i} are independent.

Note that the map $\sigma : G^m \to G^m$, $(g_1, \ldots, g_m) \mapsto \left(g_1, g_2 g_1^{-1}, \ldots, g_m g_{m-1}^{-1}\right)$ is bijective. Moreover, we have $\tilde{\nu} = \otimes_{k=1}^m \nu_k \circ \sigma$. As the total variation norm is invariant under compositions with bijections on the state space, we have

$$\|\tilde{\nu} - \mu^{\otimes m}\|_1 = \| \otimes_{k=1}^m \nu_k \circ \sigma - \mu^{\otimes m}\|_1 = \| \otimes_{k=1}^m \nu_k - \mu^{\otimes m} \circ \sigma^{-1}\|_1 = \| \otimes_{k=1}^m \nu_k - \mu^{\otimes m}\|_1,$$

where the last equality follows from the fact that the Haar measure is invariant under bijections. We will now bound $\| \otimes_{k=1}^m \nu_k - \mu^{\otimes m}\|_1$. By Pinsker's inequality [53], we have

$$\| \otimes_{k=1}^m \nu_k - \mu^{\otimes m}\|_1^2 \leqslant 4D\left(\otimes_{k=1}^m \nu_k \,||\, \mu^{\otimes m}\right) = 4 \sum_{k=1}^m D\left(\nu_k \,||\, \mu\right). \tag{C.3}$$

Here D is the relative entropy. In [53, theorem 1] they show that

$$D\left(\nu_k \,||\, \mu\right) \leqslant \frac{\log(|G|)}{1 - |G|^{-1}} \|\mu - \nu_k\|_1$$

and from (42) it follows that

$$D\left(\nu_k \,||\, \mu\right) \leqslant \frac{\log(|G|)}{1 - |G|^{-1}} \epsilon_k. \tag{C.4}$$

Combining (C.4) with (C.3) and taking the square root yields the claim. □

Appendix D. Proof of theorem 20

Theorem D.1. *Let T be $\delta-$covariant w.r.t. a unitary representation $U : G \to \mathcal{M}_d$ of a finite group G, A a subset of G that generates G and is closed under inversion and $\delta > 0$. Suppose we run the protocol above with $b = t_1(m^{-1}\epsilon)$ for some ε and $m \geqslant b$. Then*

$$\|\mathcal{T}(T)^m - \mathbb{E}(\mathcal{S}_{b,b+m+1})\|_{1\to 1} \leqslant \epsilon + \mathrm{O}\left(\delta^2 bm\right). \tag{D.1}$$

Proof. Let T_c and T_n be as in definition 19. Then we have

$$\mathcal{T}(T) = (1 - \delta)T_c + \delta \mathcal{T}(T_n),$$

as T_c is already covariant, and

$$\mathcal{T}(T)^m = (1-\delta)^m T_c^m + \delta(1-\delta)^{m-1}\sum_{j=0}^{m-1} T_c^j \mathcal{T}(T_n) T_c^{m-j-1} + \delta^2(1-\delta)^{m-2}\sum_{j_1+j_2+j_3=m-2} T_c^{j_1}\mathcal{T}(T_n)T_c^{j_2}\mathcal{T}(T_n)T_c^{j_3} + \mathrm{O}(\delta^3). \tag{D.2}$$

Moreover, as T_c is covariant w.r.t. this unitary representation, we have

$$\mathbb{E}(\mathcal{S}_{b,m+b+1}) = (1-\delta^m)T_c + \delta(1-\delta)^{m-1}\sum_{j=0}^{m-1}\mathbb{E}\left(T_c^j \mathcal{D}_{m-j} T_n \mathcal{D}_{m-j}^* T_c^{m-j-1}\right)$$
$$+\delta^2(1-\delta)^{m-2}\sum_{j_1+j_2+j_3=m-2}\mathbb{E}\left(T_c^{j_1}\mathcal{D}_{j_2+1}T_n\mathcal{D}_{j_2+1}^* T_c^{j_2}\mathcal{D}_{j_3+1}T_n\mathcal{D}_{j_3+1}^* T_c^{j_3}\right) + \mathrm{O}(\delta^3). \tag{D.3}$$

