UNIVERSITY OF LEEDS

Entropic Constraints in Quantum Resource Theories of Magic and Asymmetry

Si Chen

Submitted in accordance with the requirements for the degree of Doctor of Philosophy

The University of Leeds

Faculty of Environmental and Physical Sciences

School of Physics and Astronomy

October 2025

Intellectual Property Statement

The candidate confirms that the work submitted is her own, except where work which has formed part of jointly authored publications has been included. The candidate's contribution and the other authors to this work has been explicitly indicated below. The candidate confirms that appropriate credit has been given within the thesis where reference has been made to the work of others.

Chapters 3 and 4 are based on *General Entropic Constraints on Calderbank-Shor-Steane Codes within Magic Distillation Protocols* by Rhea Alexander, **Si Gvirtz-Chen**, Nikolaos Koukoulekidis and David Jennings, PRX Quantum **4** 020359 (30 June 2023). David Jennings supervised the project and provided high-level advice and guidance alongside Nikolaos Koukoulekidis. The bulk of research was jointly undertaken by Rhea Alexander and the candidate as follows.

- Contributions directly attributed to the candidate: Proof of properties of chosen qubit Wigner representation. Proof that CSS circuits are completely CSS-preserving and therefore stochastically represented. Entropic majorization conditions on completely CSS-preserving magic distillation protocols for qubits. Decomposition of CSS circuits into code projection protocols. Proof that CSS code projection protocols are stochastically represented. Proof that CSS/stabilizer state prepared after a failed run of a CSS/stabilizer code projection protocol can be freely chosen. Trade-off relations on parameters governing the performance of CSS/stabilizer code projection protocols (including upper and lower bounds on code length). Interpretation of aforementioned upper bounds on code length. Proof that similar upper bounds arise from the Data-Processing Inequality.

- Contributions directly attributable to Rhea Alexander: Proof of properties of chosen qubit Wigner representation. Using relative majorization of Wigner distributions to generate magic monotones. Entropic majorization conditions on completely CSS-preserving

magic distillation protocols for qubits. Properties of aforementioned entropic conditions. Choice of reference process for code projection protocols. Trade-off relations on parameters governing the performance of CSS/stabilizer code projection protocols (including upper and lower bounds on code length). Comparison of performance of bounds on code length and acceptance probability bound against those due to state of the art magic monotones. Comparison of upper bounds against those produced by the Data-Processing Inequality.

All authors contributed to the writing of the manuscript.

Chapters 5 and 6 are based on *Infinitesimal reference frames suffice to determine the asymmetry properties of a quantum system* by Rhea Alexander, **Si Gvirtz-Chen** and David Jennings, New J. Phys. **24** 053023 (10 May 2022). David Jennings supervised the project and provided high-level advice and guidance. The bulk of the research was jointly undertaken by Rhea Alexander and the candidate as follows.

- Contributions directly attributable to the candidate: results on basic reference frame redundancies. Result on sufficient surfaces of reference frames. Single minimality condition for asymmetry theory and conical behaviour of conditional min-entropy at the maximally mixed reference frame. Computation of conditional min-entropy for time-symmetric state transitions in a non-degenerate qubit. Results on sufficient depolarisation conditions on modes of asymmetry for G-covariant state conversion. Performance comparison of aforementioned sufficient depolarisation conditions against complete conditions for unitarily covariant state transitions in a qubit.
- Contributions directly attributable to Rhea Alexander: Simulations of states deemed accessible by individual reference frames for time-symmetric transition in a qubit. Result on epsilon-smoothing sufficient conditions for G-covariant state conversion. Results on sufficient depolarisation conditions on modes of asymmetry for G-covariant state conversion. Performance comparison of aforementioned sufficient depolarisation conditions against complete conditions for time-symmetric transitions in a non-degenerate qubit.

All authors contributed to the writing of the manuscript.

Acknowledgements

First and foremost, my deepest gratitude goes to my supervisor, David Jennings. Thank you for teaching me how to push an idea as far as it can go, how to avoid getting lost in the thickets of theory-building (where mistakes often lurk) by building intuition on the simplest non-trivial examples, and how much there is to learn even from ideas that do not work out. Above all, thank you for your steadfast encouragement, your unwavering patience, your immense generosity. I know how incredibly lucky I am to have been your student.

My second biggest thanks goes to my main collaborator on the two papers forming the research constituting this thesis, Rhea Alexander. I am really grateful to have worked with her and benefited from her eye for detail and rapid mastery of new mathematical machinery. I would also like to thank my collaborator Nikolaos Koukoulekidis for wonderful guidance on extending his statistical framework for quantum computing by state injection on qudits of odd prime dimensions to qubits in the first paper constituting this thesis.

Thank you to everyone in the theory group at Leeds, especially the head of group, Almut Beige, for fostering such a friendly and cheerful atmosphere. A round of drinks goes out to Tanmay, Cristi, Dan, Jean, Gabriel, Aiden and Lucy for working with me on the tutoring scheme, and a second round to Lucy for running the scheme after I left on placement. Many thanks go to the Bell Burnell Graduate Scholarship Fund for taking a chance on me and supporting this PhD.

On a personal level, I would like to thank Matt, Nick and Rhea for being my "academic siblings", and in particular to Matt and Rhea for being my "nearest and dearest" (as Rhea put it) while I lived in Leeds. I could not have made it through without your friendship and kindness. Thanks as well to my housemates Jean, Gabriel, Sam and Wura for making Leeds my favourite place to live.

Last but never least, I thank mum, dad and Damian for their love.

Abstract

This thesis describes, quantifies and applies quantum information-processing resources that go beyond observables and generalised measurements (POVMs), and must instead be defined as monotones on the pre-order established within a resource theory. The research is divided into two parts, examining monotones expressed in terms of generalised quantum entropies for the resource theories of magic and asymmetry respectively. These resources display counter-intuitive behaviour signalling the breakdown of classical descriptions for quantum computing (in terms of statistical mechanics) and symmetry-constrained open system dynamics (in terms of Noether's Theorem).

Part I of the thesis concerns magic distillation. Previous work has cast universal fault-tolerant quantum computing by magic state injection for odd-dimensional systems within a phase space setting wherein distillable magic states acquire negative α-Rényi entropies. We extend this statistical mechanics framework to the technologically important qubit case, from which we derive fundamental trade-off relations on parameters governing the performance of an elementary family of protocols that project onto CSS codes. These trade-off relations are tuned to physics specific to code projection protocols and can outperform previous monotone bounds on protocol parameters in regimes of practical interest.

Part II of this thesis concerns symmetry-constrained open system dynamics. Recently, a complete but infinite set of entropic monotones was found for the resource theory of asymmetry, given in terms of correlations with every state on a spontaneously emergent quantum reference system. We show that one can restrict to reference frames forming any surface enclosing the maximally mixed state, which implied that the possibility of state transition under symmetry-constrained general quantum channels can be determined by a single entropic minimality condition at the maximally mixed state. Building on this analysis, we provide simple,

closed conditions on the minimal depolarization needed to make a quantum state accessible under channels covariant with symmetry groups whose multidimensional representations are multiplicity-free.

Contents

List of Figures

Part I

Preliminaries

Chapter 1

Introduction

Historically, quantum mechanics seemed to present nothing but obstacles to information processing. Examples include the restriction posed by the uncertainty principle on the precision with which canonically conjugate variables may be encoded in a quantum state [1], as well as the disconcerting discovery, expressed by the No-Cloning Theorem [2], that one cannot copy an arbitrary unknown quantum state.

The advent of quantum computing has radically overturned this assessment. From algorithms that solve important problems such as prime factorisation [3] nearly exponentially faster than their classical counterparts, to protocols for fundamental cryptographic tasks such as key distribution that, for the first time, offer security on an information-theoretic basis rather than on assumptions about the technological limitations of attackers [4], we have seen that quantum mechanics in fact holds new possibilities for information processing beyond what can be achieved within classical physics.

One particularly fruitful approach to studying and characterising how the quantum properties of a physical system may be harnessed for information-processing tasks is the *resource theory framework* [5]. The basic idea is to restrict quantum information processing to a permitted or "free" subset of quantum channels, and thereby designate access to any channel outside of this subset as a resource. While such restrictions frequently originate in current technological limitations on which quantum operations are easy or even possible to perform, resource theories often came to be valued for isolating the physical ingredients essential for a given task in quantum information processing. For instance, in the first quantum resource theory ever

formulated, the resource theory of entanglement, the "free" operations are local operations and classical communication (LOCC) [6]. Restricting to such operations not only allows us to capture the practical reality that LOCC are far less challenging to carry out than the faithful transmission of quantum systems across large distances, but also isolated entanglement as the sole necessary ingredient for making such transmissions possible. Another example comes from the resource theory of asymmetry, which has shown that superselection rules on observables conserved under a symmetry group can be lifted by sharing a quantum reference frame for that symmetry group [7].

The set of free operations defining a quantum resource theory often also coincides with some notion of classicality. For example, in the field of quantum computation, the stabilizer subtheory, which can be implemented relatively easily in a fault-tolerant way [8, 9, 10], can also be simulated efficiently on a classical computer [11, 12, 13]. Non-stabilizer ("magic") states or channels can consequently be regarded as holding the resources responsible for quantum computational speedup [14]. We therefore see that the resource theory framework is highly versatile and lends itself both to practical questions probing the consequences of limited experimental capabilities and foundational questions on the sources of quantum advantage.

At a more technical level, resource theories provide a rigorous basis for identifying how many different types of resources are relevant for a given information-processing task, as well as quantifying how much of these resources are held by a given quantum state (or channel). From an operational perspective, one state possesses at least as much of a given resource as another state if it is possible to transform the former into the latter without bringing in more of the resource – i.e. via the free operations of the resource theory. Convertibility under free operations thus establishes a "ranking" – formally, a *pre-order* – on the set of quantum states, which is captured by a set of *resource monotones* that cannot increase under free operations.

Perhaps unsurprisingly, resource monotones are often expressed in terms of *entropies* – classical thermodynamics itself can be regarded as a resource theory of operations restricted by compliance with the laws of thermodynamics, within which thermodynamic entropy uniquely determines the accessibility of any equilibrium state from another (of the same energy). The Shannon entropy [15] used for quantifying classical information was developed with thermodynamic entropy as its inspiration. In recent years, many *generalised quantum entropies* have been devised in place of the Shannon entropy to characterise the irreversibility of state con-

versions within a quantum resource theory [16, 17], which typically possess far more complex structures than classical thermodynamics.

1.1 Thesis overview: motivation and contributions

This thesis is concerned with describing, quantifying and using information-theoretic resources of quantum systems that go beyond what can be captured via observables or generalised measurements (POVMs), and must instead be defined by measures capturing the pre-order established within a quantum resource theory. The research is divided into two parts, which examine measures expressed in terms of generalised quantum entropies for the resource theories of magic and asymmetry respectively. These resources display counterintuitive behaviour that signals the breakdown of classical descriptions for quantum computing (in terms of statistical mechanics) and dynamical properties induced by a symmetry constraint (in terms of Noether's Theorem) respectively. The two parts of this thesis demonstrate the wide scope of the resource-theoretic framework in characterising quantum resources, as one part studies abstract computational resources that may be held by any physical system in a fixed practical application (magic distillation), whereas the other part studies resources that must be held by specific physical systems (capable of encoding given symmetry groups) from a general foundational perspective.

We begin in Chapter 2 with a review of tools from quantum information science that will be used extensively throughout this thesis. We give particular focus to introducing the notion of *quantum resource theories*, which provides the overarching framework within which the research constituting this thesis takes place.

1.1.1 Part I: General entropic constraints on Calderbank-Shor-Steane (CSS) codes within magic distillation protocols

Quantum Computing by State Injection (QCSI) has emerged as one of the most promising approaches for achieving universal fault-tolerant quantum computation, wherein so-called *magic states* are resources that must be distilled by fault-tolerant *stabilizer circuits* in order to promote those circuits to universality [18]. First experimentally demonstrated in December 2024 [19] after long theoretical investigation, almost all known magic distillation protocols [18, 20, 21, 22, 23] are based on a subclass of quantum error-correcting codes known as *Calderbank-Shor-Steane*

(CSS) codes [24, 25]. Though significant progress has been made to reduce the overhead of such protocols [26], magic distillation is still expected to dominate the resource cost of QCSI. Understanding what fundamental trade-offs distillation must make between reducing resource costs and improving fidelity, especially in ways that could be tailored to the specific physics of protocols used, is thus of vital practical interest, and may be instructive for designing protocols with optimised parameters.

Recent work [27] developed a statistical mechanics framework for analysing magic distillation protocols by extending tools from majorization theory to phase-space representations of magic states. However, this framework was restricted to quantum systems of odd prime Hilbert space dimension, and had not yet been applied to concrete protocols for magic distillation.

In Part I of this thesis, we extend the framework of Ref. [27] to the experimentally important case of qubit systems by focussing on magic distillation in the completely CSS-preserving subset of stabilizer circuits, which we show to be similarly capable of QCSI. By exploiting this extended framework, we derive fundamental trade-off relations in parameters governing the performance of an elementary family of magic distillation protocols that project onto CSS codes. These trade-off relations yield lower bounds on code length capable of out-performing state-of-the-art lower bounds due to magic monotones in some regimes of practical interest, as well as a novel cut-off result implying, for fixed target error rate and acceptance probability, that one needs only consider CSS codes below a threshold number of qubits. These constraints are not due simply to the data-processing inequality but rely explicitly on the stochastic representation of such protocols in phase space.

1.1.2 Part II: Infinitesimal reference frames suffice to determine the asymmetry properties of a quantum system

Identifying constraints imposed by symmetry on the evolution of a system has broad applications throughout physics – when exact solutions to the laws of motion become too complex, one can often make non-trivial inferences by appealing to symmetry principles. While Noether's Theorem allows us infer conserved quantities from symmetries in closed quantum system dynamics [28], this inference is no longer valid in open quantum system dynamics [29, 30]. How symmetry principles constrain the *general* evolution of a quantum system via quantum channels therefore became a crucial question. In response, the resource theory of asym-

metry was developed [30] with the aim of understanding which states can be converted into which under channels that are covariant with respect to a symmetry group G. Most works have addressed this question by identifying *asymmetry monotones* [30, 31, 32, 33, 34, 35] that allows us to quantify the degree to which a quantum system encodes coordinates of a symmetry group; such monotones must decrease under G-covariant channels and thereby provide necessary conditions on general state evolution under a symmetry principle.

Recent work [36] identified the first complete set of asymmetry monotones in terms of correlations relative to quantum reference frames assessed by the single-shot conditional min-entropy. However, because these monotones must be evaluated at infinitely many reference frames, they are difficult to use in practice and their physical implications unclear.

In Part II of this thesis, we show that the set of asymmetry monotones identified in Ref. [36] has extensive redundancy, and that one can restrict to reference frames forming any closed surface that encloses the maximally mixed state. This result enables us to obtain a single necessary and sufficient entropic condition at the maximally mixed state for G-covariant state conversion. Though evaluating this condition in its current form poses signficant technical difficulties, it can still be interpretationally interesting because it shows that one does not need reference frames that perfectly encode a symmetry group to characterise the asymmetry properties of a quantum system – in fact, reference frames that are nearly completely degraded suffice. Building on this analysis, we provide simple, closed conditions to estimate the minimal depolarisation needed to make a given quantum state accessible under channels covariant with an arbitrary symmetry group.

Chapter 2

Quantum Resource Theories (QRTs)

2.1 Quantum information preliminaries

In this section, we rapidly review some basic concepts from quantum information theory that are relevant to quantum resource theories, and introduce notation that will be used throughout the rest of this thesis. Classic exposition of this introductory material can be found in e.g. Ref. [1] or Ref. [37]. Throughout this thesis, we work in units of $\hbar = 1$ for convenience.

A *quantum system* is any collection of physical degrees of freedom whose behaviour allow for a closed and consistent description using quantum theory [38]. Every quantum system S is associated to a *Hilbert space* $\mathcal{H}_S$, and we will denote the Hilbert space of a d-dimensional quantum system, that is, a quantum system in which at most d states are simultaneously distinguishable, by $\mathcal{H}_d$. Furthermore, we will denote the set of *bounded operators* on $\mathcal{H}_S$ by $\mathcal{B}(S)$; in the special case where S is a system of n qubits, this notation will be adapted to $\mathcal{B}(n)$. The *quantum states* (i.e. density matrices) for system S form the subset in $\mathcal{B}(S)$ of positive-semidefinite matrices with trace 1, which we denote by $\mathcal{D}(S)$.

Any physically possible evolution of a quantum system must take density matrices to density matrices in a way that respects statistical mixing, even when only acting on a subsystem of some larger system. The most general evolution of system S to system S' is therefore given by the set of *quantum channels* from S to S', which are defined as linear maps from $\mathcal{B}(S)$ to $\mathcal{B}(S')$ that are completely positive and trace-preserving (CPTP). We will denote the identity or "do nothing" channel on quantum system S by id_S.

It is worth noting at this juncture that quantum states (or, speaking more precisely, preparations of quantum states) can also be represented as quantum channels if one introduces the complex numbers $\mathbb{C}$ as the Hilbert space of the unique 0-dimensional quantum system, whose sole "state" is the number 1. The state ρ_S of system S can then be represented as the quantum channel $\Phi : \mathbb{C} \to \mathcal{B}(S)$ such that $\Phi(1) = \rho_S$; all channels of the form Φ are therefore referred to as *state preparation channels*.

2.1.1 Three representations of quantum channels

We now review three ways of representing any quantum channel that are frequently used across quantum information science. The existence of any of these representations is necessary and sufficient for a map to be CPTP, and thus constitute a valid quantum channel.

2.1.1.1 The Stinespring dilation

Possibly the most intuitive way of representing a quantum channel is the *Stinespring dilation*. Given any quantum channel $\Phi_{A\to B}$ from an input system A to an output system B, the Stinespring theorem [39] states that there always exists some ancillary system E and unitary operation U_{AE} on the input and ancillary system such that

$$\Phi_{A\to B}(\rho_A) = \mathrm{Tr}_{E'}\left[U_{AE}(\rho_A \otimes |0\rangle\langle 0|_E)U_{AE}^\dagger\right], \tag{2.1}$$

where we highlight that the discarded system E' need not be the same as the initial ancillary system E, and that the choice of U_{AE} is in general not unique. The Stinespring dilation can be interpreted as saying that every possible evolution of a quantum system is equivalent to a unitary evolution of the system and its environment, where some degrees of freedom in the joint system subsequently become inaccessible.

2.1.1.2 The Kraus representation

Probably most widely encountered method of representing a quantum channel $\Phi_{S\to S'}$ from input system S to output system S' is the *Kraus representation*, which decomposes Φ into a set

of linear operators $\{K_j : \mathcal{H}_S \to \mathcal{H}_{S'}\}$ satisfying $\sum_j K_j^\dagger K_j = \mathbb{1}_A$ such that

$$\Phi_{S\to S'}(\rho_S) = \sum_j K_j(\rho_S)K_j^\dagger. \tag{2.2}$$

As is the case with U_{AE} in the Stinespring dilation, the choice of $\{K_j\}$ is also not unique and is only defined up to unitary mixing [1].