It is clear that the terms of zero−order in δ in (D.2) and (D.3) coincide. Comparing each of the summands of first order we obtain:

$$\mathbb{E}\left(T_c^j \mathcal{D}_{m-j} T_n \mathcal{D}_{m-j}^* T_c^{m-j-1}\right) - T_c^j \mathcal{T}(T_n) T_c^{m-j-1}$$
$$= \sum_{g\in G}\left(\nu_{m-j}(g) - \frac{1}{|G|}\right) T_c^j \mathcal{U}_g T_n \mathcal{U}_g^* T_c^{m-j-1},$$

where ν_{m-j} is the distribution of $\mathcal{D}_{m-j}$. Comparing the terms of second order we obtain:

$$\mathbb{E}\left(T_c^{j_1}\mathcal{D}_{j_2+1}T_n\mathcal{D}_{j_2+1}^* T_c^{j_2}\mathcal{D}_{j_3+1}T_n\mathcal{D}_{j_3+1}^* T_c^{j_3}\right) - T_c^{j_1}\mathcal{T}(T_n)T_c^{j_2}\mathcal{T}(T_n)T_c^{j_3}$$
$$= \sum_{g_1,g_2\in G}\left(\tau_{j_3+1,j_2+1}(g_1,g_2) - \frac{1}{|G|^2}\right) T_c^{j_1}\mathcal{U}_{g_1}T_n\mathcal{U}_{g_1}^* T_c^{j_2}\mathcal{U}_{g_2}T_n\mathcal{U}_{g_2}^* T_c^{j_3}.$$

Here τ_{j_3+1,j_2+1} is the joint distribution of $\mathcal{D}_{j_3+1}$ and $\mathcal{D}_{j_2+1}$. Then, using arguments similar to those of theorem 17, we have that

$$\|\mathcal{T}(T)^m - \mathbb{E}(S_{b,m+1})\|_{1\to 1}$$
$$\leqslant \delta(1-\delta)^{m-1}\sum_{j=1}^{m}\|\nu_j - \mu\| + \delta^2(1-\delta)^{m-2}\sum_{j_1=1}^{m-1}\sum_{j_2=j_1+1}^{m}\|\tau_{j_1,j_2} - \mu^{\otimes 2}\|_1 + \mathrm{O}(\delta^3).$$

Now, from our choice of b, we have $\|\nu_j - \mu\|_1 \leqslant \frac{\epsilon}{m}$. Furthermore, we have that

$$\tau_{j_1,j_2}(g_1,g_2) = \mathbb{P}(\mathcal{D}_{j_1} = \mathcal{U}_{g_1}, \mathcal{D}_{j_2} = \mathcal{U}_{g_2}) = \mathbb{P}(\mathcal{D}_{j_2} = \mathcal{U}_{g_2} | \mathcal{D}_{j_1} = \mathcal{U}_{g_1})\mathbb{P}(\mathcal{D}_{j_1} = \mathcal{U}_{g_1}).$$

By the construction of the $\mathcal{D}_j$, it holds that

$$\mathbb{P}(\mathcal{D}_{j_2} = \mathcal{U}_{g_2} | \mathcal{D}_{j_1} = \mathcal{U}_{g_1}) = \pi^{j_2-j_1}(g_1,g_2),$$

where π is the stochastic matrix of the chain generated by A. From this we obtain

$$\sum_{g_1,g_2\in G}\left|\tau_{j_1,j_2}(g_1,g_2) - \frac{1}{|G|^2}\right| = \sum_{g_1,g_2\in G}\left|\nu_{j_1}(g_1)\pi^{j_2-j_1}(g_1,g_2) - \frac{1}{|G|^2}\right|$$
$$\leqslant \sum_{g_1,g_2\in G}\left|\nu_{j_1}(g_1) - \frac{1}{|G|}\right|\pi^{j_2-j_1}(g_1,g_2) + \left|\frac{1}{|G|}\pi^{j_2-j_1}(g_1,g_2) - \frac{1}{|G|^2}\right|. \tag{D.4}$$

As the matrix π is doubly stochastic, summing over g_2 first

$$\sum_{g_1,g_2\in G}\left|\nu_{j_1}(g_1)-\frac{1}{|G|}\right|\pi^{j_2-j_1}(g_1,g_2)=\sum_{g_1\in G}\left|\nu_{j_1}(g_1)-\frac{1}{|G|}\right|\leqslant \epsilon m^{-1},$$

which again follows from our choice of b. We now estimate the other term in (D.4),

$$\sum_{j_1=1}^{m-1}\sum_{j_2=j_1+1}^{m}\sum_{g_1,g_2\in G}\frac{1}{|G|}\left|\pi^{j_2-j_1}(g_1,g_2)-\frac{1}{|G|}\right|. \tag{D.5}$$

Note that for a fixed g_1,

$$\sum_{g_2\in G}\left|\pi^{j_2-j_1}(g_1,g_2)-\frac{1}{|G|}\right|$$

is just the total variation distance between the Markov chain starting at g_1 and the Haar measure after j_2-j_1 steps. Thus, in case $j_2-j_1\geqslant t_1\left(\epsilon m^{-1}\right)$,

$$\sum_{g_1,g_2\in G}\frac{1}{|G|}\left|\pi^{j_2-j_1}(g_1,g_2)-\frac{1}{|G|}\right|\leqslant\frac{\epsilon}{m} \tag{D.6}$$

and in case $j_2-j_1\leqslant t_1\left(\epsilon m^{-1}\right)$ we have the trivial estimate

$$\sum_{g_1,g_2\in G}\frac{1}{|G|}\left|\pi^{j_2-j_1}(g_1,g_2)-\frac{1}{|G|}\right|\leqslant 2. \tag{D.7}$$

Combining inequalities (D.6) and (D.7), we obtain

$$\sum_{j_1=1}^{m-1}\sum_{j_2=j_1+1}^{m}\sum_{g_1,g_2\in G}\frac{1}{|G|}\left|\pi^{j_2-j_1}(g_1,g_2)-\frac{1}{|G|}\right|=\mathcal{O}(mt_1(\epsilon m^{-1})).$$

Putting all inequalities together, we obtain the claim. □

Appendix E. Proof of lemma 24

Lemma E.1 (Structure of channels covariant w.r.t. monomial unitaries). *Let $MU(d,n)$ be such that $n\geqslant 3$ and $T:\mathcal{M}_d\to\mathcal{M}_d$ a quantum channel. Then the following are equivalent:*

(i) $T(\rho)=UT\left(U^\dagger\rho U\right)U^\dagger\quad\forall U\in MU(d,n),\rho\in\mathcal{S}_d$.
(ii) There are $\alpha,\beta\in\mathbb{R}$ *so that*

$$T(\cdot)=\mathrm{Tr}\,(\cdot)\,\frac{\mathbb{I}}{d}+\alpha\left(\mathrm{id}-\sum_{i=1}^{d}|i\rangle\langle i|\,\langle i|\cdot|i\rangle\right)+\beta\left(\sum_{i=1}^{d}|i\rangle\langle i|\,\langle i|\cdot|i\rangle-\mathrm{Tr}\,(\cdot)\,\frac{\mathbb{I}}{d}\right). \tag{E.1}$$

Moreover, the terms in the r.h.s. of (55) are projections of rank 1, d^2-d and $d-1$, respectively.