2.1.1.3 The Choi-Jamiolkowski representation

This representation expresses the Choi-Jamiolkowski isomorphism [40, 41] between quantum states and channels. Let us define the *maximally entangled state* between two copies of an input system S with Hilbert space dimension d_S in an orthonormal basis $\{|i\rangle_S\}$ as

$$|\phi^+\rangle_{SS} := \frac{1}{\sqrt{d_S}} \sum_{i=1}^{d_S} |i\rangle_S |i\rangle_S . \tag{2.3}$$

Given any quantum channel $\Phi_{S\to S'}$ from S to an output system S', we can construct the associated *Choi state* $\mathcal{J}(\Phi_{S\to S'}) \in \mathcal{B}(SS')$ representing the channel as the following state on the bipartite system SS',

$$\mathcal{J}(\Phi_{S\to S'}) := \mathsf{id}_S \otimes \Phi_{S\to S'}(|\phi^+\rangle\langle\phi^+|_{SS}), \tag{2.4}$$

which further has the property of being maximally mixed on the input system, i.e.

$$\mathrm{Tr}_{S'}[\mathcal{J}(\Phi_{S\to S'})] = \frac{I_S}{d_S} \tag{2.5}$$

The Choi-Jamiolkowski isomorphism states that $\mathcal{J}$ constitutes a *bijection* between CPTP maps from S to S' and states on SS' that are maximally mixed on S. Concretely, given any bipartite state $J_{SS'} \in \mathcal{B}(SS')$ on the joint system SS' satisfying Eq. (2.5), one can construct a quantum channel $\mathcal{J}^{-1}(J_{SS'})$ from S to S' that acts on a state ρ_S of S as

$$\mathcal{J}^{-1}(J_{SS'})[\rho_S] := d_S \,\mathrm{Tr}_S(J_{SS'}[\rho_S]^T \otimes I_{S'}), \tag{2.6}$$

where transposition is performed relative to the basis $\{|i\rangle_S\}$.

2.1.2 Inner products, norms and distance measures between quantum states

The set of bounded linear operators $\mathcal{B}(\mathcal{H})$ on a finite-dimensional Hilbert space $\mathcal{H}$ can constitute a (complex) linear vector space, and can itself be turned into a Hilbert space upon being equipped with the *Hilbert-Schmidt inner product*, which is defined as

$$\langle A, B\rangle \coloneqq \mathrm{Tr}\Big[A^\dagger B\Big]. \tag{2.7}$$

The norm induced by the Hilbert-Schmidt inner product on $\mathcal{B}(\mathcal{H})$ is also known as the *trace norm* or Schatten-1 norm, which has been extended into a whole family of operator norms used extensively throughout this thesis: the *Schatten-p norms* defined as

$$\|X\|_p \coloneqq \left[\mathrm{Tr}\Big(X^\dagger X\Big)^{\frac{p}{2}}\right]^{\frac{1}{p}} \tag{2.8}$$

for $p \in [1, \infty]$ and bounded operator X. The Schatten-2 norm is also known as the *Frobenius norm*. Given a Hermitian input, the Schatten-∞ norm returns the largest absolute value of its eigenvalues.

We can construct *distance measures* between two quantum states ρ and σ of the same system to quantify how far they are from being identical. This thesis uses two of the most important distance measures in quantum information science, the *trace distance*

$$D(\rho||\sigma) \coloneqq \frac{1}{2}\|\rho - \sigma\|_1, \tag{2.9}$$

and the *infidelity*,

$$1 - F(\rho, \sigma) \text{ where } F(\rho, \sigma) \coloneqq \left\|\sqrt{\rho}\sqrt{\sigma}\right\|_1^2. \tag{2.10}$$

When σ is a pure state given by $|\psi\rangle\langle\psi|$, its fidelity with respect to ρ becomes simplified to $F(\rho, |\psi\rangle\langle\psi|) = \langle\psi|\,\rho\,|\psi\rangle$. In accordance with the definition of distance measures, both the trace distance the infidelity becomes 0 if and only if $\rho = \sigma$.

2.1.3 Universal quantum computation

A quantum computer on qudits of the same dimension d is *universal* when it is able to execute any quantum channel between such qudits with arbitrary precision. From the Stinespring

dilation of quantum channels, we see that universality can be reduced to the ability to discard any qudit, to initialise any qudit in a *computational basis state* that is the eigenstate of the Pauli Z operator, and to perform any unitary operation on qudits with arbitrary precision. In quantum information, unitary operations are usually referred to as *gates* following terminology from classical computer science. *Universal quantum computation* is then typically defined as the ability to approximate any unitary channel using a number of elementary gates (drawn from a finite set) that scales polynomially with the desired accuracy [1]. A classic example of a universal gateset for qubits is the CNOT-gate, the Hadamard gate H, the phase-gate S and the T-gate, which are reviewed in section 3.2.3 of Chapter 3.

2.2 The QRT framework

In this section, we briefly review the QRT framework. Despite its name, a quantum resource theory is not a scientific theory in the sense of falsifiability, but rather a framework for casting physical properties as "resources" relative to a set of quantum channels. This framework allowed researchers a method to precisely quantify fundamental quantum properties that could not be expressed as Hermitian observables or even POVMs, such as entanglement [42, 6], coherence [43, 33, 44, 45], thermodynamics [46, 47, 48], non-Gaussianity [49, 50], magic [14, 51] and many more [5]. Outstanding and much more thorough exposition of the QRT formalism can be found at e.g. Ref. [52].

Consider a quantum system given to an agent or distributed amongst a group of agents. A QRT essentially models what the parties can physically accomplish using this system given restrictions imposed by technological or experimental limitations, the rules of some game, or even the laws of physics [5]. What operations the agents can still perform is then described by a (typically much smaller) subset $\mathcal{R}$ of all the quantum channels by which this system could evolve, which forms the basis for defining a QRT.

Definition 1 (Quantum Resource Theory). *A quantum resource theory is defined by a subset $\mathcal{R}$ of quantum channels, called free or permitted, with the following properties:*

(R1) *contains all identity channels, i.e. given any quantum system A, $\mathsf{id}_A \in \mathcal{R}$.*

(R2) *is closed under composition, i.e. given any three quantum systems A, B and C, if $\Phi_{A\to B}$ is a quantum channel from A to B and $\Phi_{B\to C}$ is a quantum channel from B to C such*

that $\Phi_{A\to B} \in \mathcal{R}$ *and* $\Phi_{B\to C} \in \mathcal{R}$*, then* $\Phi_{A\to B} \circ \Phi_{B\to C} \in \mathcal{R}$*.*

The *free states* of a QRT are the subset $\mathcal{F}$ of quantum states produced by state preparation channels in $\mathcal{R}$, i.e. given any quantum state ρ_A of system A, we have $\rho_A \in \mathcal{F}$ if and only if there exists a quantum channel $\Phi_A : \mathbb{C} \to \mathcal{B}(A)$ such that $\Phi_A \in \mathcal{F}$ and $\Phi_A(1) = \rho_A$. Any state or channel that is *not* free is called a *resource* state or channel.

It is worth taking a few moments to unpack the physical interpretation and consequences of Definition 1. Property **(R1)** simply says that doing nothing is always permitted. Property **(R2)** then formalises what it means for quantum channels to be permitted: that they can be performed in whatever order one wishes and as often as one wishes. The major consequence of Property **(R2)** is

The Golden Rule of QRTs

Free operations cannot convert free states to resource states.

More formally, if ρ_A is a free state of quantum system A prepared by some state preparation channel $\Phi_A \in \mathcal{R}$, then given any channel $\Phi_{A\to B} \in \mathcal{R}$ from system A to system B, Property **(R2)** implies that $\Phi_{A\to B} \circ \Phi_A \in \mathcal{R}$, which is equivalent to $\Phi_{A\to B}(\rho_A) \in \mathcal{F}$. Therefore, in the absence of any resource states or channels, a QRT leaves agents "stuck" in a *subtheory* – i.e. closed subset of channels – of quantum theory that consists exclusively of its free operations.

Given a set of free operations $\mathcal{R}$ alongside the free states $\mathcal{F}$ it induces, we will use $\mathcal{R}(A \to B)$ to denote the set of free operations from input system A to output system B, and use $\mathcal{F}_A$ to denote the set of free states on system A.

2.2.1 Resource Monotone

The golden rule of QRTs does not mean that resource states are totally excluded. On the contrary, if the quantum system originally given to the parties happens to be a resource state, it could potentially be used to lift the restrictions on the allowed operations. More formally, given any two physical systems A and B, access to a *resource state* ρ_X on system X allows one to perform channels from A to B of the form $\Phi_{A\to B}(\cdot) := \Phi_{AX\to B}((\cdot)_A \otimes \rho_X)$ where $\Phi_{AX\to B} \in \mathcal{R}$, which may not necessarily also lie in $\mathcal{R}$. The most celebrated example of using a resource state to circumvent operational restrictions come from the QRT of entanglement, where the free op-

erations are restricted to local operations and classical communication (LOCC), yet sharing a Bell state allows one agent to transmit a qubit perfectly to another.

However, the Bell state is "used up" by the perfect transmission of a qubit, in the sense that after the transmission, the Bell state is degraded to a separable state [6]. One basic question that QRTs were developed to answer is therefore: how quickly must a resource state be used up, i.e. what free operations can it go through before ending up as a free state? Addressing this question naturally leads us to explore the structure of state conversion within a QRT.

Given some quantum states ρ and σ, if there exists a free operation $\Phi \in \mathcal{R}$ that can convert σ to ρ as $\Phi(\sigma) = \rho$, then we write $\rho \succ_{\mathcal{R}} \sigma$. Generally, the relation $\succ_{\mathcal{R}}$ is only a *pre-order*, because there can exist ρ and σ such that neither state can be converted to the other via free operations, i.e. we have $\rho \not\succ_{\mathcal{R}} \sigma$ and $\sigma \not\succ_{\mathcal{R}} \sigma$, and moreover $\rho \succ_{\mathcal{R}} \sigma$ and $\sigma \succ_{\mathcal{R}} \rho$ does not necessarily imply $\rho = \sigma$. This pre-order is captured by *resource monotone*:

Definition 2 (Resource monotone)**.** *Given a resource theory with free operations $\mathcal{R}$, a resource monotone $M : \cup_S \mathcal{D}(S) \to \mathbb{R}^+$ is a function that assigns a non-negative real number to every state of every quantum system S that cannot be increased by free operations, i.e.*

$$\rho \succ_{\mathcal{R}} \sigma \implies M(\rho) \geq M(\sigma). \tag{2.11}$$

Informally, resource monotones allow us to quantify "how much" of a resource a state possesses, as well as how many different "flavours" of resource are required by a particular information-processing task. Though this is not the case with the monotones studied in this thesis, we note that typically resource monotones are also expected to be *faithful*, i.e. to vanish on free states.

Part II

General Entropic Constraints on Calderbank-Shor-Steane (CSS) Codes within Magic Distillation Protocols

Chapter 3

A Statistical Mechanics Framework for Qubit Magic Distillation

3.1 Introduction

3.1.1 Quantum computation by state injection (QCSI)

The components of any real quantum computer are invariably noisy. To prevent errors introduced by noise from multiplying and spreading until computation becomes unreliable, a practical quantum computer must have a *fault-tolerant architecture*, which rests upon two ingredients. First and foremost, the quantum computer must be able to protect information stored in its memory using *quantum error-correcting codes* (QECCs) [53]. Similar to its classical counterpart, a QECC typically encodes any state on a single qubit redundantly across a block of *several* qubits, which allows the original state to be reconstructed even after a few qubits in the block are disturbed by noise [54]. To maintain this protection as information is processed, the quantum computer then encodes a universal set of channels to operate *directly* on encoded states so they need never be decoded, and periodically performs error correction on each encoded block of qubits [54]. In August 2024, Google provided the first experimental demonstration of below-threshold quantum error correction using surface code constructions [9, 10, 55], which exponentially suppressed the logical relative to the physical error rate by a factor 2 [56].

However, error correction alone is insufficient for fault tolerance, because encoded operations can cause errors to *propagate* from one encoded block to another, and the correction procedure

may itself introduce errors as it is carried out using the same noisy components as the rest of the computer. The second ingredient of fault-tolerant architecture is designing each encoded operation such that a failure anywhere in the procedure only propagates to a small number of qubits in each encoded block, so error-correction can remain effective at removing them [1]. To this end, it is highly desirable if encoded operations can be *transversal*, i.e. act on each qubit in an encoded block independently[1], as this ensures a failure can produce at most one qubit error per encoded block [54]. Unfortunately, hopes that transversal encoding can directly provide universal fault-tolerant quantum computing were dashed by the Eastin-Knill Theorem [57], which proved that no universal gateset can be transversally encoded.

Quantum Computation by State Injection (QCSI) has emerged as one of the most promising frameworks for circumventing the Eastin-Knill Theorem. Its basic idea is restricting computation to a *subtheory* that can be encoded transversally at the expense of only offering classical computational power, and then promoting this subtheory to universal quantum computation by injecting special, pre-prepared ancillary states known as *magic states*. In the first and still most popular concrete scheme for QCSI, the *Bravyi-Kitaev model*, computation is restricted to *stabilizer circuits*, which can be transversally encoded via surface code constructions but were shown in the Gottesman-Knill theorem to be efficiently simulable on a classical probabilistic computer [12, 13]. Stabilizer circuits can be promoted to universal quantum computation through the injection of *any* pure state that they cannot prepare [58, 59]. We see that a QCSI scheme is naturally structured as a *resource theory of magic*, wherein the transversally-encoded subtheory constitutes the free operations, while states (and channels) outside this subtheory are defined as magic and contain the resources necessary for quantum computation. The split between free and resourceful operations in a QCSI scheme approximates[2] the classical-quantum boundary in computational power.

Because magic states are prepared outside the fault-tolerant subtheory to which QCSI computation is restricted, they suffer error rates that are too high for them to be used directly. QCSI therefore requires an initial stage of *magic state distillation*, which converts many noisy copies

[1] Formally, a quantum operation Φ on m blocks of n qubits is transversal if there exists a set of m-qubit operations $\{\Phi_i\}_{i=1,\dots n}$ such that $\Phi = \bigotimes_{i=1}^{n} \Phi_i$.

[2] Whether a subtheory can be transversally encoded is not straightforwardly related to whether it only provides classical computational power. Taking the Bravyi-Kitaev model as an example, the subtheory of operations with a non-negative Gross's Wigner representation (reviewed in the next section) is strictly larger than stabilizer circuits, but can still be efficiently simulated by a classical probabilistic computer [60]. Resource theories of magic with a more foundational outlook than those reformulating QCSI schemes have been constructed by taking such larger sets of free operations [14, 61].

of a desired magic state into fewer copies at higher fidelity using only operations from the fault-tolerant subtheory. Magic distillation was experimentally demonstrated for the first time in December 2024 [19].

3.1.2 Phase-space formulations of QCSI

On qudits of odd prime dimension, the Bravyi-Kitaev model has a beautiful phase-space formulation that has powerful applications for resource theories of magic and other sorts of non-classicality. Invented by David Gross [60] as an extension of a similar representation for continuous variable systems in quantum optics by Eugene Wigner [62], this formalism provides the closest quantum analogue to the classical conception of dynamical systems as probability distributions evolving over phase space, with the exception that some quantum states are represented by *quasiprobability distributions* that have *negative* values. Nevertheless, under Gross's Wigner representation, every state that can be prepared by a stabilizer circuit becomes a legitimate probability distribution over phase space; more generally, every stabilizer circuit becomes a stochastic map between phase spaces. By contrast, distillable magic states are represented by quasiprobability distributions that always contain negative entries [63]. *Negativity* under Gross's Wigner representation thus emerges as a necessary (and, for pure states, sufficient) resource for quantum computational speedup.

Gross's Wigner representation can serve as a *realist model* for stabilizer circuits in odd prime dimension, i.e. we can interpret points in the phase space of a single qudit as the "true" states of that qudit, such that preparing a stabilizer state would, "in reality", place each qubit in one of these states, which then evolve under stabilizer operations that, "in reality", sends these states to others. This is partly because Gross's Wigner representation becomes stochastic for stabilizer circuits, but also because it respects the sequential and parallel composition of processes – that is, if a channel Φ_1 operates after, or on a separate system from, another channel Φ_2, Gross's Wigner representation W of Φ_1 also acts after, or on a separate phase space from, the representation of Φ_2, i.e. $W_{\Phi_1 \otimes \Phi_2} = W_{\Phi_1} \otimes W_{\Phi_2}$ and $W_{\Phi_2 \circ \Phi_1} = W_{\Phi_1} \circ W_{\Phi_2}$.

Gross's Wigner representation goes further by being a realist model for stabilizer circuits that satisfies a form of classicality known as *generalised noncontextuality* [64] – it explains operational equivalence between stabilizer circuits as due to these procedures being *identical* on the "true" states of systems (e.g. no matter which mixtures were used to prepare a maximally mixed state,

the probability distribution over phase space is the same). In fact, it is impossible to construct a generalised noncontextual model incorporating both stabilizer circuits and any distillable magic state [65], making negativity in Gross's Wigner representation a resource under this more philosophical notion of nonclassicality, and forms a pleasing connection with the more practical notion of nonclassicality from computational speedup.

3.1.3 Finding statistical mechanics constraints on magic distillation

Recent work [27] has developed a statistical mechanics framework for analysing magic distillation in qudits of odd prime dimension by taking key insights from a rich literature of majorization theory and applying them to Gross's Wigner representation. The Bravyi-Kitaev model is transformed by Gross's Wigner representation into the stochastic updating of quasiprobability distributions, but when computation is restricted to stabilizer circuits, the resultant classical stochastic model can be studied using entropic theory and in particular *relative majorization* [66, 47, 67, 68, 69]. Ref. [27] extended majorization to the quasiprobability distributions representing magic states, and found that a dense subset of α-Rényi entropies H_α defined on phase space remains capable of characterising their pre-order under stochastic processing, leading to fundamental constraints on magic distillation protocols that can be tailored to their particular physics[3].

Since most quantum algorithms are formulated for systems of qubits, we would like to extend this framework to qubits with the aim of identifying trade-off relations between physical parameters characterising distillation protocols. There are, however, many well-known obstacles in constructing valid Wigner representations for qubits (related to the fact that 2^{-1} does not exist modulo 2 [71]) – in particular, a recent no-go theorem has shown that it is impossible to construct a quasiprobability representation of the qubit Bravyi-Kitaev model that respects the sequential and parallel composition of processes while representing all stabilizer circuits stochastically [65]. As we will see, relaxing the former property significantly complicates the extraction of trade-off relations, so we instead make progress by identifying subsets of qubit stabilizer operations that can be represented stochastically, while nevertheless remaining capable of universal quantum computation via magic state injection.

[3] As a point of clarification, these α-Rényi entropies are distinguished from the stabilizer Rényi entropies introduced in Ref. [70] by the fact that they are defined on Gross's Wigner representations of a quantum state, rather than on the quantum state itself.

3.1.4 Chapter summary

The rest of this chapter proceeds as follows. We begin with a review of stabilizer circuits in section 3.2, which forms the main technical background for this chapter. In sections 3.3 and 3.4, we develop the properties of a qubit Wigner representation introduced by Ref. [72] and extend it to arbitrary quantum channels, which we verify as respecting the parallel and sequential composition of processes. Drawing on results from Ref. [73], we show in Section 3.4 that a stabilizer circuit is stochastically represented by our chosen qubit Wigner representation if it is completely CSS-preserving (that is, preserves a subset of stabilizer states known as CSS states even when acting on a subsystem), and that such circuits are capable of QCSI. In section 3.5, we extend the statistical mechanics framework for universal quantum computing developed in Ref. [27] from qudits of odd prime dimension to qubits in the completely CSS-preserving setting, and show how this constitutes a pleasing physical picture of universal qubit quantum computing as statistical mechanics with negative entropies. All arithmetic is modulo 2 unless stated otherwise.

3.2 Technical background: stabilizer circuits

We first provide a brief review of stabilizer circuits. While the core presentation will focus on qubits, we will also sketch how the subtheory is to be extended to qudits of odd prime dimension d. Brilliant and much more comprehensive exposition of the stabilizer formalism by its inventor, Daniel Gottesman, can be found at Ref. [54].

3.2.1 The Pauli group and its stabilizer subgroups

The formalism of qubit stabilizer circuits is constructed around Pauli groups. Formally, the *n-qubit Pauli group* P_n consists of all n-fold tensor products of Pauli matrices with phases of ± 1 or $\pm i$,

$$\mathsf{P}_n \coloneqq \{i^k\{\mathbb{1}, X, Y, Z\}^{\otimes n} | k \in \mathbb{Z}_4\}, \tag{3.1}$$

where X, Y and Z are *Pauli matrices* represented in the computational basis as

$$X \coloneqq \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \; Y \coloneqq \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \; Z \coloneqq \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}. \tag{3.2}$$

It is immediately apparent that the n-qubit Pauli group is simply the n-fold tensor product of the single-qubit Pauli group – formally, $\mathsf{P}_n = \bigotimes_{i=1}^{n} \mathsf{P}_1$. We further highlight the following properties of Pauli operators, which arise because X, Y and Z are unitary, have eigenvalues ± 1 and anti-commute with each other.