Proof. (2) ⇒ (1) can be seen by direct inspection. In order to prove the converse, we consider the Choi–Jamiolkowski state $\tau_T := \frac{1}{d}\sum_{i,j=1}^{d} T(|i\rangle\langle j|) \otimes |i\rangle\langle j|$. Then (1) is equivalent to the statement that τ_T commutes with all unitaries of the form $U \otimes \bar{U}$, $U \in MU(d,n)$. That is, we have

$$\sum_{i,j=1}^{d} U \otimes \bar{U}\left(T(|i\rangle\langle j|) \otimes |i\rangle\langle j|\right)(U \otimes \bar{U})^\dagger = \sum_{i,j=1}^{d} T(|i\rangle\langle j|) \otimes |i\rangle\langle j|\,.$$

Restricting to the subgroup of diagonal unitaries in $MU(d,n)$, for which $U^\dagger = \bar{U}$, we have

$$\sum_{i,j=1}^{d} \mathrm{e}^{\mathrm{i}(\phi_j-\phi_i)} UT(|i\rangle\langle j|)\bar{U} \otimes |i\rangle\langle j| = \sum_{i,j=1}^{d} T(|i\rangle\langle j|) \otimes |i\rangle\langle j|\,,$$

where $\mathrm{e}^{\mathrm{i}\phi_i}$ is the ith diagonal entry of U. Comparing the tensor factors it follows that

$$\mathrm{e}^{\mathrm{i}(\phi_j-\phi_i)} UT(|i\rangle\langle j|)\bar{U} = T(|i\rangle\langle j|). \tag{E.2}$$

We will now show that we have

$$\tau_T = \sum_{i,j=1}^{d} A_{ij}\,|i\rangle\langle i| \otimes |j\rangle\langle j| + B_{ij}\,|i\rangle\langle j| \otimes |i\rangle\langle j|\,. \tag{E.3}$$

We have

$$T(|i\rangle\langle j|) = \sum_{k,l=1}^{d} a_{k,l}\,|k\rangle\langle l|$$

for some $a_{k,l} \in \mathbb{C}$. From (E.2) it follows that

$$\sum_{k,l=1}^{d} \mathrm{e}^{\mathrm{i}(\phi_k-\phi_l)} a_{k,l}\,|k\rangle\langle l| = \mathrm{e}^{\mathrm{i}(\phi_i-\phi_j)} \sum_{k,l=1}^{d} a_{k,l}\,|k\rangle\langle l| \tag{E.4}$$

for all diagonal unitaries. Again comparing both sides of (E.4) we have $a_{k,l}\mathrm{e}^{\mathrm{i}(\phi_i-\phi_j)} = \mathrm{e}^{\mathrm{i}(\phi_k-\phi_l)} a_{k,l}$. Suppose now $i \neq j$. For $a_{k,l} \neq 0$ we have

$$\mathrm{e}^{\mathrm{i}(\phi_i-\phi_j)} = \mathrm{e}^{\mathrm{i}(\phi_k-\phi_l)} \tag{E.5}$$

for all diagonal entries of diagonal unitaries. If k, l, i and j are all pairwise distinct, we have $i = k$ and $j \neq l$ or $i \neq k$ and $j = l$, then it is clear that we may always find a combination of ϕ_k, ϕ_l, ϕ_i and ϕ_j such that (E.5) is not satisfied, a contradiction. For $i = l$ and $k = j$, it is only possible to find such a combination for $n > 2$, as otherwise $\phi_i - \phi_j = -(\phi_i - \phi_j)$ always holds. This proves that we have

$$T(|i\rangle\langle j|) = B_{ij}\,|i\rangle\langle j| \tag{E.6}$$

for $i \neq j$. For $i = j$ we have analogously that

$$UT(|i\rangle\langle i|)\bar{U} = \sum_{k,l=1}^{d} \mathrm{e}^{\mathrm{i}(\phi_k-\phi_l)} a_{k,l}\,|k\rangle\langle l| = \sum_{k,l=1}^{d} a_{k,l}\,|k\rangle\langle l|\,.$$