Lemma 3. *A Pauli operator is unitary and either has eigenvalues ± 1, in which case it constitutes a valid observable, or eigenvalues $\pm i$. Two Pauli operators either commute or anti-commute.*

A state vector $|\psi\rangle$ in the Hilbert space $\mathcal{H}_2^n \coloneqq \mathcal{H}_2^{\otimes n}$ of n qubits is said to be *stabilized* by an n-qubit Pauli operator P if $|\psi\rangle$ is an eigenstate of P of eigenvalue 1, i.e.

$$P\,|\psi\rangle = |\psi\rangle \tag{3.3}$$

Evidently, only Pauli operators that are also observables, which we henceforth refer to as Pauli observables, can stabilize a non-zero state vector. This notion can be used to pick out a special class of Pauli subgroups known as *stabilizer groups*.

Definition 4 (Stabilizer group)**.** *Given any subgroup S of the n-qubit Pauli group P_n, the vectors in $\mathcal{H}_2^n$ stabilized by every element of S form a Hilbert subspace in $\mathcal{H}_2^n$, referred to as the stabilized subspace $\mathcal{H}_\mathsf{S}$. Then S is called a stabilizer group when $\mathcal{H}_\mathsf{S}$ is not empty.*

The abstract definition of stabilizer groups conceals the fact that their structure is in fact highly restrictive, as revealed in the following theorem (for completeness, a proof is provided in Appendix A.1).

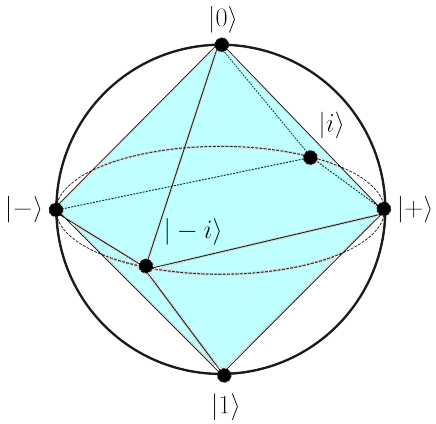
Theorem 5. *A subgroup S of the n-qubit Pauli group P_n is a stabilizer group if and only if it is an Abelian group generated by a set of $n \geq m \geq 0$ commuting, independent and non-trivial Pauli observables such that $-I \notin \mathsf{S}$ (with $m = 0$ corresponding to S being the trivial group). The stabilized subspace $\mathcal{H}_\mathsf{S}$ of S then has dimension $\dim(\mathcal{H}_\mathsf{S}) = 2^{n-m}$.*

3.2.2 Stabilizer states

We are now in a position to present the states that can be prepared by stabilizer circuits. By Theorem 5, we see that the stabilizer spaces of the largest stabilizer groups on n qubits – those generated by n commuting, independent and non-trivial Pauli observables – are 1-dimensional, and therefore spanned by a single *unique* pure state.

Definition 6 (Stabilizer state)**.** *An n-qubit stabilizer state is the unique pure state stabilized by a stabilizer group generated from n independent and non-trivial Pauli observables, or any statistical mixture of such pure states.*

As a simple example, we now identify all the stabilizer states of a single qubit. To do so, we first need to identify all the single-qubit stabilizer groups generated by a single non-trivial Pauli observable. These are evidently groups generated by $\pm X, \pm Y$ and $\pm Z$, so the pure single-qubit stabilizer states are the eigenstates of X, Y and Z.



stabilizer group	stabilizer state
$\langle X \rangle$	$\|+\rangle := \frac{1}{\sqrt{2}}(\|0\rangle + \|1\rangle)$
$\langle -X \rangle$	$\|-\rangle := \frac{1}{\sqrt{2}}(\|0\rangle - \|1\rangle)$
$\langle Y \rangle$	$\|i\rangle := \frac{1}{\sqrt{2}}(\|0\rangle + i\|1\rangle)$
$\langle -Y \rangle$	$\|-i\rangle := \frac{1}{\sqrt{2}}(\|0\rangle - i\|1\rangle)$
$\langle Z \rangle$	$\|0\rangle$
$\langle -Z \rangle$	$\|1\rangle$

Figure 3.1: The stabilizer states of a single qubit (blue octahedron).

As shown in Fig. 3.1, the stabilizer states of a single qubit therefore form an octahedron in the Bloch sphere, with vertices given by the eigenstates of X, Y and Z.

3.2.3 Unitary operations and measurements in stabilizer circuits

Unitary operations in stabilizer circuits are restricted to those executed using members of the *Clifford group*. Roughly speaking, these are unitaries that map Pauli operators to Pauli operators under conjugation.

Definition 7 (Clifford group)**.** *The Clifford group C_n for n qubits is the normalizer of P_n in the*

unitary group $\mathsf{U}(2^n)$*, that is*[4]

$$\mathsf{C}_n \coloneqq \{U \in \mathsf{U}(2^n) | UPU^\dagger \in \mathsf{P}_n \; \forall P \in \mathsf{P}_n\}/\{e^{i\theta} I | \theta \in [0, 2\pi)\}. \tag{3.4}$$

The Clifford group is more conveniently characterised by a standard set of generators. To do so, we must first define the *Hadamard* gate H and *phase* gate S on a single qubit, which are represented in the computational basis as

$$H \coloneqq \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}, \; S \coloneqq \begin{pmatrix} 1 & 0 \\ 0 & i \end{pmatrix}, \tag{3.5}$$

as well as the CNOT-gate with control qubit i and target qubit j as

$$\mathrm{CNOT}(i, j) \coloneqq |0\rangle\langle 0|_i \otimes \mathbb{1}_j + |1\rangle\langle 1|_i \otimes X_j. \tag{3.6}$$

Theorem 8 (Generators of the Clifford Group)**.** *The Clifford group on n qubits, C_n, can be generated from all Hadamard, phase and* CNOT *gates on n qubits, i.e.*

$$\mathsf{C}_n = \langle H_i, S_i, \mathrm{CNOT}_{i,j} \rangle_{i,j=1,\ldots,n, i \neq j} \tag{3.7}$$

The Clifford group can be promoted by a universal gateset through the addition of any gate outside it [58, 59], with a typical choice being the T-gate $\sqrt[4]{Z}$.

Measurements in stabilizer circuits are restricted to projective measurements in the eigenbasis of Pauli observables. In general, a stabilizer circuit is any sequence of preparing stabilizer states, performing Clifford unitaries, measuring Pauli observables, adaptive decision making conditioned on outcomes, and classical randomness, as well as statistical mixtures amongst such sequences.

3.2.4 Stabilizer circuits in higher prime dimensions

The formalism for stabilizer circuits in a higher prime dimension d is constructed analogously around the *Heisenberg-Weyl group* as the formalism for qubit stabilizer circuits are constructed

[4] Modding out by unitaries of the form $e^{i\theta} I$ is necessary as otherwise C_n would become infinite for the trivial reason that if $U \in \mathsf{C}_n$, then $e^{i\theta} U \in \mathsf{C}_n$ for any phase θ.

around the Pauli group. Let us first define the following generalisations of qubit X and Z operators to qudits of odd prime dimension d,

$$X_d \coloneqq \sum_{k=0}^{d-1} |k+1\rangle\langle k|\,,\ Z_d \coloneqq \sum_{k=0}^{d-1} \omega^k\, |k\rangle\langle k|\,, \tag{3.8}$$

wherein $\{|0\rangle\,, \ldots, |d-1\rangle\}$ is the computational basis, $\omega \coloneqq \exp\left(\frac{2\pi i}{d}\right)$ is the dth root of unity and addition should be interpreted modulo d. The n-qudit Heisenberg-Weyl group is then simply $\mathsf{HW}_{d,n} \coloneqq \{\langle X_d, Z_d\rangle^{\otimes n}\}$. While there is also a Heisenberg-Weyl group defined for a qubit, it differs from the single-qubit Pauli group in only containing $\pm iY$ and not $\pm Y$ as well (and therefore also does not contain $\pm i\mathbb{1}$). The allowed unitary operations in stabilizer circuits for qudits of odd prime dimension d are members of the normalizer for $\mathsf{HW}_{d,n}$ up to a global phase, while allowed measurements are projective measurements in the eigenbasis of any element drawn from $\mathsf{HW}_{d,n}$.

3.3 A Qubit Wigner Representation Respecting Parallel and Sequential Process Composition

In this section, we review the Wigner representation of qubit states W_ρ introduced in Ref. [72]. We then expand upon its properties and extend it to the representation of channels. In the process, we confirm that the channel representation we have produced respects the sequential and parallel composition of processes, the former crucially giving rise to an input-output relation $W_{\Phi(\rho)} = W_\Phi W_\rho$ in phase space mirroring how the channel Φ acts on ρ in state space, which justifies W_Φ as a meaningful phase-space representation of the channel Φ. Moreover, respect for the parallel composition of processes implies that the representation of product states factorizes over subsystems, a property which will prove computationally advantageous given that inputs to magic distillation protocols typically take on the form $\rho^{\otimes n}$.

3.3.1 Wigner representation of qubit states

We first establish some convenient notation. Let $\boldsymbol{u} \coloneqq (u_1, \ldots, u_n) \in \mathbb{Z}_2^n$ denote a binary vector. Furthermore, given any single qubit operator O, let us define the following abbreviation

$$O(\boldsymbol{u}) \coloneqq O^{u_1} \otimes \cdots \otimes O^{u_n}. \tag{3.9}$$

We can now present the phase space of our chosen qubit Wigner representation.

Definition 9 (Phase space). *Consider an n-qubit quantum system with total Hilbert space $\mathcal{H}_2^n$. We associate to this system a phase space $\mathcal{P}_n := \mathbb{Z}_2^n \times \mathbb{Z}_2^n$, where $\mathcal{P}_n$ consists of all pairs of n-bit strings $(\boldsymbol{u}_x, \boldsymbol{u}_z)$, and has a symplectic inner product $[\boldsymbol{u}, \boldsymbol{v}]$ defined as*

$$[\boldsymbol{u}, \boldsymbol{v}] := \boldsymbol{u}_z \cdot \boldsymbol{v}_x - \boldsymbol{v}_z \cdot \boldsymbol{u}_x \equiv \boldsymbol{u}_z \cdot \boldsymbol{v}_x + \boldsymbol{v}_z \cdot \boldsymbol{u}_x. \tag{3.10}$$

We note that the equivalence between the two forms of the symplectic inner product presented in this definition originates from the fact that we are working in modulo 2 arithmetic.

Because each point in the phase space of a single qubit is a pair of bits (u_x, u_z), it is often visualised as a 2×2 square grid, as shown in Fig. 3.2(a). Because $\mathcal{P}_n = \mathcal{P}_1^{\times n} = (\mathbb{Z}_2 \times \mathbb{Z}_2)^{\times n}$, $\mathcal{P}_n$ simply consists of all n-fold Cartesian products of points in a single-qubit phase space. Concretely, representing a phase-space point $\boldsymbol{u} \in \mathcal{P}_n$ as the pair of n-bit strings $\boldsymbol{u} = (\boldsymbol{u}_x, \boldsymbol{u}_z)$, we can regard $((u_x)_i, (u_z)_i)$ as the phase-space co-ordinates of the ith qubit, as seen for two qubits in Fig. 3.2(b).

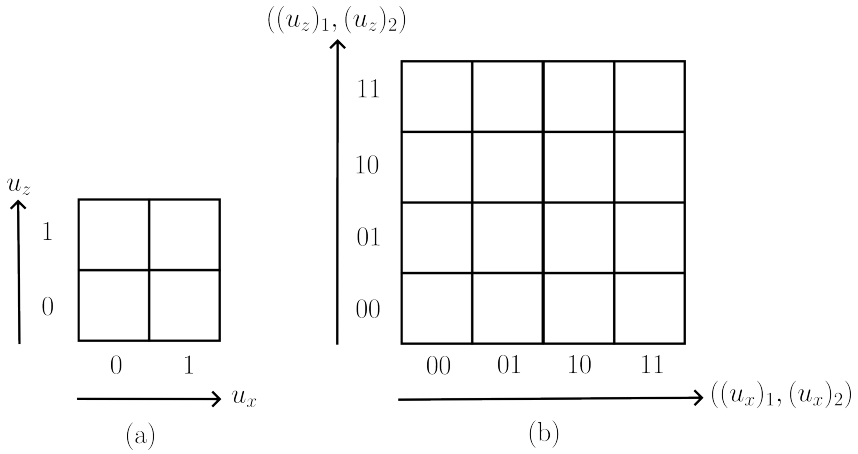


Figure 3.2: Phase spaces for (a) a single qubit and (b) two qubits.

We are now in a position to present our chosen representation over $\mathcal{P}_n$.

Definition 10 (Displacement Operators). *The displacement operators of the n-qubit phase space $\mathcal{P}_n$ are the set of operators $\{D_{\boldsymbol{u}}\}_{\boldsymbol{u} \in \mathcal{P}_n}$ where $D_{\boldsymbol{u}}$ is defined for $\boldsymbol{u} = (\boldsymbol{u}_x, \boldsymbol{u}_z) \in \mathcal{P}_n$ as*

$$D_{\boldsymbol{u}} := Z(\boldsymbol{u}_z) X(\boldsymbol{u}_x). \tag{3.11}$$

We can now construct the qubit Wigner representation used throughout this chapter as follows.

Definition 11 (Chosen qubit Wigner representation). *Given any n-qubit state ρ, we define its Wigner representation as*

$$W_\rho(\boldsymbol{u}) \coloneqq \frac{1}{2^n}\langle A_{\boldsymbol{u}}, \rho\rangle, \tag{3.12}$$

where $\{A_{\boldsymbol{u}}\}$ are the set of 2^{2n} phase-point operators on n-qubits, which are defined as

$$A_{\boldsymbol{0}} \coloneqq \frac{1}{2^n}\sum_{\boldsymbol{v}\in\mathcal{P}_n} D_{\boldsymbol{v}}, \quad A_{\boldsymbol{u}} \coloneqq D_{\boldsymbol{u}} A_{\boldsymbol{0}} D_{\boldsymbol{u}}^\dagger. \tag{3.13}$$

To gain some visual intuition for how our chosen Wigner representation works, Fig. 3.3 provides several examples showing the Wigner representations of commonly-encountered states.

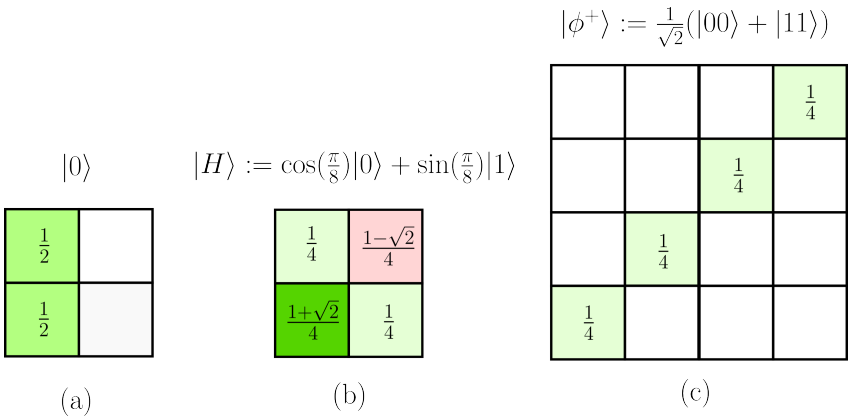


Figure 3.3: **Three examples for Wigner representations of qubit states.** (a) the computational basis state $|0\rangle$. (b) the Hadamard state (+1 eigenstate of the Hadamard gate), a typical magic state chosen for the Bravyi-Kitaev model, and so unsurprisingly represented with negative components (c) The ϕ^+ Bell state on two qubits.

As an aside, we can now motivate the naming of $\{D_{\boldsymbol{u}}\}_{\boldsymbol{u}\in\mathcal{P}_n}$ as displacement operators. Consider the Wigner representation of $D_{\boldsymbol{v}}\rho D_{\boldsymbol{v}}^\dagger$ for an arbitrary phase space displacement $\boldsymbol{v}$,

$$\begin{aligned} W_{D_{\boldsymbol{v}}\rho D_{\boldsymbol{v}}^\dagger}(\boldsymbol{u}) &\coloneqq \frac{1}{2^n}\langle A^\dagger, D_{\boldsymbol{v}}\rho D_{\boldsymbol{v}}^\dagger\rangle \equiv \frac{1}{2^n}\operatorname{Tr}\left[A_{\boldsymbol{u}}^\dagger D_{\boldsymbol{v}}\rho D_{\boldsymbol{v}}^\dagger\right] \\ &= \frac{1}{2^n}\operatorname{Tr}\left[D_{\boldsymbol{v}}^\dagger A_{\boldsymbol{u}}^\dagger D_{\boldsymbol{v}}\rho\right] \\ &= \frac{1}{2^n}\operatorname{Tr}\left[\sum_{\boldsymbol{a}\in\mathcal{P}_n}(-1)^{[\boldsymbol{u},\boldsymbol{a}]}(D_{\boldsymbol{a}}D_{\boldsymbol{v}})^\dagger D_{\boldsymbol{v}}\rho\right] \\ &= \frac{1}{2^n}\operatorname{Tr}\left[\left(\sum_{\boldsymbol{a}\in\mathcal{P}_n}(-1)^{[\boldsymbol{u}+\boldsymbol{v},\boldsymbol{a}]}D_{\boldsymbol{a}}^\dagger\right)\rho\right] = W_\rho(\boldsymbol{u}+\boldsymbol{v}), \end{aligned} \tag{3.14}$$

from which we see that the unitary channel executed by $D_{\boldsymbol{v}}$ has the effect of *displacing* the Wigner representation of ρ in phase space by $-\boldsymbol{v}$ (equivalent to $\boldsymbol{v}$ modulo 2).

If we were to change $X \to X_d$, $Z \to Z_d$ and $\mathcal{P}_n \to (\mathbb{Z}_d \times \mathbb{Z}_d)^{\times n}$ for some odd prime d, we would find that the construction of W_ρ has exactly copied that of Gross's Wigner representation except in one detail – every displacement operator in Gross's Wigner representation, $D^{(G)}_{\boldsymbol{u}}$ for some $\boldsymbol{u} \coloneqq (\boldsymbol{u}_x, \boldsymbol{u}_z) \in \mathbb{Z}_d \times \mathbb{Z}_d$, is multiplied by an additional phase factor to ensure that they are Hermitian as follows

$$D^{(G)}_{\boldsymbol{u}} \coloneqq \omega^{2^{-1}(\boldsymbol{u}_z \cdot \boldsymbol{u}_x)} Z_d(\boldsymbol{u}_z) X_d(\boldsymbol{u}_x), \tag{3.15}$$

where $\omega^{2^{-1}}$ denotes the unique dth root of unity that, when squared, produce ω [60]. The absence of a similar phase factor in Eq. (3.11) means the displacement operators of our chosen qubit Wigner representation, such as $D_{1,1} = iY$, are not Hermitian. We will shortly see that this produces some startling consequences.

3.3.2 Properties of the Wigner representation of qubit states

The Wigner representation we have chosen for qubit states behave very similarly to Gross's Wigner representation for the states of qudits with arbitrary odd prime dimension, which is unsurprising given their nearly identical construction. To begin with, we can show (see Appendix A.2.1) that the phase point operators defining W_ρ have the following properties, which are shared by phase-point operators defining Gross's Wigner representation [60].

Lemma 12. *The phase point operators $\{A_{\boldsymbol{u}}\}_{\boldsymbol{u} \in \mathcal{P}_n}$ have the following properties:*

(A1) *Factorizability: $A_{\boldsymbol{u}_X \oplus \boldsymbol{u}_Y} = A_{\boldsymbol{u}_X} \otimes A_{\boldsymbol{u}_Y}$ on a bipartite system XY, where $\boldsymbol{u}_X$ and $\boldsymbol{u}_Y$ are respectively points in the phase spaces of subsystems X and Y.*

(A2) *Orthogonality: $\langle A_{\boldsymbol{u}}, A_{\boldsymbol{v}} \rangle = 2^n \delta_{\boldsymbol{u},\boldsymbol{v}}$,*

(A3) *Unit trace: $\mathrm{Tr}[A_{\boldsymbol{u}}] = 1$,*

(A4) *Completeness: $\sum_{\boldsymbol{u} \in \mathcal{P}_n} A_{\boldsymbol{u}} = 2^n I$.*

As proved in Appendix A.2.2, these properties further ensure that our chosen Wigner representation for qubit states shares the following properties with Gross's Wigner representation.

Lemma 13. *The Wigner representation W_ρ for qubit states has the following properties*

(W1) *Informational completeness: Given any qubit state ρ, one can decompose it uniquely as $\rho = \sum_{\boldsymbol{u}\in\mathcal{P}_n} W_\rho(\boldsymbol{u}) A_{\boldsymbol{u}}$ because $\{A_{\boldsymbol{u}}\}_{\boldsymbol{u}\in\mathcal{P}_n}$ form a complete orthogonal basis for $\mathcal{B}(n)$ under the Hilbert-Schmidt inner product.*

(W2) *Normalization: the function representing any qubit state ρ sums to 1 over phase-space, i.e. $\sum_{\boldsymbol{u}\in\mathcal{P}_n} W_\rho(\boldsymbol{u}) = 1$.*

The lack of phase factors ensuring the Hermiticity of displacement operators means that the phase-point operators of our qubit Wigner representation, unlike those of Gross's Wigner representation, are *not Hermitian*. For instance, the single-qubit phase-point operator at the origin is $A_{0,0} = \frac{1}{4}(\mathbb{1} + X + Z + iY)$, which is evidently not Hermitian. Therefore, W_ρ is generally *complex*, and not guaranteed to be real on all states the way Gross's Wigner representation is. As a simple example, the +1 eigenstate of Y, $|i\rangle \coloneqq \frac{1}{\sqrt{2}}(|0\rangle + i\,|1\rangle)$, appears in our chosen Wigner representation as $W_{|i\rangle\langle i|} \coloneqq \frac{1}{4}(1+i, 1-i, 1-i, 1+i)$.

Nevertheless, even though the qubit displacement operators we have chosen are not Hermitian, they are constructed by Eq. (3.11) to always be *real* in the computational basis, which implies by Eq. (A.12) that qubit phase-point operators are real as well. This fact leads a direct correspondence between the real/imaginary parts of a state's Wigner representation and those of its density matrix in the computational basis (a proof is given in Appendix A.2.3).

Lemma 14. *Given any n-qubit quantum state ρ,*

$$\mathfrak{Re}[W_\rho(\boldsymbol{u})] = W_{\mathfrak{Re}(\rho)}(\boldsymbol{u}) \tag{3.16}$$

$$\mathfrak{Im}[W_\rho(\boldsymbol{u})] = W_{\mathfrak{Im}(\rho)}(\boldsymbol{u}) \tag{3.17}$$

for all $\boldsymbol{u} \in \mathcal{P}_n$, where $\mathfrak{Re}(\rho)$ and $\mathfrak{Im}(\rho)$ are respectively the real and imaginary parts of the density matrix of ρ in the computational basis.

We therefore obtain the following corollary.

Corollary 1. *The Wigner representation W_ρ of an n-qubit state ρ is real if and only if ρ is an n-rebit state, i.e., the density matrix ρ is real in the computational basis.*

In the next section, we extend our chosen qubit representation to quantum channels and verify

that the result respects the sequential and parallel compositions of processes.

3.3.3 Wigner representation of qubit channels

Denoting the set of bounded operators on the Hilbert space of n qubits by $\mathcal{B}(n)$ for the rest of this chapter, we can extend the qubit state Wigner representation defined in Eq. (3.12) to quantum channels between qubits.

Definition 15 (Wigner representation of qubit channels)**.** *The Wigner representation W_Φ of a quantum channel $\Phi : \mathcal{B}(n) \to \mathcal{B}(m)$ from n to m qubits is defined as the following linear map $W_\Phi : \mathcal{P}_n \to \mathcal{P}_m$ from the phase space of n qubits to that of m qubits*

$$W_\Phi(\boldsymbol{v}|\boldsymbol{u}) \coloneqq 2^{2n} W_{\mathcal{J}(\Phi)}(\boldsymbol{u} \oplus \boldsymbol{v}), \tag{3.18}$$

where $\mathcal{J}(\Phi)$ is the state on $n+m$ qubits constituting the Choi-Jamiolkowski representation of Φ.

This construction copies the extension of Gross's Wigner representation for qudits of odd prime dimension d from states to channels in Ref. [61] except for the substitution of 2^{2n} in place of the original d^{2n}. Unsurprisingly given their extremely similar construction, the Wigner representation for channels between qubits that we constructed share the following properties (proof in Appendix A.3) with the extension of Gross's Wigner representation to channels [61].

Lemma 16. *The Wigner representation W_Φ of a quantum channel $\Phi : \mathcal{B}(n) \to \mathcal{B}(m)$ from n to m qubits has the following properties:*

(W3) *(Input-Output Relation). If $\sigma = \Phi(\rho)$, then $W_\sigma(\boldsymbol{v}) = \sum_{\boldsymbol{u} \in \mathcal{P}_n} W_\Phi(\boldsymbol{v}|\boldsymbol{u}) W_\rho(\boldsymbol{u})$.*

(W4) *(Respects the sequential composition of channels). Given a channel $\Lambda : \mathcal{B}(n') \to \mathcal{B}(n)$ from n' to n qubits, we have that $W_{\Phi \circ \Lambda} = W_\Phi W_\Lambda$.*

(W5) *(Respects the parallel composition of channels). Given a channel $\Lambda : \mathcal{B}(n') \to \mathcal{B}(m')$ from n' to m' qubits, we have that $W_{\Phi \otimes \Lambda} = W_\Phi \otimes W_\Lambda$.*

(W6) *(Preserves normalization). Each column of W_Φ sums to 1, i.e. $\sum_{\boldsymbol{v} \in \mathcal{P}_m} W_\Phi(\boldsymbol{v}|\boldsymbol{u}) = 1$.*

It is worth spending a few moments unpacking the channel representation properties listed in the Lemma above. Property **(W3)** reveals that W_Φ turns every channel into a *matrix* mapping the phase-space representation of an input state to the channel onto the phase-space representation of the resultant output state from that channel, which justifies accepting W_Φ as a mean-

ingful phase-space representation of quantum channels. By taking Λ to be the state preparation channel for ρ, we can see that Property **(W3)** is a special case of the respect for the sequential composition of channels given in Property **(W4)**.

One useful implication of the respect for parallel composition guaranteed by Property **(W5)** is that, by taking Φ and Λ to be preparation channels for states ρ and σ, we obtain

$$W_{\rho\otimes\sigma} = W_\rho \otimes W_\sigma, \tag{3.19}$$

so our chosen qubit state representation factorizes over subsystems for product states.

Finally, Property **(W6)** is so called because it guarantees that the action of W_Φ will preserves the sum over phase space of any complex Wigner distribution over phase space $W(\boldsymbol{u}) : \mathcal{P}_n \to \mathbb{C}$, as it can be used to show

$$\begin{aligned}\sum_{\boldsymbol{v}\in\mathcal{P}_m} [W_\Phi W](\boldsymbol{v}) &= \sum_{\boldsymbol{v}\in\mathcal{P}_m} \left(\sum_{\boldsymbol{u}\in\mathcal{P}_n} W_\Phi(\boldsymbol{v}|\boldsymbol{u}) W(\boldsymbol{u}) \right) \\ &= \sum_{\boldsymbol{u}\in\mathcal{P}_n} \left(\sum_{\boldsymbol{v}\in\mathcal{P}_m} W_\Phi(\boldsymbol{v}|\boldsymbol{u}) \right) W(\boldsymbol{u}) = \sum_{\boldsymbol{u}\in\mathcal{P}_n} W(\boldsymbol{u}).\end{aligned} \tag{3.20}$$

We can interpret Property **(W6)** as the phase-space equivalent of Φ being trace-preserving – indeed, this property would not hold were Φ not trace-preserving.

Property **(W6)** further implies that a quantum channel Φ from n qubits to m qubits is represented by a stochastic matrix if and only if $W_\Phi(\boldsymbol{v}|\boldsymbol{u}) \geq 0$ for all $\boldsymbol{u} \in \mathcal{P}_n$ and $\boldsymbol{v} \in \mathcal{P}_m$. By Eq. (3.18), we equivalently have that Φ is stochastically represented if and only if its Choi-Jamiolkowski state $\mathcal{J}(\Phi)$ on $n+m$ qubits is represented by a genuine probability distribution on the phase space $\mathcal{P}_{n+m}$. In the next section, we will show that Φ is stochastically represented if $\mathcal{J}(\Phi)$ belongs to an important subset of stabilizer states known as *CSS states*.

3.4 Stochastic representation of CSS circuits

Because our qubit Wigner representation respects the sequential and parallel composition of processes, a recent no-go theorem [65] prohibits it from representing all stabilizer circuits

stochastically. As an example, the Hadamard gate is represented with negative entries as

$$W_H = \frac{1}{2}\begin{pmatrix} 1 & 1 & 1 & -1 \\ 1 & -1 & 1 & 1 \\ 1 & 1 & -1 & 1 \\ -1 & 1 & 1 & 1 \end{pmatrix}. \tag{3.21}$$

Nevertheless, we can identify a smaller subtheory that arises naturally in fault-tolerant quantum computing, is sufficiently large to enable universal quantum computation, and admits a stochastic representation. In particular, we show that a channel is stochastically represented if its Choi-Jamiolkowski state is CSS. Building on this result, we construct a stochastically-represented subset of stabilizer circuits wherein, roughly speaking, CSS states play the role of stabilizer states, and quickly review QCSI schemes built around such circuits wherein they act as the fault-tolerant but computationally classical subtheory that is promoted to universal quantum computing by the injection of magic states.

3.4.1 CSS states

We begin by clarifying what makes a stabilizer state CSS.

Definition 17 (CSS state). *A CSS state on n qubits is a pure stabilizer state whose stabilizer group can be generated by Pauli observables that are individually a tensor product of X or Z only, or a statistical mixture of such states.*

As a simple example, the Bell state $|\phi^+\rangle \coloneqq \frac{1}{\sqrt{2}}(|00\rangle + |11\rangle)$ has the stabilizer group

$$\mathsf{S}(|\phi^+\rangle) = \langle X_1X_2, Z_1Z_2\rangle, \tag{3.22}$$

and is therefore CSS. By contrast, $|\psi\rangle \coloneqq \mathbb{1} \otimes H\,|\phi^+\rangle$ is stabilized by

$$\mathsf{S}(|\psi\rangle) = \langle X_1Z_2, Z_1X_2\rangle, \tag{3.23}$$

and is not CSS because its stabilizer generators necessarily mix X and Z. As they are generators of stabilizer groups defining CSS states, we group X- and Z-type Pauli observables together under the name *CSS observables*.

It is worth noting at this juncture that *all CSS states are rebit states*. It can be shown (see Lemma 74) that every pure CSS state can be generated from a tensor product of $|+\rangle$ and $|0\rangle$ states by some sequence of CNOT, X and Z gates. As $|+\rangle$ and $|0\rangle$ have real density matrices in the computational basis, and unitary matrices representing CNOT, X and Z in the computational basis are also real, we can conclude that all pure CSS states, and therefore any statistical combination of them, have density matrices that are real in the computational basis.

3.4.2 Probabilistic representation of CSS states

The Discrete Hudson's Theorem for Gross's Wigner representation proved that a pure state on qudits of odd prime dimension is represented as a genuine probability distribution if and only if it is a stabilizer state [60]. Ref. [73] introduced a Wigner representation for rebits *alone* wherein a Discrete Hudson's Theorem could be recovered – every pure state on rebits is represented as a genuine probability distribution if and only if it is CSS. This representation coincides with W_ρ on all rebit states [74] (an original proof is provided in Appendix A.2.3 for completeness). We can therefore carry over the recovery of the Discrete Hudson's Theorem for rebits in Ref. [73] and conclude that W_ρ represents all CSS states as probability distributions.

3.4.3 Stochastic representation of completely CSS-preserving channels

Since we saw at the end of section 3.3.3 that a qubit channel Φ is stochastically represented if and only if $\mathcal{J}(\Phi)$ is probabilistically represented, we arrive at the following theorem.

Theorem 18. *A channel Φ between qubits is stochastically represented if $\mathcal{J}(\Phi)$ is a CSS state.*

Theorem 18 can be leveraged to identify stochastically-represented qubit stabilizer operations in a systematic way. A channel Φ from a system of qubits B is CSS-preserving if $\Phi(\rho)$ is always CSS whenever ρ is, and completely CSS-preserving if, given any CSS state ρ_{AB} on another system of qubits A as well as B, $\mathsf{id}_A \otimes \Phi_B(\rho_{AB})$ is always CSS.

We now note that the canonical maximally entangled state $|\phi_n^+\rangle$ over two sets of n qubits, which is defined as $|\phi_n^+\rangle := \frac{1}{\sqrt{2^n}} \left(\sum_{\boldsymbol{k} \in \{0,1\}^n} |\boldsymbol{k}\rangle \otimes |\boldsymbol{k}\rangle \right)$, is CSS because one can verify that is stabilized by the group

$$\mathsf{S}(|\phi_n^+\rangle) = \langle Z_i Z_{n+i}, X_i X_{n+i} \rangle_{i=1,\ldots,n}. \tag{3.24}$$

Therefore, if Φ is completely CSS-preserving, $\mathcal{J}(\Phi)$ must be CSS, and so by Theorem 18, we conclude that

Corollary 2. *A completely CSS-preserving channel between qubits is stochastically represented.*

This corollary is the qubit equivalent of a similar result for qudits of odd prime dimension. Under Gross's Wigner representation, stabilizer states on such qudits are represented as valid probability distributions over phase space [60]. Therefore, if a channel between qudits of the same odd prime dimension is completely stabilizer-preserving, that channel will be stochastically represented under the extension of Gross's Wigner representation to channels [61].

3.4.4 QCSI with completely CSS-preserving channels

Completely CSS-preserving channels turn out to also be capable of QCSI. To motivate this, we first highlight (proof in Appendix A.4.3) that such channels must at least include the following subset of stabilizer circuits.

Lemma 19 (CSS circuits)**.** *Any sequence of the following operations:*

1. *introducing a CSS state on any number of qubits,*
2. *performing a completely CSS-preserving gate on any number n of qubits, i.e., any gate drawn from the group $\mathsf{G}(n) := \langle \mathrm{CNOT}(i,j), Z_i, X_i \rangle_{i,j=1,\ldots,n, i\neq j}$.*
3. *projectively measuring a CSS observable (with the possibility of classical control conditioned on outcome),*
4. *discarding any number of qubits,*

as well as statistical mixtures of such sequences, is completely CSS-preserving.

By using CSS-preserving rather than completely CSS-preserving gates, channels covered by Lemma 19 can be promoted to the subset of stabilizer operations wherein CSS states play the role of stabilizer states. However, both groups of gates are equally powerful for magic distillation (see discussion in Appendix A.4.3.1). Thus we directly refer to the set of channels covered by Lemma 19 as *CSS circuits*, and conclude that every operation within this subtheory is stochastically represented.

CSS circuits are capable of universal quantum computation [73, 74], and we direct interested readers to a summary of the QCSI scheme presented in Ref. [73] at Appendix A.5. In fact, we

will see in the next chapter that CSS circuits form the basis of many existing magic distillation protocols constructed around a family of QECCs known as *CSS codes*.

3.5 Entropic constraints on qubit magic distillation

We now generalise the statistical mechanics framework of magic distillation introduced by Ref. [27] from the Bravyi-Kitaev model for qudits of odd prime dimensions to completely CSS-preserving QCSI on rebits. The framework considers how a class of magic distillation protocols transforms the Wigner representations for a *pair* of quantum states – one a noisy magic state, the other a non-magic state singled out by the characteristic physics of those protocol – and identifies a family of relative entropies to characterise which transformations are ruled out. In this section, we briefly review the approach taken in Ref. [27] to extend relative majorization to quasiprobability distributions, and how that leads to the extension of a dense subset of α-Rényi divergences from classical statistical mechanics to quantify the non-classical order in magic states under distillation. We then adapt this work for rebit magic state distillation using completely CSS-preserving circuits.

3.5.1 Majorization of Wigner representations

Standard majorization [68, 67] is a tool for capturing a notion of stochastic ordering amongst probability distributions that regards the uniform probability distribution as the "most disordered". This kind of stochastic ordering underpins classical statistical mechanics, wherein the macrostate of an isolated system tends towards a uniform distribution over its microstates as time goes by, corresponding to an increase in the system's Gibbs entropy.

Definition 20 ((Standard) majorization [75])**.** *Let $\boldsymbol{p}$ and $\boldsymbol{p}'$ be probability distributions over the sample space Ω respectively. We say that $\boldsymbol{p}$ majorizes $\boldsymbol{p}'$, denoted $\boldsymbol{p} \succ \boldsymbol{p}'$, when there exists a stochastic map M on Ω that transforms $\boldsymbol{p}$ into $\boldsymbol{p}'$ while preserving the uniform probability distribution $\boldsymbol{\mu}$ over Ω, i.e.*

$$M(\boldsymbol{p}) = \boldsymbol{p}' \tag{3.25}$$

$$M(\boldsymbol{\mu}) = \boldsymbol{\mu}. \tag{3.26}$$

Necessary conditions on majorization are provided by the family of α-*Rényi entropies*, which

becomes the Shannon entropy constructed on the basis of the Gibbs entropy at $\alpha = 1$.

Definition 21 ((Classical) α-Rényi entropies). *The (classical) α-Rényi divergence D_α for a probability distribution $\boldsymbol{p}$ at $\alpha \in (0, \infty]$ is*

$$H_\alpha(\boldsymbol{p}) \coloneqq \begin{cases} \frac{1}{1-\alpha} \log \sum_i p_i^\alpha & \text{for } \alpha \in (0,\infty) \cup (1,\infty) \\ -\sum_i p_i \log p_i & \text{for } \alpha = 1 \\ -\log \max_i p_i & \text{for } \alpha = \infty \end{cases} \tag{3.27}$$

Each α-Rényi entropy has the crucial property that it cannot be decreased by any classical channel, which at its most generic is simply a stochastic mixing of states, that preserves the uniform probability distribution. We consequently obtain the following necessary conditions on majorization [76]:

$$\boldsymbol{p} \succ \boldsymbol{p}' \implies \forall \alpha \in (0, \infty] : H_\alpha(\boldsymbol{p}) \leq H_\alpha(\boldsymbol{p}'). \tag{3.28}$$

In the next section, we will see that H_α monotonically decreases in a measure of statistical distance between a given probability distribution and the uniform probability distribution over the same sample space. We therefore see that majorization describes stochastic processing that brings probability distributions closer to the uniform probability distribution.

Ref. [27] made the highly non-trivial discovery that the majorization relations of Eq. (3.28) can in fact be extended from probability distributions to quasiprobability distributions for a dense subset of α-Rényi entropies.

Theorem 22 ([27]). *Let $\boldsymbol{w} = (w_1, \ldots, w_n)$ and $\boldsymbol{w}' = (w'_1, \ldots, w'_n)$ be two quasiprobability distributions. If $\boldsymbol{w} \succ \boldsymbol{w}'$ then*

$$\forall \alpha \in \mathcal{A} \coloneqq \left\{ \frac{2a}{2b-1} \ : \ a, b \in \mathbb{N}^+, a \geq b \right\} : H_\alpha(\boldsymbol{w}) \geq H_\alpha(\boldsymbol{w}'), \tag{3.29}$$

for all $\alpha \in \mathcal{A} \coloneqq \left\{ \frac{2a}{2b-1} \ : \ a, b \in \mathbb{N}^+, a \geq b \right\}$.