In this case, we have $a_{k,l} = \mathrm{e}^{\mathrm{i}(\phi_k - \phi_l)} a_{k,l}$ for all possible phases of the form $e^{\mathrm{i}(\phi_k - \phi_l)}$. It is then clear that $a_{k,l} = 0$ unless $k = l$ by a similar argument as before. This gives

$$T\left(|i\rangle\langle i|\right) = \sum_{j=1}^{d} A_{ij} \, |j\rangle\langle j| \, . \tag{E.7}$$

Putting together (E.7) and (E.6) implies (E.3). Next, we will exploit that τ_T commutes in addition with permutations of the form $U_\pi \otimes U_\pi$ for all $\pi \in S_d$. For $i \neq j$ this implies that $A_{i,j} = A_{\pi(i),\pi(j)}$ and $B_{i,j} = B_{\pi(i),\pi(j)}$ so that there is only one independent off-diagonal element for each A and B. The case $i = j$ leads to a third parameter that is a coefficient in front of $\sum_\alpha |ii\rangle\langle ii|$. Translating this back to the level of projections then yields (55). The fact that the terms of (55) are projections can be seen by direct inspection. Note that the term corresponding to α is the difference of two projections, the identity and projection onto diagonal matrices. As the rank of the identity is d^2 and the space of diagonal matrices has dimension d, we obtain the claim. The same reasoning applies to the term corresponding to β, as it is the difference of the projection onto diagonal matrices and the projection onto the maximally mixed state. The latter is a projection of rank 1, which yields a rank of $d - 1$ for their difference. □

ORCID iDs

D S França https://orcid.org/0000-0001-9699-5994
A K Hashagen https://orcid.org/0000-0002-8682-6510

References

[1] Poyatos J F, Cirac J I and Zoller P 1997 *Phys. Rev. Lett.* **78** 390–3
[2] Chuang I L and Nielsen M A 1997 *J. Mod. Opt.* **44** 2455–67
[3] Knill E, Leibfried D, Reichle R, Britton J, Blakestad R B, Jost J D, Langer C, Ozeri R, Seidelin S and Wineland D J 2008 *Phys. Rev.* A **77** 012307
[4] Emerson J, Silva M, Moussa O, Ryan C, Laforest M, Baugh J, Cory D G and Laflamme R 2007 *Science* **317** 1893–6
[5] Lévi B, López C C, Emerson J and Cory D G 2007 *Phys. Rev.* A **75** 022314
[6] Emerson J, Alicki R and Zyczkowski K 2005 *J. Opt.* B **7** S347
[7] Wallman J J and Flammia S T 2014 *New J. Phys.* **16** 103032
[8] Helsen J, Wallman J J, Flammia S T and Wehner S 2017 in preparation
[9] Chow J M, Gambetta J M, Tornberg L, Koch J, Bishop L S, Houck A A, Johnson B R, Frunzio L, Girvin S M and Schoelkopf R J 2009 *Phys. Rev. Lett.* **102** 090502
[10] Ryan C A, Laforest M and Laflamme R 2009 *New J. Phys.* **11** 013034
[11] Olmschenk S, Chicireanu R, Nelson K D and Porto J V 2010 *New J. Phys.* **12** 113007
[12] Brown K R, Wilson A C, Colombe Y, Ospelkaus C, Meier A M, Knill E, Leibfried D and Wineland D J 2011 *Phys. Rev.* A **84** 030303
[13] Gaebler J P *et al* 2012 *Phys. Rev. Lett.* **108** 260503
[14] Barends R *et al* 2014 *Nature* **508** 500–3
[15] Xia T, Lichtman M, Maller K, Carr A W, Piotrowicz M J, Isenhower L and Saffman M 2015 *Phys. Rev. Lett.* **114** 100503
[16] Muhonen J T *et al* 2015 *J. Phys.: Condens. Matter* **27** 154205
[17] Asaad S, Dickel C, Langford N K, Poletto S, Bruno A, Rol M A, Deurloo D and DiCarlo L 2016 *nph Quantum Inf.* **2** 16029
[18] Hashagen A K, Flammia S T, Gross D and Wallman J J 2018 in preparation
[19] Brown W G and Eastin B 2018 *Phys. Rev.* A **97** 062323