The rigorous framework established in Eq. (3.28) for describing stochastic order in classical statistical mechanics therefore extends to QCSI via stochastically-represented quantum channels that preserves the uniform distribution over phase space, even though the $\alpha = 1$ case (Shannon

entropy) is no longer well-defined on quasiprobability distributions. Under the qubit Wigner representation we chose (or Gross's Wigner representation for qudits of odd prime dimension), this restricts us to stochastically-represented *unital* quantum channels, as the maximally mixed state is uniquely represented as a uniform distribution over phase space.

We highlight that the entropies used within this framework are not Hermitian observables or even POVMs, but resource monotones characterising a pre-order defined by which states are accessible from which other states via unital stochastic processing of their phase-space representations. These entropies thereby furnish examples of physical properties isolated by the QRT formalism that cannot be captured by more traditional notions focussed on what can be directly measured.

Furthermore, under the qubit Wigner representation we chose, there always exists $\alpha \in \mathcal{A}$ such that $H_\alpha(W_\rho) < 0$ when ρ is a (distillable) rebit magic state for completely CSS-preserving QCSI (see Appendix A.6.3). Thus this chapter paints a conceptual picture of universal qubit computing as classical statistical mechanics where one gains access to resources with non-classically low – i.e. negative – entropies, which "fuel" its computational advantage.

3.5.2 Relative majorization of Wigner representations

Standard *relative* majorization [68, 67] allows us to further capture the stochastic ordering amongst *pairs* of probability distributions.

Definition 23 ((Standard) relative majorization)**.** *Let $(\boldsymbol{p}, \boldsymbol{r})$ and $(\boldsymbol{p}', \boldsymbol{r}')$ be pairs of probability distributions over sample spaces Ω and Ω' respectively. We say that $(\boldsymbol{p}, \boldsymbol{r})$ relatively majorizes $(\boldsymbol{p}', \boldsymbol{r}')$, which we denote by $(\boldsymbol{p}, \boldsymbol{r}) \succ (\boldsymbol{p}', \boldsymbol{r}')$, when there exists a stochastic map $M : \Omega \to \Omega'$ that transforms the first pair of distributions into the second, i.e.*

$$M(\boldsymbol{p}) = \boldsymbol{p}' \tag{3.30}$$

$$M(\boldsymbol{r}) = \boldsymbol{r}'. \tag{3.31}$$

Similar to the case of standard majorization, a family of measures for statistical distance based on the α-Rényi entropies, known as the *α-Rényi divergences*, provide necessary conditions for relative majorization.

Definition 24 ((Classical) α-Rényi divergence)**.** *The (classical) α-Rényi divergence D_α for an input probability distribution $\boldsymbol{p}$ and a reference probability distribution $\boldsymbol{r}$ at $\alpha \in (0, \infty]$ is*

$$D_\alpha(\boldsymbol{p}||\boldsymbol{r}) \coloneqq \begin{cases} \frac{1}{\alpha-1} \log \sum_i p_i^\alpha r_i^{1-\alpha} & \text{for } \alpha \in (0,1) \cup (1,\infty) \\ \sum_i p_i \log\left(\frac{p_i}{r_i}\right) & \text{for } \alpha = 1 \\ \log \max_i \left(\frac{p_i}{r_i}\right) & \text{for } \alpha = \infty \end{cases} \tag{3.32}$$

For $\alpha > 1$, we read $p_i^\alpha r_i^{1-\alpha}$ as $p_i^\alpha / r_i^{(\alpha-1)}$, and adopt the conventions $0/0 \coloneqq 0$ and $x/0 \coloneqq \infty$ for $x > 0$ to accommodate points at which probability distributions are 0. We note that the α-Rényi entropy for a probability distribution $\boldsymbol{p}$ on a sample space Ω is related to the α-Rényi divergence of $\boldsymbol{p}$ relative to the uniform probability distribution $\boldsymbol{\mu}$ over Ω as

$$H_\alpha(\boldsymbol{p}) = \log[\dim(\Omega)] - D_\alpha(\boldsymbol{p}||\boldsymbol{\mu}). \tag{3.33}$$

The distance between two probability distributions as measured by an α-Rényi entropy cannot be increased by any classical channel. Formally, given any stochastic map $M : \Omega \to \Omega'$, it has been shown [76] that

$$\forall \alpha \in (0, \infty] : D_\alpha(\boldsymbol{p}||\boldsymbol{r}) \geq D_\alpha(M(\boldsymbol{p})||M(\boldsymbol{r})). \tag{3.34}$$

from which we obtain the following necessary conditions on when one pair of probability distributions relatively majorize another:

$$(\boldsymbol{p}, \boldsymbol{r}) \succ (\boldsymbol{p}', \boldsymbol{r}') \implies \forall \alpha \in (0, \infty] : D_\alpha(\boldsymbol{p}||\boldsymbol{r}) \geq D_\alpha(\boldsymbol{p}'||\boldsymbol{r}'). \tag{3.35}$$

Ref. [27] also discovered that the relative majorization relations of Eq. (3.35) can be extended to input *quasiprobability* distributions for the same dense subset of α-Rényi entropies identified in Theorem 22, provided that the reference distributions are strictly positive probability distributions. In fact, this requirement can be relaxed as follows (see proof in Appendix A.6).

Theorem 25. *Let $\boldsymbol{w} = (w_1, \ldots, w_N)$ and $\boldsymbol{w}' = (w'_1, \ldots, w'_{N'})$ be any two quasiprobability distributions and let $\boldsymbol{r} = (r_1, \ldots, r_N)$ and $\boldsymbol{r}' = (r'_1, \ldots, r'_{N'})$ be any two probability distributions. If $(\boldsymbol{w}, \boldsymbol{r}) \succ (\boldsymbol{w}', \boldsymbol{r}')$ such that $\mathrm{supp}(\boldsymbol{w}) \subseteqq \mathrm{supp}(\boldsymbol{r})$ and therefore $\mathrm{supp}(\boldsymbol{w}') \subseteqq \mathrm{supp}(\boldsymbol{r}')$, then*

$$D_\alpha(\boldsymbol{w}||\boldsymbol{r}) \geq D_\alpha(\boldsymbol{w}'||\boldsymbol{r}'), \tag{3.36}$$

or all $\alpha \in \mathcal{A} \coloneqq \left\{ \frac{2a}{2b-1} \,:\, a, b \in \mathbb{N}^+, a \geq b \right\}$.

We note that since the set $\mathcal{A}$ is dense in $\alpha \in (1, \infty)$, one could extend the definition of $D_\alpha(\boldsymbol{w}||\boldsymbol{r})$ to all $\alpha \in (1, \infty]$ by continuity. We therefore speculate that it may be possible to extend all results in this part of the thesis from $\alpha \in \mathcal{A}$ to all $\alpha \in (1, \infty]$.

We now apply the generalised relative majorization relations in Theorem 25 to the Wigner representations of magic states. To do so, we first introduce the following definition to pull together the shared ingredients of Gross's Wigner representation and the qubit representation developed in this chapter that enable us to effectively draw on the results of Ref [27].

Definition 26 (Generalised Gross's Wigner representation)**.** *Let $\mathcal{R}$ be a quantum subtheory for d-dimensional qudits, and let $\mathcal{M}$ be a set of magic states that can promote $\mathcal{R}$ to universal quantum computation such that $\mathcal{F} \cup \mathcal{M}$ is closed under $\mathcal{R}$, where $\mathcal{F}$ are the free states in $\mathcal{R}$ (though $\mathcal{M}$ need not be the full set of resource states in the QRT defined by $\mathcal{R}$). We then call W a generalised Gross's Wigner representation with respect to the QCSI scheme $(\mathcal{R}, \mathcal{F}, \mathcal{M})$ if*

1. *each free state $\rho_S \in \mathcal{F}_S$ of a system S is represented as a unique probability distribution W_ρ over a phase space $\mathcal{P}_S$ associated to S,*
2. *each magic state $\sigma_S \in \mathcal{M}$ of a system S is represented as a unique quasiprobability distribution W_σ over $\mathcal{P}_S$,*
3. *each free channel $\Phi_{S\to S'} \in \mathcal{R}$ from S to S' is represented as a unique stochastic map W_Φ from $\mathcal{P}_S$ to $\mathcal{P}_{S'}$,*
4. *W respects the sequential and parallel composition of processes.*

We then have the following Lemma (see proof in Appendix A.6.1), which generalises Theorem 11 of Ref. [27] to qubits with the slight upgrade suggested by Theorem 25.

Lemma 27. *Let W be a generalised Gross's Wigner representation for d-dimensional qudits with respect to the QCSI scheme $(\mathcal{R}, \mathcal{F}, \mathcal{M})$. Consider states $\rho \in \mathcal{F} \cup \mathcal{M}$, $\tau \in \mathcal{F}$ such that $\mathrm{supp}(W_\rho) \subseteqq \mathrm{supp}(W_\tau)$, and a stochastically-represented channel $\Phi \in \mathcal{R}$. Then the α-Rényi divergence $D_\alpha(\cdot||\cdot)$ is well-defined and satisfies the following properties for $\alpha \in \mathcal{A}$:*

(D1) *Non-negativity: $D_\alpha(W_\rho||W_\tau) \geq 0$.*

(D2) *Additivity: $D_\alpha(W_{\rho^{\otimes n}}||W_{\tau^{\otimes n}}) = nD_\alpha(W_\rho||W_\tau)$ for all $n \in \mathbb{N}^+$.*

(D3) *Data-Processing Inequality: $D_\alpha(W_\rho||W_\tau) \geq D_\alpha(W_{\Phi(\rho)}||W_{\Phi(\tau)})$, where by Lemma 70 we also have that $\mathrm{supp}(W_{\Phi(\rho)}) \subseteqq \mathrm{supp}(W_{\Phi(\tau)})$.*

3.5.3 Using relative majorization to generate magic monotones

The standard approach to obtaining constraints on magic distillation is tracking a *magic monotone* [14, 51, 77, 78, 79], which is any property of a quantum system that cannot be increased under some class of magic non-generating operations (e.g. stabilizer operations). The paradigmatic example of a magic monotone is mana [14], the total negativity in Gross's Wigner representation $W^{(G)}$ of a state ρ on qudits of odd prime dimension:

$$\mathcal{M}(\rho) := \log \sum_{\boldsymbol{u} \in \mathcal{P}_n} \left| W_\rho^{(G)}(\boldsymbol{u}) \right|. \tag{3.37}$$

We can leverage Lemma 27 to systematically identify the following family of magic monotones for any QCSI scheme by using τ as a *variational parameter* to optimise over non-magic states (see proof in Appendix A.6.2).

Theorem 28. *Let W be a generalised Gross's Wigner representation for the QCSI scheme $(\mathcal{R}, \mathcal{F}, \mathcal{M})$. We then have that*

$$M_\alpha(\rho) := \inf_{\substack{\tau \in \mathcal{F} \text{ s.t.} \\ \mathrm{supp}(W_\rho) \subseteqq \mathrm{supp}(W_\tau)}} D_\alpha(W_\rho||W_\tau), \tag{3.38}$$

is a magic monotone for any $\alpha \in \mathcal{A}$. Furthermore, M_α can be used to lower-bound the overhead n of a magic distillation process $\rho^{\otimes n} \to \sigma$ within this QCSI scheme as

$$\forall \alpha \in \mathcal{A} : n \geq \frac{M_\alpha(\sigma)}{M_\alpha(\rho)}. \tag{3.39}$$

In fact, in the limit of $\alpha \to 1$, these magic monotones tend towards a generalisation of mana for an arbitrary generalised Gross's Wigner representation W (see proof in Appendix A.6.2).

Lemma 29. *Let W be a generalised Gross's Wigner representation for a QCSI scheme $(\mathcal{R}, \mathcal{F}, \mathcal{M})$. Then given any state $\rho \in \mathcal{F} \cup \mathcal{M}$ covered by the QCSI scheme, we have that*

$$\lim_{\alpha \to 1^+} (\alpha - 1) M_\alpha(\rho) = \mathcal{M}_W(\rho) \coloneqq \log \sum_{\boldsymbol{u}} |W_\rho(\boldsymbol{u})| \tag{3.40}$$

where the limit is taken through a sequence of rational values drawn from $\mathcal{A}$.

3.5.4 Entropic conditions on completely CSS-preserving magic distillation

We can exploit the relative majorization of Wigner representations to obtain more targeted constraints on magic distillation than the magic monotone approach described in the previous section. Alongside the magic distillation process $\rho^{\otimes n} \to \rho'$, we can identify a *reference process* $\tau^{\otimes n} \to \tau'$ between non-magic states τ and τ' that singles out a *particular class* of distillation protocols – that is, while protocols within this class will always perform $\tau^{\otimes n} \to \tau'$, protocols outside this class may not. We now apply Lemma 27 to completely CSS-preserving magic distillation on rebits to demonstrate how it can generate entropic constraints specialised for what protocols singled out by such a reference process can achieve.

Theorem 30. *Let ρ be a noisy multirebit magic state and τ be a CSS state where we have that $\mathrm{supp}(W_\rho) \subseteqq \mathrm{supp}(W_\tau)$. If there exists a completely CSS-preserving magic distillation protocol Φ such that $\Phi(\rho^{\otimes n}) = \rho'$ and $\tau' \coloneqq \Phi(\tau^{\otimes n})$, then*

$$\Delta D_\alpha \coloneqq n D_\alpha(W_\rho || W_\tau) - D_\alpha(W_{\rho'} || W_{\tau'}) \geq 0 \tag{3.41}$$

for the qubit Wigner representation of Eq. (3.18) and all $\alpha \in \mathcal{A} \coloneqq \{\frac{2a}{2b-1}, | a, b \in \mathbb{N}^+, a \geq b\}$.

In the next chapter, we will show how to the reference process may be chosen to produce entropic constraints specialised to a family of protocols based on CSS codes. For now, we note that the reference process $\tau^{\otimes n} \mapsto \tau'$ can also be used to take into account limitations in the hardware carrying out magic distillation. For instance, Ref. [27] uses the reference process to preserve the state at equilibrium with a heat bath in order to encode a background temperature

or free energy production in the distillation hardware.

Theorems 28 and 30 motivate why we demanded a Wigner representation for qubits that respected the parallel composition of processes. One of the key aims of the present work is to identify fundamental limitations on the overhead of magic distillation, and we can straightforwardly extract a lower bound on such an overhead in Eq. (3.41) and Eq. (3.39) only because $D_\alpha(W_{\rho^{\otimes n}}||W_{\tau^{\otimes n}}) = nD_\alpha(W_\rho||W_\tau)$, which (as seen in the proof of Lemma 27) relies on the respect of W for the parallel composition of processes.

Chapter 4

Entropic trade-off relations in stabilizer code distillation

4.1 Introduction

Broadly speaking, the aim of magic distillation is to convert many copies of a noisy magic state into fewer copies at any desired higher fidelity with respect to a target pure magic state. Each run of any magic distillation protocol usually only succeeds probabilistically, and a sequence of successful runs may be needed to go from a typical input error rate of $\sim 10^{-4}$ [26] to a desired error rate of at most $\sim 10^{-9}$ at which useful and classically intractable quantum computation starts to become possible (particular computations may require error rates that are many orders of magnitude lower) [80]. Consequently, distillation is expected to dominate the resource cost of QCSI, though recently significant progress has been made to reduce the overhead they require [26]. Understanding what fundamental trade-offs must be made in each round of magic distillation between reducing resource costs and improving fidelity, especially in ways that could be tailored to the specific physics of the protocols used, is therefore of vital practical interest, and may be instructive for designing protocols with optimised parameters.

Almost all magic distillation protocols to-date [18, 20, 21, 22, 23] are based on a subclass of QECCs known as *Calderbank-Shor-Steane (CSS) codes* [24, 25]. CSS codes can be constructed from two classical linear codes, allowing one to draw on a plethora of results from classical coding theory to construct quantum codes with symmetries (such as triorthogonality [81])

that are useful for encoding operations transversally. Furthermore, CSS codes are examples of *stabilizer codes*, the class of QECCs defined by codespaces that are stabilized by stabilizer groups. One family of protocols that project onto the codespaces of such codes, known as *stabilizer code projection protocols*, broadly underpin the possibility of magic distillation in the Bravyi-Kitaev model, where every protocol producing a single copy of the desired magic state can be decomposed into stabilizer code projection protocols and simply re-preparing stabilizer states [82]. We therefore apply the statistical framework for QCSI that was extended to qubits in the previous chapter to obtain trade-off relations in parameters governing the performance of stabilizer code projection protocols.

4.1.1 Chapter Summary

The rest of this chapter proceeds as follows. We begin with a review of stabilizer codes in section 4.2, which forms the main technical background for this chapter. In section 4.3, we establish that qubit CSS code projection protocols can be stochastically represented by the Wigner representation defined in the previous chapter, and identify a reference process for such protocols by exploiting how they act on the maximally mixed state. In section 4.4, we use the statistical framework for qubit QCSI obtained in the previous chapter to identify entropic trade-off relations on parameters controlling qubit CSS code projection protocols, and demonstrate how such constraints give rise to lower *and* upper bounds on the code length ($\sim$ resource cost) of such protocols. To our knowledge, these are the first fundamental *upper* bounds on resource cost for a family of distillation protocols identified in the literature. For an example application of distilling Hadamard states via CSS code projection protocols, our lower bounds on code length are competitive with those generated using state-of-the-art magic monotones in some parameter regimes of interest, which can be combined with our upper bounds to produce no-go conditions on the code length of CSS code projection protocols that could achieve a desired distillation process. We conclude in section 4.5 by extending our trade-off relations to stabilizer code projection protocols in qudits of odd prime dimension.

4.2 Technical background: stabilizer codes

We begin with a brief review of stabilizer codes on qudits of prime dimension, of which CSS codes are a subclass. Much more detailed and insightful exposition of this material can be

found at Ref. [54]. For simplicity, we will use the term "Pauli operators" throughout this section to refer to members of the multiqudit Heisenberg-Weyl groups for qudits of odd prime dimension as well as members of the multiqubit Pauli groups.

Though a quantum computer could simply store a qudit of information, or *logical qudit*, on a single *physical qudit* of the same dimension, this approach makes error detection impossible because no state on that physical qudit could only arise through error. Like its classical counterpart, quantum error correction is only possible if we encode our logical qudit redundantly as a *subspace*, known as the *codespace*, within the Hilbert space of a physical system, usually *several* logical qudits of the same dimension. In the case of stabilizer codes, these codespaces are subspaces stabilized by stabilizer groups.

Definition 31 (Stabilizer code). *An $[[n, k]]$ stabilizer code for qudits of prime dimension d is a QECC that encodes k logical qudits across n physical qudits in a codespace $\mathcal{C}$ that constitutes the subspace stabilized by an n-qudit stabilizer group requiring $(n-k)$ independent and non-trivial generators, where we have $n \geq k \geq 0$.*

We will denote a stabilizer code directly by its codespace $\mathcal{C}$ and the stabilizer group of $\mathcal{C}$ by $\mathsf{S}(\mathcal{C})$. When $\mathsf{S}(\mathcal{C})$ can be generated from Pauli operators that are individually products of single-qudit X operators only or single-qudit Z operators only, then $\mathcal{C}$ is known as a *CSS code*. As an example, the $[[7, 1]]$ qubit *Steane code* is a CSS code because its codespace is stabilized by

$$\mathsf{S}_{\text{steane}} \coloneqq \langle X_4X_5X_6X_7, X_2X_3X_6X_7, X_1X_3X_5X_7, Z_4Z_5Z_6Z_7, Z_2Z_3Z_6Z_7, Z_1Z_3Z_5Z_7\rangle. \tag{4.1}$$

An $[[n, k]]$ stabilizer code $\mathcal{C}$ for qudits of prime dimension d carries out its encoding process via an *encoding unitary* $U_{\mathcal{C}} : \mathcal{H}_d^{\otimes k} \otimes |0\rangle^{\otimes(n-k)} \to \mathcal{C}$, which is always a *Clifford unitary*. An example encoding unitary for the Steane code is shown in Fig. 4.1.

Given any state $|\psi\rangle \in \mathcal{H}_d^{\otimes k}$ encoded by the stabilizer code as $|\psi_L\rangle \coloneqq U_{\mathcal{C}}(|\psi\rangle |0\rangle^{\otimes(n-k)}) \in \mathcal{C}$, we can calculate that

$$\forall j \in \{n-k+1, \ldots, n\} : (U_{\mathcal{C}} Z_j U_{\mathcal{C}}^\dagger) |\psi_L\rangle = (U_{\mathcal{C}} Z_j U_{\mathcal{C}}^\dagger U_{\mathcal{C}} |\psi\rangle |0\rangle^{\otimes(n-k)} = |\psi_L\rangle \tag{4.2}$$

which implies that $\mathsf{S}(\mathcal{C})$ can be generated as $\mathsf{S}(\mathcal{C}) = \langle U_{\mathcal{C}}(Z_{n-k+1})U_{\mathcal{C}}{}^\dagger, \ldots, U_{\mathcal{C}}(Z_n)U_{\mathcal{C}}{}^\dagger\rangle$. We will

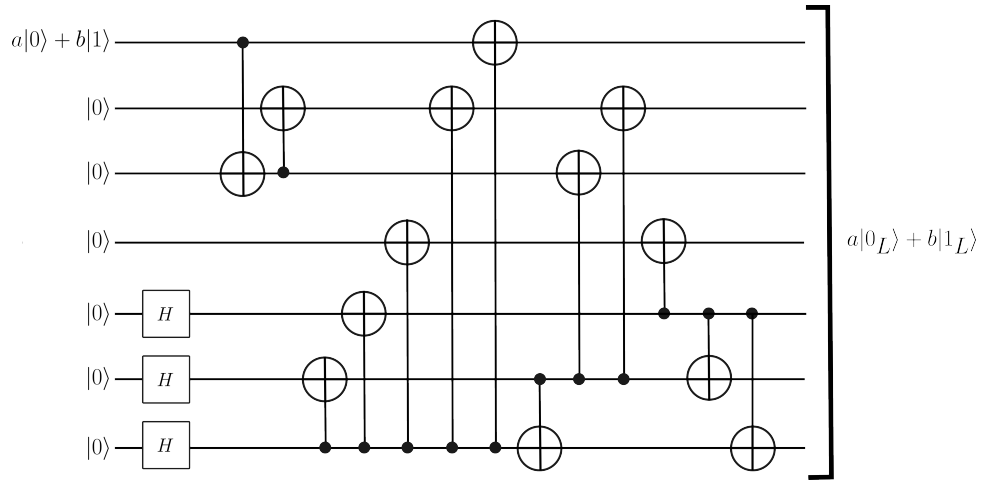


Figure 4.1: An encoding circuit for the Steane code, where $|0_L\rangle \coloneqq \frac{1}{\sqrt{8}}(|0000000\rangle + |1010101\rangle + |0110011\rangle + |1100110\rangle + |0001111\rangle + |1011010\rangle + |0111100\rangle + |1101001\rangle)$ and $|1_L\rangle \coloneqq \frac{1}{\sqrt{8}}(|1111111\rangle + |0101010\rangle + |1001100\rangle + |0011001\rangle + |1110000\rangle + |0100101\rangle + |1000011\rangle + |0010110\rangle)$.

refer to the channel $\mathcal{U}_\mathcal{C}(\cdot) \coloneqq U_\mathcal{C}(\cdot)U_\mathcal{C}{}^\dagger$ as the *encoding channel*, and its inverse $\mathcal{U}_\mathcal{C}^\dagger(\cdot) \coloneqq U_\mathcal{C}{}^\dagger(\cdot)\mathcal{U}_\mathcal{C}$ as the *decoding channel*.

In general, we will denote the encoding of states and channels under a QECC by the subscript $(\cdot)_L$ (for logical as opposed to physical states and channels). An $[[n, k]]$ stabilizer code encodes an arbitrary k-qudit state $|\psi\rangle$ as $|\psi_L\rangle \coloneqq U_\mathcal{C} |\psi\rangle \otimes |0\rangle^{\otimes(n-k)}$. Of particular note is the *logical basis* $\{|\boldsymbol{k}_L\rangle \coloneqq U_\mathcal{C} |\boldsymbol{k}\rangle \otimes |0\rangle^{\otimes(n-k)} \, | \boldsymbol{k} \in \mathbb{Z}_d^n\}$, which encode the computational basis states of k qudits.

Stabilizer codes were designed for protection against Pauli errors, i.e. any error process described by a unitary channel $E(\cdot)E^\dagger$ where E is a Pauli operator on n qudits. From section 3.2, we see that any two Pauli operators P_1 and P_2 on qudits of prime dimension d have a conjugation relation of the form

$$P_1 P_2 = \omega^s P_2 P_1, \text{ where } \omega \coloneqq \exp\left(\frac{2\pi i}{d}\right) \text{ and } s \in \mathbb{Z}_d. \tag{4.3}$$

Consequently, given any codespace state $|\psi_L\rangle \in \mathcal{C}$ and set of independent and non-trivial generators $\langle S_1, \ldots, S_{n-k}\rangle$ for $\mathsf{S}(\mathcal{C})$, we have

$$\forall j \in \{1, \ldots, n-k\} : E|\psi_L\rangle = ES_j|\psi_L\rangle = (ES_jE^\dagger)E|\psi_L\rangle = \omega^{s_j} S_j(E|\psi_L\rangle) \text{ for } s_j \in \mathbb{Z}_d. \tag{4.4}$$

Eq. (4.4) implies that, if E does *not* commute with the generators of $\mathsf{S}(\mathcal{C})$, it transforms $\mathcal{C}$ into an *orthogonal* subspace $\mathcal{C}'$ stabilized by $\langle \omega^{s_1} S_1, \ldots, \omega^{s_{n-k}} S_{n-k}\rangle$ for an n-dimensional vector of

d-its $\boldsymbol{s} \coloneqq (s_1, \ldots, s_{n-k}) \neq \mathbf{0}$, known as the *error syndrome*. This further implies that $\mathcal{C}'$ is a simultaneous eigenspace of $\{S_1, \ldots, S_{n-k}\}$ characterised by eigenvalues $\{\omega^{-s_1}, \ldots, \omega^{-s_{n-k}}\}$.

An $[[n, k]]$ stabilizer code $\mathcal{C}$ therefore enables us to detect any Pauli error process $E(\cdot)E^\dagger$ where E does *not* commute with $\mathsf{S}(\mathcal{C})$ by projectively measuring every generator $\{S_1, \ldots, S_{n-k}\}$ of $\mathsf{S}(\mathcal{C})$ in its eigenbasis, which is known as the *syndrome measurement*, and obtaining outcomes $\{\omega^{-s_1}, \ldots, \omega^{-s_{n-k}}\}$ that can be summarised in a *non-zero* error syndrome $\boldsymbol{s} \coloneqq (s_1, \ldots, s_{n-k})$. Crucially, because E transforms $\mathcal{C}$ into a simultaneous eigenspace of $\{S_1, \ldots, S_{n-k}\}$, *the syndrome measurement detects error without disturbing the physical qubits*. The error can be perfectly corrected by applying the channel $F^\dagger(\cdot)F$ where F is an n-qudit Pauli operator such that $FS_jF^\dagger = \omega^{s_j} S_j$ for all j and $F^\dagger E \notin N[\mathsf{S}(\mathcal{C})] - \mathsf{S}(\mathcal{C})$, where $N[\mathsf{S}(\mathcal{C})]$ are n-qubit Pauli operators that commute with $\mathsf{S}(\mathcal{C})$ (i.e. the normalizer of $\mathsf{S}(\mathcal{C})$ in P_n). This works even if F was not the actual error E because, defining the projector onto the codespace

$$P_{\mathcal{C}} \coloneqq \prod_{i=1}^{n-k} \left[\frac{1}{d} \sum_{j=0}^{n-1} (S_j)^n \right], \tag{4.5}$$

we have by by construction that $EP_{\mathcal{C}}E^\dagger = FP_{\mathcal{C}}F^\dagger$ and therefore $F^\dagger E P_{\mathcal{C}} E^\dagger F = P_{\mathcal{C}}$. Since this implies $F^\dagger E \in \mathsf{S}(\mathcal{C})$, we are assured that $F^\dagger E \left|\psi_L\right\rangle = \left|\psi_L\right\rangle$ for any $\left|\psi_L\right\rangle \in \mathcal{C}$, so $F^\dagger$ has corrected the error produced by E.

The most interesting scenario occurs when *every* (non-trivial) Pauli operator on m qudits does not commute with $\mathsf{S}(\mathcal{C})$. For example, every (non-trivial) single-qubit Pauli operator anticommutes with at least one generator of the Steane code given in Eq. (4.1). Concretely, one can verify by direct calculation that single-qubit Pauli operators have the following commutation relations with the Steane code generators.

	$X_4X_5X_6X_7$	$X_2X_3X_6X_7$	$X_1X_3X_5X_7$	$Z_4Z_5Z_6Z_7$	$Z_2Z_3Z_6Z_7$	$Z_1Z_3Z_5Z_7$
X_j	0	0	0	s_1	s_2	s_3
Z_j	s_1	s_2	s_3	0	0	0
Y_j	s_1	s_2	s_3	s_1	s_2	s_3

Table 4.1: Commutation relations between single-qubit Pauli operators on seven qubits and generators of the Steane code, where s_1, s_2 and s_3 be three bits such that $s_1s_2s_3$ is the binary representation of $j \in \{1, \ldots, 7\}$, a bit value of 0 indicates commutation and a bit value of 1 indicates anticommutation.

Because[1] any QECC capable of correcting all Pauli errors on m qubits is capable of correcting

[1]The reasoning is essentially that if a QECC can perfectly correct a set of error processes $\{E_j(\cdot)E_j\}$, it can also

arbitrary errors on m qubits [54], the Steane code is an example of a stabilizer code able to correct an arbitrary single-qubit error.

4.3 CSS code projection protocols for qubit magic distillation

An elementary protocol for magic distillation, first proposed in the seminal work launching the Bravyi-Kitaev model [18], is based on projection onto the codespace of a stabilizer code. Very often, the stabilizer code used will be a CSS code (e.g. see the code projection protocols based on the $[[7,1]]$ Steane and $[[23,1]]$ Golay CSS codes analysed in Ref. [20].) This protocol, known as a *stabilizer code projection*, consists of the following four steps.

Protocol for a stabilizer code projection

1. Take in n copies of a noisy magic state ρ_{in} on a qudit of prime dimension d, i.e. prepare the state $\rho_{\text{in}}^{\otimes n}$.
2. Perform the syndrome measurement of an $[[n,k]]$ stabilizer code $\mathcal{C}$, i.e. projectively measure in the eigenbasis of every generator in a set of $(n-k)$ independent and non-trivial generators for $\mathsf{S}(\mathcal{C})$, and obtain the error syndrome $\boldsymbol{s}$.
3. If errors were found, i.e. if $\boldsymbol{s} \neq \mathbf{0}$, we discard the output state and return to step 1.
4. Otherwise, we accept the output state, decode it onto k qudits and discard the remaining $(n-k)$ syndrome qudits. Concretely, steps 2-3 is equivalent to projecting onto $\mathcal{C}$, so reaching step 4 implies the accepted output state ρ_{acc} lies in $\mathcal{C}$. We can therefore perform the decoding channel to obtain $\mathcal{U}_{\mathcal{C}}^{\dagger}(\rho_{\text{acc}}) = \rho_{\text{out}} \otimes |0\rangle\langle 0|^{\otimes(n-k)}$ for some k-qudit state ρ_{out}, and then discard the final $(n-k)$ qudits.

In the qubit case, an n-to-k CSS code projection protocol can be formalised as a quantum channel $\Phi_{\mathcal{C}} : \mathcal{B}(n) \to \mathcal{B}(k)$ from n to k qubits that acts as

$$\Phi_{\mathcal{C}}(\cdot) \coloneqq \mathrm{Tr}_{k+1,\ldots,n}[\mathcal{U}_{\mathcal{C}}^{\dagger} \circ \Pi_{\mathcal{C}}(\cdot)], \tag{4.6}$$

where $\mathcal{U}_{\mathcal{C}}^{\dagger}$ and $\Pi_{\mathcal{C}} \coloneqq P_{\mathcal{C}}(\cdot)P_{\mathcal{C}}$ are respectively the decoding and codespace projection channels

correct any error process described by Kraus operators that are linear combinations of $\{E_j\}$. Since Pauli operators on m qudits form a basis for all bounded operators on m qudits, being able to correct all Pauli operators implies being able to correct all errors.

for an $[[n, k]]$ CSS code $\mathcal{C}$. Given n copies of a noisy magic state ρ on a single qubit, $\Phi_{\mathcal{C}}$ acts as

$$\Phi_{\mathcal{C}}(\rho_{\text{in}}^{\otimes n}) = p\rho_{\text{out}}, \tag{4.7}$$

where $p \coloneqq \text{Tr}[P_{\mathcal{C}}(\rho_{\text{in}}^{\otimes n})]$ is the probability that the output state from the protocol is accepted (often abbreviated to *acceptance probability*) and ρ_{out} is the resultant magic state on k qubits distilled out by the protocol. Distillation towards a target single-qubit pure magic state ψ is successful if the per-qubit output error defined as $\epsilon_{\text{out}} \coloneqq k^{-1}\left\|\rho_{\text{out}} - \psi^{\otimes k}\right\|_1$ is less than the per-qubit input error defined as $\epsilon_{\text{in}} \coloneqq \|\rho_{\text{in}} - \psi\|_1$. As an example, repeated iterations of code projection protocols using the Steane and Golay CSS codes have been shown to decrease error with respect to the magic state

$$|H\rangle \coloneqq \cos\frac{\pi}{8}|0\rangle + \sin\frac{\pi}{8}|1\rangle, \tag{4.8}$$

which is the $+1$ eigenstate of the Hadamard gate, exponentially quickly in the number of iterations whenever $F_H(\rho) \coloneqq \max_{U\in\mathsf{C}_1}\langle H|U^\dagger\rho U|H\rangle \geq \left[\frac{1}{2}(1+\sqrt{\frac{1}{2}})\right]^{\frac{1}{2}}$.

It has long been known that any n-to-1 qubit magic distillation protocol yields a statistical mixture of stabilizer states and outputs from n-to-1 stabilizer code projection protocols [82]. This result demonstrates that stabilizer code projection protocols provide the core machinery that makes qubit magic distillation possible at all in the n-to-1 case, and further implies that optimal fidelity with respect to a target magic state (though not necessarily optimal acceptance probability) is always achieved by a stabilizer code projection protocol. In a similar way, we can show (Theorem 83) that any CSS circuit carrying out an n-to-k magic distillation protocol can be decomposed into CSS code projection protocols followed by completely CSS-preserving post-processing (the proof line we give also allows one to generalise the result of Ref. [82] to arbitrary n-k qubit magic distillation protocols). Many state-of-the-art protocols are based on CSS codes, such as the celebrated 15-to-1 protocol [18] constructed from the $[[15, 1]]$ punctured Reed-Muller code [83, 84].

4.3.1 Trace-preserving CSS code projections and their stochasticity

Because Theorem 30 was derived for *trace-preserving* magic distillation protocols, we cannot directly apply it to a CSS code projection $\Phi_{\mathcal{C}}$ as defined in Eq. (4.6). However, this can be

remedied by preparing a specially designated CSS state σ on k qubits whenever the output of an n-to-k CSS code projection protocol is rejected, while continuing to distinguish between accepted (labelled '0') and rejected (labelled '1') runs of the protocol by recording this information in an ancillary qubit. We can therefore extend $\Phi_{\mathcal{C}}$ into the following trace-preserving operation $\Lambda_{\mathcal{C}} : \mathcal{B}(n) \to \mathcal{B}(k+1)$ from n to $(k+1)$ qubits

$$\begin{aligned}\Lambda_{\mathcal{C}}(\cdot) &\coloneqq \Phi_{\mathcal{C}}(\cdot) \otimes |0\rangle\langle 0| + \mathrm{Tr}\big[\overline{\Pi}_{\mathcal{C}}(\cdot)\big]\sigma \otimes |1\rangle\langle 1| \\ &= \mathrm{Tr}_{k+1,\ldots,n}[\mathcal{U}_{\mathcal{C}}^{\dagger} \circ \Pi_{\mathcal{C}}(\cdot)] \otimes |0\rangle\langle 0| + \mathrm{Tr}\big[\overline{\Pi}_{\mathcal{C}}(\cdot)\big]\sigma \otimes |1\rangle\langle 1| \end{aligned} \tag{4.9}$$

where once again $\mathcal{U}_{\mathcal{C}}^{\dagger}$ and $\Pi_{\mathcal{C}}$ are the decoding and codespace projection channels of an $[[n,k]]$ qubit CSS code $\mathcal{C}$, $\overline{\Pi}_{\mathcal{C}} \coloneqq (I - P_{\mathcal{C}})(\cdot)(I - P_{\mathcal{C}})$ is the projection channel onto the orthogonal complement of $\mathcal{C}$, and σ is an arbitrary k-qubit CSS state. We conclude that there exists an n-to-k qubit CSS code projection $\Phi_{\mathcal{C}}$ such that $\Phi_{\mathcal{C}}(\rho_{\mathrm{in}}^{\otimes n}) = p\rho_{\mathrm{out}}$ if and only if there exists a trace-preserving n-to-k qubit CSS code projection $\Lambda_{\mathcal{C}}$ as defined in Eq. (4.9) such that

$$\Lambda_{\mathcal{C}}[\rho_{\mathrm{in}}^{\otimes n}] = p\rho_{\mathrm{out}} \otimes |0\rangle\langle 0| + (1-p)\sigma \otimes |1\rangle\langle 1| \coloneqq \rho_p. \tag{4.10}$$

Crucially, trace-preserving qubit CSS code projections defined according to Eq. (4.9) can be implemented as a CSS circuit (see Appendix B.1 for proof), which leads to the following Lemma.

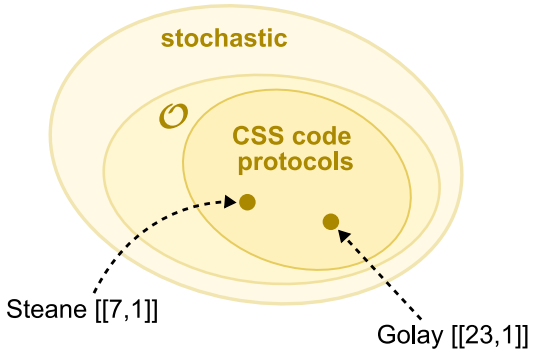


Figure 4.2: **(Schematic of our approach)**. We find that the set of completely CSS-preserving protocols $\mathcal{O}$ are stochastically represented. Such protocols contain the family of CSS code projections as a subset, examples of which include 7-1 and 23-1 protocols based respectively on the Steane $[[7,1]]$ and Golay $[[23,1]]$ codes [20].