[20] Cross A W, Magesan E, Bishop L S, Smolin J A and Gambetta J M 2016 *npj Quantum Inf.* **2** 16012
[21] Carignan-Dugas A, Wallman J J and Emerson J 2015 *Phys. Rev.* A **92** 060302
[22] Van den Nest M 2011 *New J. Phys.* **13** 123004
[23] Nielsen M A and Chuang I L 2009 *Quantum Computation and Quantum Information* (Cambridge: Cambridge University Press)
[24] Heinosaari T and Ziman M 2012 *The Mathematical Language of Quantum Theory: from Uncertainty to Entanglement* (Cambridge: Cambridge University Press)
[25] Simon B 1996 *Representations of Finite and Compact Groups* (*Graduate Studies in Mathematics* vol 10) (Providence, RI: American Mathematical Society)
[26] Mendl C B and Wolf M M 2009 *Commun. Math. Phys.* **289** 1057–86
[27] Mozrzymas M, Studziński M and Datta N 2017 *J. Math. Phys.* **58** 052204
[28] Levin D A and Peres Y 2017 *Markov Chains and Mixing Times* vol 107 (Providence, RI: American Mathematical Society)
[29] Saloff-Coste L 2004 *Probability on Discrete Structures* (New York: Springer) pp 263–346
[30] Wehrl A 1978 *Rev. Mod. Phys.* **50** 221–60
[31] Nielsen M A 2002 *Phys. Lett.* A **303** 249–52
[32] Horn R A and Johnson C R 2009 *Matrix Analysis* (Cambridge: Cambridge University Press)
[33] Burgarth D, Chiribella G, Giovannetti V, Perinotti P and Yuasa K 2013 *New J. Phys.* **15** 073045
[34] Wallman J J 2018 *Quantum* **2** 47
[35] Proctor T, Rudinger K, Young K, Sarovar M and Blume-Kohout R 2017 *Phys. Rev. Lett.* **119** 130502
[36] Gambetta J M *et al* 2012 *Phys. Rev. Lett.* **109** 240504
[37] Magesan E, Gambetta J M and Emerson J 2012 *Phys. Rev.* A **85** 042311
[38] Magesan E, Gambetta J M and Emerson J 2011 *Phys. Rev. Lett.* **106** 180504
[39] Dankert C, Cleve R, Emerson J and Livine E 2009 *Phys. Rev.* A **80** 012304
[40] Helsen J, Xue X, Vandersypen L M K and Wehner S 2018 in preparation
[41] Pérez-García D, Wolf M M, Petz D and Ruskai M B 2006 *J. Math. Phys.* **47** 083506
[42] Harrow A W and Low R A 2009 *Commun. Math. Phys.* **291** 257–302
[43] Kliuchnikov V 2014 *PhD Thesis* University of Waterloo
[44] Leditzky F, Kaur E, Datta N and Wilde M M 2018 *Phys. Rev.* A **97** 012332
[45] Randall D 2006 *Comput. Sci. Eng.* **8** 30–41
[46] Gottesman D 1997 in preparation
[47] Koenig R and Smolin J A 2014 *J. Math. Phys.* **55** 122202
[48] Hokanson J 2013 *PhD Thesis* Rice University
[49] Holmström K and Petersson J 2002 *Appl. Math. Comput.* **126** 31–61
[50] Hokanson J 2014 Numerically stable and statistically efficient algorithms for large scale exponential fitting *PhD Thesis* Rice University
[51] Hincks I, Wallman J J, Ferrie C, Granade C and Cory D G 2018 in preparation
[52] Hoeffding W 1963 *J. Am. Stat. Assoc.* **58** 13–30
[53] Sason I 2015 Upper bounds on the relative entropy and Rényi divergence as a function of total variation distance for finite alphabets *2015 IEEE Information Theory Workshop–Fall (ITW) (11–15 October 2015)* (Piscataway, NJ: IEEE)