Lemma 32. *Every trace-preserving qubit CSS code projection can be executed as a CSS circuit, and is therefore stochastically represented by the Wigner representation defined in Eq. (3.18).*

4.3.2 Unital reference process for CSS code projection protocols

In order to apply Theorem 30 to trace-preserving qubit CSS code projection protocols, we must first identify a *reference process* common to all such protocols. Concretely, we must find a single-qubit CSS state τ_{in} whose Wigner representation is full-rank (to ensure $\text{supp}(W_{\rho_{\text{in}}}) \subseteq \text{supp}(W_{\tau_{\text{in}}})$ for any single-rebit state ρ_{in}) and a $(k+1)$-qubit CSS state τ_{out} such that, given any $[[n,k]]$ qubit CSS code $\mathcal{C}$,

$$\tau_{\text{in}}^{\otimes n} \xrightarrow{\Lambda_{\mathcal{C}}} \tau_{\text{out}}. \tag{4.11}$$

A natural reference process can be chosen for trace-preserving CSS code projection protocols by exploiting the fact that their acceptance components (describing the case in which the protocol output is accepted) are *sub-unital*. To see this, we first note that the identity operator on n qubits can be decomposed as $\mathbb{1}^{\otimes n} = P_{\mathcal{C}} + \overline{P}_{\mathcal{C}}$ for the codespace projector $P_{\mathcal{C}}$ of *any* $[[n,k]]$ CSS code $\mathcal{C}$. Since $P_{\mathcal{C}}$ and $\overline{P}_{\mathcal{C}}$ project onto orthogonal subspaces by definition, the acceptance component $\Phi_{\mathcal{C}}$ in the trace-preserving code projection of $\mathcal{C}$ acts as

$$\Phi_{\mathcal{C}}(\mathbb{1}^{\otimes n}) = \Phi_{\mathcal{C}}(P_{\mathcal{C}} + \overline{P}_{\mathcal{C}}) = \Phi_{\mathcal{C}}(P_{\mathcal{C}}) = \text{Tr}_{k+1,\ldots,n}[\mathcal{U}_{\mathcal{C}}^{\dagger}(P_{\mathcal{C}})] \tag{4.12}$$

Since $P_{\mathcal{C}}$ is the logical identity operator on the codespace $\mathcal{C}$, i.e.

$$P_{\mathcal{C}} = \sum_{\boldsymbol{k}\in\{0,1\}^k} |\boldsymbol{k}_L\rangle\langle\boldsymbol{k}_L| \equiv I_L, \tag{4.13}$$

the decoding of $\mathcal{P}_{\mathcal{C}}$ in Eq. (4.12) must produce the identity operator on k qubits, i.e.

$$\begin{aligned}\forall \boldsymbol{k} \in \mathcal{H}_2^{\otimes k} : U_{\mathcal{C}}^{\dagger} |\boldsymbol{k}_L\rangle = |\boldsymbol{k}\rangle \otimes |0\rangle^{\otimes(n-k)} \implies \mathcal{U}_{\mathcal{C}}^{\dagger}(P_{\mathcal{C}}) &= \left(\sum_{\boldsymbol{k}\in\{0,1\}^k} |\boldsymbol{k}\rangle\langle\boldsymbol{k}|\right) \otimes |0\rangle\langle 0|^{\otimes(n-k)} \\ &= \mathbb{1}^{\otimes k} \otimes |0\rangle\langle 0|^{\otimes(n-k)}. \end{aligned} \tag{4.14}$$

Substituting into Eq. (4.12), we confirm that the acceptance component of any trace-preserving qubit CSS code projection is unital, so the acceptance component takes n copies of the (single-

qubit) maximally mixed state to k copies with probability 2^{k-n} because

$$\Phi_{\mathcal{C}}(\mathbb{1}^{\otimes n}) = \mathbb{1}^{\otimes k} \implies \Phi_{\mathcal{C}}\left[\left(\frac{\mathbb{1}}{2}\right)^{\otimes n}\right] = 2^{k-n}\left[\left(\frac{\mathbb{1}}{2}\right)^{\otimes k}\right]. \tag{4.15}$$

Consequently, every n-to-k trace-preserving qubit CSS code projection $\Lambda_{\mathcal{C}}$ maps n copies of the (single-qubit) maximally mixed state to

$$\Lambda_{\mathcal{C}}\left[\left(\frac{\mathbb{1}}{2}\right)^{\otimes n}\right] = 2^{k-n}\left(\frac{\mathbb{1}}{2}\right)^{\otimes k} \otimes |0\rangle\langle 0| + (1-2^{k-n})\sigma \otimes |1\rangle\langle 1| =: \tau_{n,k} \tag{4.16}$$

Since the Wigner representation of the maximally mixed state can be verified by direct calculation to be $W_{\frac{\mathbb{1}}{2}} = (\frac{1}{4}, \frac{1}{4}, \frac{1}{4}, \frac{1}{4})$ and is therefore full-rank, we can choose Eq. (4.16) to be our reference process for *all* trace-preserving CSS code projections.

4.4 Entropic constraints on qubit CSS code projections

We now apply Theorem 30 to trace-preserving qubit CSS code projection protocols with the reference process defined by Eq. (4.16), and obtain the following entropic constraints on *all* n-to-k CSS code projection protocols on qubits.

Theorem 33. *If there exists an n-to-k CSS code projection protocol that distils n copies of a noisy rebit magic state ρ_{in} with acceptance probability p into the k-rebit magic state ρ_{out}, then*

$$\forall \alpha \in \mathcal{A} : \Delta D_\alpha \coloneqq n D_\alpha\left(W_{\rho_{\text{in}}}\middle\|W_{\frac{\mathbb{1}}{2}}\right) - D_\alpha\left(W_{\rho_p}||W_{\tau_{n,k}}\right) \geq 0 \tag{4.17}$$

We highlight the satisfying fact that the choice of CSS state σ to prepare when the output from a CSS code projection protocol is rejected has no impact on the entropic conditions in Eq. (4.17).

Lemma 34. *The entropic constraints $\Delta D_\alpha \geq 0$ on n-to-k CSS code projection protocols are independent of which CSS state σ is prepared following failed runs.*

Proof. This Lemma follows from resource-theoretic considerations. Consider a channel Φ on $(k+1)$ qubits that performs a Z-basis measurement on the final qubit and, if the -1 outcome

is obtained, re-prepares the first k qubits in a CSS state ω instead. Thus one can write

$$\Phi_\omega(\cdot) \coloneqq \mathsf{id} \otimes \Pi_0(\cdot) + \omega \operatorname{Tr} \otimes \Pi_1(\cdot), \tag{4.18}$$

where $\Pi_k(\cdot) \coloneqq |k\rangle\langle k| (\cdot) |k\rangle\langle k|$, from which one straightforwardly verifies that $\Phi_\omega(\rho_p)$ and $\Phi_\omega(\tau_{n,k})$ simply replaces σ in ρ_p and $\tau_{n,k}$ respectively by ω. Since Φ is a CSS circuit and therefore stochastically represented by Lemma 19, given any CSS state ω, we have by Property **(D3)** in Lemma 27 that

$$\begin{aligned} D_\alpha \left(W_{\rho_p} || W_{\tau_{n,k}}\right) &\geq D_\alpha \left(W_{\Phi_\omega[\rho_p]} || W_{\Phi_\omega[\tau_{n,k}]}\right) \\ D_\alpha \left(W_{\Phi_\omega(\rho_p)} || W_{\Phi_\omega(\tau_{n,k})}\right) \geq D_\alpha \left(W_{\Phi_\sigma \circ \Phi_\omega(\rho_p)} || W_{\Phi_\sigma \circ \Phi_\omega(\tau_{n,k})}\right) &= D_\alpha \left(W_{\rho_p} || W_{\tau_{n,k}}\right). \end{aligned} \tag{4.19}$$

We therefore conclude that

$$D_\alpha \left(\Phi_\omega(W_{\rho_p}) || \Phi_\omega(W_{\tau_{n,k}})\right) = D_\alpha \left(W_{\rho_p} || W_{\tau_{n,k}}\right), \tag{4.20}$$

for any CSS state ω, so the entropic constraints on qubit CSS code projections are unaffected by changing the choice of CSS state to prepare when the protocol is rejected. □

For convenience, we will take σ to be the maximally mixed state on k qubits to make $\tau_{n,k}$ the product state

$$\tau_{n,k} \coloneqq \left[\frac{\mathbb{1}}{2}\right]^{\otimes k} \otimes (2^{k-n} |0\rangle\langle 0| + (1 - 2^{k-n}) |1\rangle\langle 1|). \tag{4.21}$$

In the rest of this section, we specialise the entropic constraints obtained in Theorem 33 to identifying bounds on the code length used in qubit CSS code projection protocols.

4.4.1 Bounds on the code length of qubit CSS code projection protocols

The entropic constraints on qubit CSS code projection protocols given by Theorem 33 can be used to bound many metrics on their performance as magic distillation protocols. In this section, we apply these constraints to bounding the code length of any CSS code projection protocol that could achieve some target combination of noise reduction and acceptance probability from a given supply of noisy magic states. In some parameter regimes, the new lower bounds

we identify on code length outperform those due to a state-of-the-art approach to constructing magic monotones based on the notion of *robustness* [85], in particular *generalised robustness*[2] [86] and *projective robustness*[3] [87]. Such monotones formalise answers to the question: "how much of a stabilizer state can be mixed into a given state ρ before it becomes another stabilizer state?", where we can intuitively regard ρ as highly resourceful if the answer is "a lot".

We first highlight some properties of the relative entropy difference ΔD_α in the following lemma, proof of which can be found in Appendix B.2.2. To this end, it is helpful to momentarily extend the domain for code length n of ΔD_α to the *continuous* range $n \in [k, \infty]$.

Lemma 35. *Given any noisy input rebit magic state ρ_{in}, output k-rebit magic state ρ_{out}, acceptance probability p and $\alpha \in \mathcal{A}$, we have that*

(i) *ΔD_α is concave over the domain $n \in [k, \infty]$.*

(ii) *ΔD_α becomes negative as $n \to k$, i.e. $\lim_{n\to k} \Delta D_\alpha < 0$.*

(iii) *If $H_\alpha[W_{\rho_{\text{in}}}] > 1$, then we have that ΔD_α also becomes negative in the asymptotic limit, i.e. $\lim_{n\to\infty} \Delta D_\alpha < 0$.*

An immediate consequence of Lemma 35 is that ΔD_α is either negative for all n, which implies *no* CSS code projection protocol can distil ρ_{out} from a supply of ρ_{in} with acceptance probability p, or ΔD_α has one or two roots located at n_L^α and n_U^α. These roots then constitute lower and upper bounds on the code length n of any CSS code projection that can carry out the desired distillation. We formalise these observations in the following theorem.

Theorem 36. *Any n-to-k CSS code projection that can distil out the k-rebit magic state ρ_{out} from a supply of the noisy rebit magic state ρ_{in} at acceptance probability p must have a code length n such that for all $\alpha \in \mathcal{A}$, we have that*

$$n \geq n_L^\alpha := \begin{cases} \inf_n\{n : \Delta D_\alpha \geq 0\} & \text{if } \exists n : \Delta D_\alpha \geq 0, \\ \infty & \text{otherwise.} \end{cases} \tag{4.22}$$

$$n \leq n_U^\alpha := \begin{cases} \sup_n\{n : \Delta D_\alpha \geq 0\} & \text{if } \exists n : \Delta D_\alpha \geq 0, \\ -\infty & \text{otherwise.} \end{cases} \tag{4.23}$$

Moreover, given any α such that $H_\alpha[W_{\rho_{\text{in}}}] > 1$, n_U^α provides a finite upper bound on n.

For sufficiently low k, these bounds can be computed numerically using basic root-finding methods. However, we also find analytic upper and lower bounds on n in Section 4.5.

We emphasise that n in Theorem 36 refers to the code length (related to the resource cost C by $C = \frac{n}{pk}$) in a single run of a distillation protocol, as opposed to the the asymptotic overhead. However, single-run n still constitutes a useful metric for analysing the actual resource cost of a given stage of a protocol. Moreover, distillation costs are typically dominated by the final round of a multi-stage distillation protocol (see Ref. [53] and references contained therein), so we expect the above bounds to be particularly informative in this context.

4.4.2 Example application: Hadamard distillation

We can apply Theorem 36 to n-to-1 rebit distillation carried out by CSS code projection protocols targeting the Hadamard state $|H\rangle$ introduced in Eq. (4.8). Regardless of which error processes affect our initial supply of noisy single-qubit magic states, it is sufficient to consider *partially-depolarised* input states of the form

$$\rho(\epsilon_{\text{in}}) \coloneqq (1-\epsilon_{\text{in}})\,|H\rangle\langle H| + \epsilon_{\text{in}}\frac{\mathbb{1}}{2}. \tag{4.24}$$

where the input error $\epsilon_{\text{in}} \in [0,1]$ is the depolarisation noise, because any noisy input magic state ρ_{in} can be converted into this canonical form[4] by fault-tolerant stabilizer operations. Somewhat less obviously, it is also sufficient to consider output states $\rho(\epsilon_{\text{out}})$ of the same canonical form for output error $\epsilon_{\text{out}} \in [0,1]$. This is largely because, if there is an n-to-1 CSS code projection protocol whose accepted run sends n copies of $\rho(\epsilon_{\text{in}})$ to ρ_{out} with acceptance probability p, then there exists another CSS code projection protocol whose accepted run sends the same input to $H\rho_{\text{out}}H$ with the same acceptance probability (see Appendix B.3 for proof and further details). Above a threshold input error $\epsilon_{\text{in}} \geq 1-\frac{1}{\sqrt{2}} \approx 0.3$, the input state $\rho(\epsilon_{\text{in}})$ becomes CSS and no longer magic.

[3]The projective robustness of a state ρ on quantum system S is $\Omega(\rho) \coloneqq \min_{\sigma\in\mathcal{D}_{\text{STAB}(S)}} R_{\max}(\rho||\sigma)R_{\max}(\sigma||\rho)$, where $R_{\max}$ is the non-logarithmic *max-relative entropy* defined as $R_{\max} \coloneqq \min\{\lambda|\rho \leq \lambda\sigma\}$ [88].

[3]The generalised robustness of a state ρ on a quantum system S is $\Lambda^{+}(\rho) \coloneqq \min_{\lambda\geq 1,\tau\in\mathcal{D}(S)}\{\lambda|\frac{1}{\lambda}\rho + (1-\frac{1}{\lambda})\tau \in \mathcal{D}_{\text{STAB}}(\mathcal{H})\}$, where $\mathcal{D}_{\text{STAB}(S)}$ is the set of stabilizer states on S.

[4]Concretely, the channel $\mathcal{G}_H$ executing gates from the group generated by the single Hadamard gate at random,

$$\mathcal{G}_H(\rho_{\text{in}}) \coloneqq \frac{1}{2}(\mathbb{1}\rho_{\text{in}}\mathbb{1} + H\rho_{\text{in}}H), \tag{4.25}$$

projects ρ_{in} onto the Bloch sphere axis from $|H\rangle$ to its orthogonal state $|\overline{H}\rangle$, resulting in $\rho(\epsilon_{\text{in}})$ for some error probability ϵ_{in} when the fidelity of ρ_{in} with respect to $|H\rangle$ is greater than that with respect to $|\overline{H}\rangle$, and $\overline{\rho}(\epsilon_{\text{in}}) \coloneqq (1-\epsilon_{\text{in}})\,|\overline{H}\rangle\langle\overline{H}| + \epsilon_{\text{in}}\frac{\mathbb{1}}{2}$ otherwise. In the latter case, one can perform the gate XZ to take $\overline{\rho}(\epsilon_{\text{in}})$ to $\rho(\epsilon_{\text{in}})$.

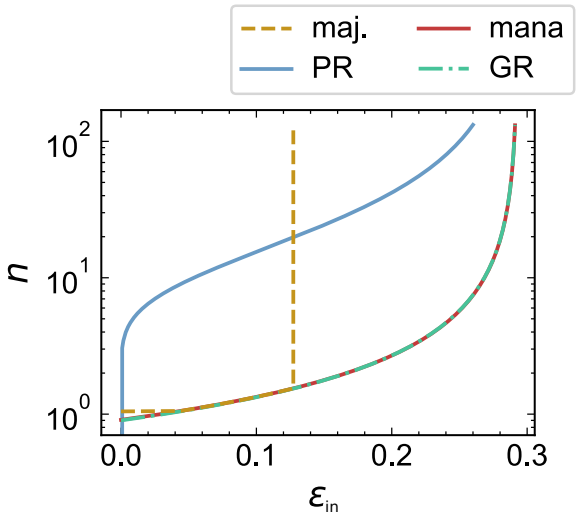


Figure 4.3: **(Lower bound comparison [by Rhea Alexander]).** We plot lower bounds on the number of copies n of the noisy Hadamard state $\rho(\epsilon_{\text{in}})$ required to distil a single Hadamard state $\rho(\epsilon_{\text{out}})$ with output error rate $\epsilon_{\text{out}} = 10^{-9}$ and acceptance probability $p = 0.9$ under a CSS code projection protocol as a function of input error rate ϵ_{in}. Our tightest lower bound from majorization (maj.) is shown to be tighter than those from mana and generalized robustness (GR). However, it only outperforms the lower bound from projective robustness (PR) at high p, high ϵ_{in}.

In Fig. 4.3, we plot the tightest[5] lower bound produced by Theorem 36 on the code length of n-to-1 CSS code projection protocols carrying out Hadamard distillation. In all parameter regimes, our lower bound is tighter than that produced by mana generalized robustness. Furthermore, in the high p, high ϵ regime our lower bound gives tighter constraints than projective robustness [87]. In particular, there is a cut-off input error rate $\epsilon \approx 0.12$ at which our lower bound shoots up to infinity because, for any input error greater than this cut-off, one can always find some α such that $\Delta D_\alpha < 0$ for all n, so no CSS code projection can carry out the desired distillation given a higher input error rate.

In the low p regime, our *upper bounds* are still able to give additional constraints on code length beyond those given by the projective robustness bound. In particular, Fig. 4.4 puts together information from the entropic upper bounds and the lower bound from projective robustness to show that no CSS code projection can achieve some target combinations of output error and acceptance probability beyond a cut-off input error rate.

[5] as numerically identified by basin-hopping in α.

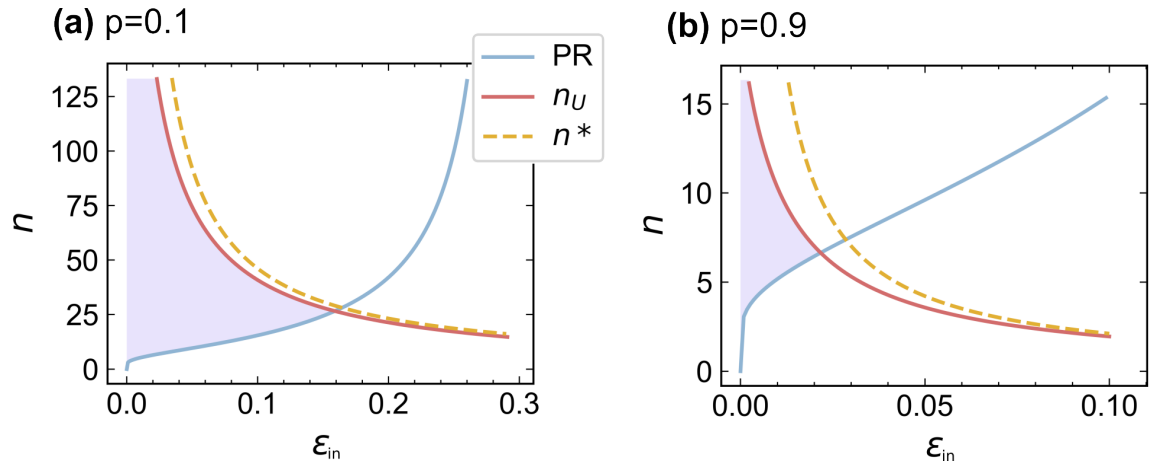


Figure 4.4: **(Finite range on CSS code lengths for magic state distillation protocols [by Rhea Alexander]).** We plot upper and lower bounds on the number of copies n of the noisy Hadamard state $\rho(\epsilon_{\text{in}})$ required to distil a single output state $\rho(\epsilon_{\text{out}})$ with output error rate $\epsilon_{\text{out}} = 10^{-9}$ by projecting onto an $[[n, 1]]$ CSS code. The shaded purple region shows the range of code lengths allowed by the tightest numeric upper bound (red curve) from Theorem 36 and the lower bound from projective robustness (PR) (blue curve). The analytic upper bound n^* (dashed yellow curve) defined in Eq. (4.28) is shown to form a good approximation to the numeric bound. **(a)** When target acceptance probability p is low ($p = 0.1$) the upper bounds are less constraining; **(b)** By increasing to $p = 0.9$, the upper bounds become considerably tighter. In both cases, there is a cut-off input error ϵ beyond which no CSS code projection protocol can achieve the desired combination of output error and acceptance probability.

4.5 Extension to non-qubit stabilizer code projections

Because Gross's Wigner representation stochastically represents *every* stabilizer operation for qudits of odd prime dimension [60], we can extend Theorem 33 to all stabilizer code projection protocols on qudits of odd prime dimension d (see proof at the start of Appendix B.4).

Theorem 37. *If there exists an n-to-k stabilizer code projection protocol for qudits of odd prime dimension d that distils n copies of a noisy magic state ρ_{in} with acceptance probability p into the k-qudit magic state ρ_{out}, then we have that*

$$\forall \alpha \in \mathcal{A} : \Delta D_\alpha := n D_\alpha \left(W^{(G)}_{\rho_{\text{in}}} \middle|\middle| W^{(G)}_{\frac{\mathbb{1}}{d}} \right) - D_\alpha \left(W^{(G)}_{\rho_p} || W^{(G)}_{\tau_{n,k}} \right) \geq 0, \tag{4.26}$$

where $W^{(G)}$ denote Gross's Wigner representation.

The concavity of ΔD_α in code length n, which leads to upper bounds on n, is not peculiar to qubits and leads to the following analytic bounds on code length (proof in Appendix B.4).

Theorem 38 (Qudit code length bounds)**.** *Consider the distillation of a target pure magic state ψ from a supply of the noisy magic state ρ_{in}, where ψ and ρ_{in} are states of a single rebit (qudit of odd prime dimension d). Any n-to-k CSS (stabilizer) code projection protocol that distils n copies of ρ_{in} into a k-rebit (qudit) state ρ_{out} with acceptance probability p and per-rebit (qudit) output error ϵ_{out} must have a code length n such that*

$$n \geq \frac{k\left[\log d - H_\alpha(W_\psi)\right] - \frac{\alpha}{1-\alpha}\log\left(\frac{p}{1+k\epsilon_{\text{out}}d^{5/2}}\right)}{\left[\log d - H_\alpha(W_{\rho_{\text{in}}})\right]}, \tag{4.27}$$

for all $\alpha \in \mathcal{A}$ for which $H_\alpha(W_{\rho_{\text{in}}}) < \log d$, and

$$n \leq \frac{k\left[H_\alpha(W_\psi) - \log d\right] + \frac{\alpha}{1-\alpha}\log\left(\frac{p}{1+k\epsilon_{\text{out}}d^{5/2}}\right)}{\left[H_\alpha(W_{\rho_{\text{in}}}) - \log d\right]}, \tag{4.28}$$

for all $\alpha \in \mathcal{A}$ for which $H_\alpha(W_{\rho_{\text{in}}}) > \log d$, where W is the qubit Wigner representation defined in Eq. (3.18) (Gross's Wigner representation).

One might be concerned that the conditions $H_\alpha(W_{\rho_{\text{in}}}) > \log d$ given in Theorems 36 and 38 for the existence of a finite upper bound on n are never actually satisfied. Fortunately, this turns out not to be the case. As an example, we return to n-to-1 qubit CSS code projection protocols carrying out Hadamard distillation examined from a supply of partially depolarised Hadamard states $\rho(\epsilon_{\text{in}})$. The partial depolarisation channel $\Lambda_{\epsilon_{\text{in}}}$ on a single qubit, which re-prepares that qubit in the CSS maximally mixed state with probability ϵ_{in} and does nothing otherwise, is a unital quantum channel that can be carried out as a stabilizer circuit. Consequently, we can transform $W_{\rho(0)}$ into $W_{\rho(\epsilon_{\text{in}})}$ by the stochastic representation $W_{\Lambda_{\epsilon_{\text{in}}}}$ of the partial depolarisation channel, which also preserves the uniform distribution over phase space that constitutes the Wigner representation of the maximally mixed state. Since H_α for $\alpha \in \mathcal{A}$ is well-defined on quasiprobability distributions and cannot decrease under their stochastic processing when the uniform distribution over phase space is preserved (Theorem 22), we can upper-bound the α-Rényi entropy of $W_{\rho(\epsilon_{\text{in}})}$ as

$$\forall a \in \mathcal{A} : H_\alpha\left[W_{\rho(\epsilon_{\text{in}})}\right] \geq H_\alpha\left[W_{\rho(0)}\right] \tag{4.29}$$

Fig. 4.5 shows that there exists α for which $H_\alpha > 1$ even at $\epsilon_{\text{in}} = 0$, and therefore at any ϵ_{in}.

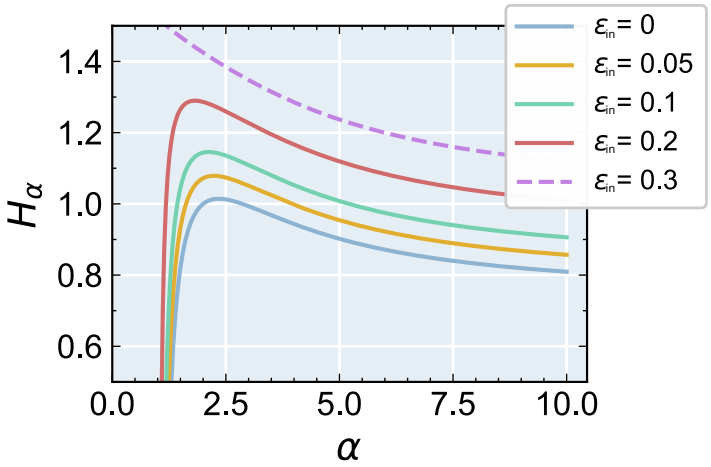


Figure 4.5: **(Wigner-Rényi entropies & magic distillation [by Rhea Alexander])** We examine the condition $H_\alpha[W_{\rho(\epsilon_{\text{in}})}] > 1$ in Theorem 36 for the existence of finite upper bounds on n for n-to-1 qubit CSS code projection protocols carrying out Hadamard distillation. Even in the limit of zero input error $\epsilon_{\text{in}} = 0$, there exist α at which this condition is met, with H_α attaining a maximum of ~ 1.012 at $\alpha \approx 2.2$. Therefore Hadamard distillation under n-to-1 qubit CSS code projection protocols is ruled out in the asymptotic limit $n \to \infty$ at any target acceptance probability and output error. We highlight that at the error rate ~ 0.3 (dashed curve) beyond which $\rho(\epsilon_{\text{in}})$ is no longer magic, H_α starts satisfying the standard property of monotonically decreasing with α because $W_{\rho(\epsilon_{\text{in}})}$ has become a proper probability distribution. We also highlight that the $\alpha \to 1$ divergence corresponds to a pole in $H_\alpha[W_{\rho_{\text{in}}}]$ for magic state ρ, and its residue is mana.

By evaluating the $\alpha = 2$ condition explicitly, we find that, if there exists a CSS code projection that distils n copies of the ϵ_{in}-depolarised Hadamard state $\rho(\epsilon_{\text{in}})$ into a k-rebit state ρ_{out} with acceptance probability p and per-qubit output error ϵ_{out}, then n is upper-bounded as

$$n \leq n^* := 2 \log_{f(\epsilon_{\text{in}})} \left[\frac{1 + 6k\epsilon_{\text{out}}}{p} \right], \tag{4.30}$$

where the logarithm base is $f(\epsilon_{\text{in}}) := \left[1 - \epsilon_{\text{in}} + \frac{\epsilon_{\text{in}}^2}{2}\right]^{-1}$. This expression reveals that, if we use CSS code projection protocols for Hadamard distillation, we must accept a fundamental trade-off between acceptance probability and output fidelity. In particular, Eq. (4.30) shows that given a supply of noisy magic states ($\epsilon_{\text{in}} > 0$), no CSS code projection can distil a perfect Hadamard state ($\epsilon_{\text{out}} = 0$) with certainty ($p = 1$), which was first demonstrated in Ref. [78].

This example highlights that Theorem 38, and indeed the stronger results in Theorem 36, need not simply be viewed providing bounds on code length n, but as *trade-off relations* between the three parameters that govern the performance of a stabilizer code projection protocol at distilling a target ψ from a supply ρ_{in}: the code length n, the acceptance probability p and the

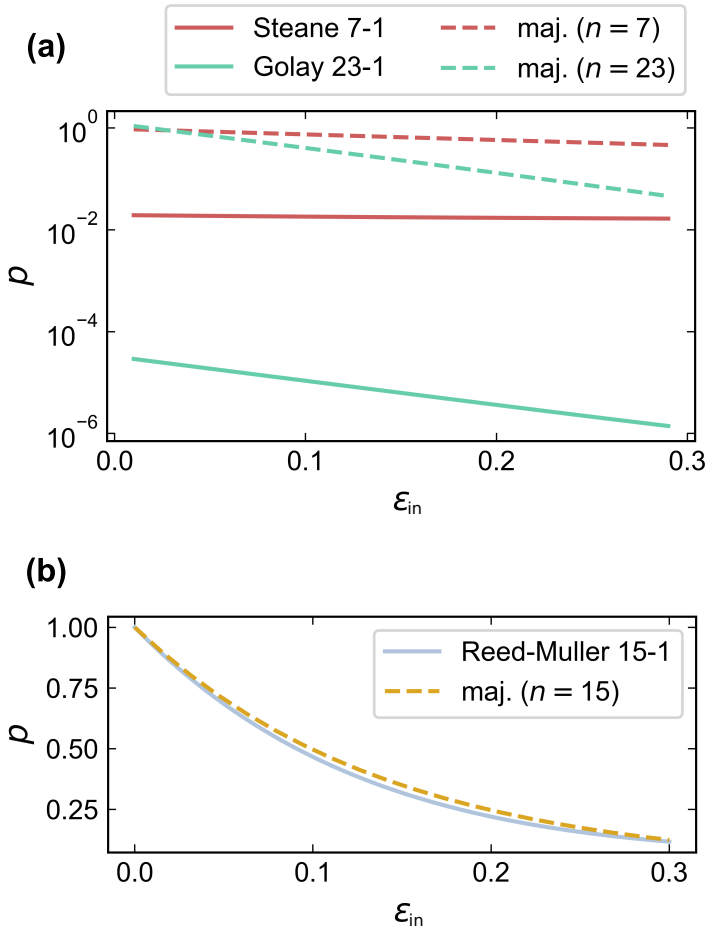


Figure 4.6: **(Explicit protocol comparison [by Rhea Alexander]).** (a) We compare the entropic upper bounds (dashed lines) on the acceptance probability p with which one can distil a noisy Hadamard state $\rho(\epsilon_{\text{in}})$ via an n-to-1 code projection against actual acceptance probabilities attained using the Steane code (purple) at $n = 7$ and the Golay code (green) at $n = 23$ (detailed in Ref.[20]). Attained acceptance probabilities are orders of magnitude less than our upper bounds. (b) We plot the entropic upper bound (dashed line) on the acceptance probability p of any 15-to-1 CSS code projection protocol with which one can distil the noisy magic state $(1-\epsilon)\left|A\right\rangle\!\left\langle A\right| + \epsilon\frac{\mathbb{1}}{2}$. Interestingly, our bound is very close to the actual acceptance probability for the 15-to-1 protocol (blue line) given in [18], though we emphasise this latter protocol is *not* just a straightforward CSS code projection.

output error ϵ_{out}.

As an example, Eq. (4.27) and Eq. (4.28) can be rearranged to upper-bound the acceptance probability p at a target code length and output error. In Fig. 4.6, we compare the tightest upper bound on p produced by Theorem 38 for n-to-1 Hadamard distillation to those attained in existing protocols based on CSS codes given in Refs. [18] and [20]. We see that the acceptance probabilities of code projection protocols using the $[[7,1]]$ Steane and $[[23,1]]$ Golay codes in Fig. 4.6(a) are orders of magnitude less than our upper bounds, suggesting that substantial room for improvement is not ruled out. Interestingly, in Fig. 4.6(b) our upper bound is very

close to the actual acceptance probability of the protocol based on the $[[15, 1]]$ Reed-Muller code in Ref. [18], which we speculate may hint at something fundamental about the role of the intermediate Clifford corrections used in that protocol. To further investigate this trade-off, in Fig. 4.7(b) we plot the *maximum output fidelity* with respect to the Hadamard state $F_{\max}$ that can be achieved by any n-to-1 CSS code projection according to Theorem 36.

4.5.1 Why do we expect upper bounds on code length?

The appearance of upper bounds on the number of input copies n might first seem to contradict a resource-theoretic perspective. Concretely, if we take CSS (stabilizer) operations on qubits (qudits of odd prime dimension d) to be our free operations, we might expect $n + 1$ copies of a noisy magic state ρ_{in} to be at least as good as n copies at distilling magic, since discarding is itself a CSS (stabilizer) operation.

However, by specialising to code projection protocols, we necessarily introduce a trade-off between n and acceptance probability p, which can be seen from the following back-of-the-envelope calculation. For any n-to-k stabilizer code projection protocol used for d-dimensional qudit distillation, p is given by how much of n copies of the noisy input magic state ρ_{in} projects onto the d^k-dimensional codespace spanned by the logical basis $\{|\boldsymbol{j}_L\rangle\}_{j=0,\ldots,d^k-1}$ of the code. Letting $\lambda_{\max}(\cdot)$ denote a state's largest eigenvalue, we find the following upper bound on p,

$$p = \sum_{j=0}^{d^k-1} \langle \boldsymbol{j}_L | \, \rho_{\text{in}}^{\otimes n} \, | \boldsymbol{j}_L \rangle \leq d^k \lambda_{\max}(\rho^{\otimes n}) = d^k [\lambda_{\max}(\rho_{\text{in}})]^n, \tag{4.31}$$

which falls monotonically towards 0 as $n \to \infty$ whenever ρ_{in} is impure.

At an intuitive level, the trade-off between n and p occurs because the codespaces of $[[n, k]]$ stabilizer codes remain the same size as we increase n, and so take up a vanishingly small part in the support of all the noisy input magic states used. Under the requirement that we have some threshold acceptance probability (below which the expected overhead would be too large), a corresponding upper bound on n is then expected.

4.5.2 Comparison with the data-processing inequality (DPI)

We have seen how upper bounds on code length such as those identified by Theorem 33 constitute a generic feature of code projection protocols (once other parameters governing distil-

lation performance are fixed). As the derivation of these upper bounds encoded the restriction to code projection protocols by demanding the reference process $\left(\frac{\mathbb{1}}{d}\right)^{\otimes n} \to \tau_{n,k}$, we can ask how far this demand *alone* accounts for the strength of these upper bounds.

Concretely, we can invoke the *data-processing inequality* (DPI) of quantum channels, which states that, if there exists *any* quantum channel, stabilizer or otherwise, that can carry out a desired distillation process $\rho_{\text{in}} \to \rho_p$ *and* the reference process $\left(\frac{\mathbb{1}}{d}\right)^{\otimes n} \to \tau_{n,k}$, then

$$\Delta\tilde{D}_\alpha \coloneqq n\tilde{D}_\alpha\left(\rho_{\text{in}}\middle\|\frac{\mathbb{1}}{d}\right) - \tilde{D}_\alpha\left(\rho_p\|\tau_{n,k}\right) \geq 0, \tag{4.32}$$

for $\alpha \geq \frac{1}{2}$ [89], where $\tilde{D}_\alpha(\rho\|\tau)$ is the sandwiched α-Rényi divergence [90, 91] on the normalized quantum states ρ and τ, which is defined as

$$\tilde{D}_\alpha(\rho\|\tau) \coloneqq \begin{cases} \frac{1}{\alpha-1}\log \text{Tr}\left[\left(\tau^{\frac{1-\alpha}{2\alpha}}\rho\tau^{\frac{1-\alpha}{2\alpha}}\right)^\alpha\right] \text{ for } \alpha \in (1,\infty) & \text{for } \alpha \in (0,1)\cup(1,\infty) \\ \text{Tr}[\rho\log\rho - \log\sigma] & \text{for } \alpha = 1 \\ \log\left\|\sigma^{-\frac{1}{2}}\rho\sigma^{-\frac{1}{2}}\right\|_\infty & \text{for } \alpha = \infty. \end{cases} \tag{4.33}$$

Because $\left(\rho, \frac{\mathbb{1}}{d}\right)$ are simultaneously diagonalisable in the eigenbasis of ρ_{in}, and $(\rho_p, \tau_{n,k})$ are simultaneously diagonalisable in $\{\text{eigenbasis}(\rho_{\text{out}})\otimes|0\rangle\langle 0|, \text{eigenbasis}(\rho_{\text{out}})\otimes|1\rangle\langle 1|\}$, we can in fact replace the sandwiched α-Rényi divergences in the above equation by *(classical) α-Rényi divergences* evaluated on *eigenvalue distributions*. Letting $\boldsymbol{\lambda}(\cdot)$ denote a probability distribution formed from the eigenvalues, we have that

$$\Delta\tilde{D}_\alpha = nD_\alpha\left(\boldsymbol{\lambda}(\rho_{\text{in}})\middle\|\boldsymbol{\lambda}\left(\frac{\mathbb{1}}{d}\right)\right) - D_\alpha\left(\boldsymbol{\lambda}(\rho_p)\|\boldsymbol{\lambda}(\tau_{n,k})\right) \geq 0. \tag{4.34}$$

Following a similar proof strategy to Lemma 35, we find that, just like the majorization constraint $\Delta D_\alpha \geq 0$, the DPI constraint $\Delta\tilde{D}_\alpha \geq 0$ above also generates upper bounds on code length n for a given supply of noisy input magic state ρ_{in} and a combined target of output magic state ρ_{out} and acceptance probability p. This motivates defining the following DPI upper bound on n:

$$n \leq n^{DPI} \coloneqq \min_{\alpha\geq\frac{1}{2}} \max_n \{n : \Delta\tilde{D}_\alpha \geq 0\}. \tag{4.35}$$

Loosely speaking, by examining $\Delta n_U := n^{maj} - n^{DPI}$ (the extent to which majorization "beats" DPI), where n^{maj} is the tightest upper bound due to majorization from Theorem 36, we can see whether stochasticity imposes additional constraints beyond demanding a quantum channel sufficiently similar to code projection as to reproduce the reference process $\left(\frac{\mathbb{1}}{d}\right)^{\otimes n} \to \tau_{n,k}$.

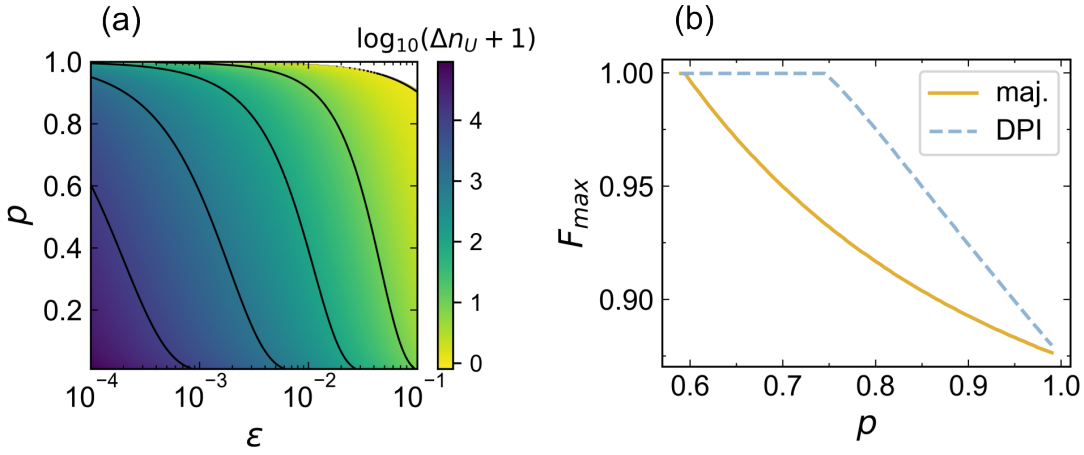


Figure 4.7: **(Majorization gives independent constraints over DPI [by Rhea Alexander]). (a)** Variation of (scaled) $\Delta n_U := n_U^{DPI} - n_U^{maj}$ over all possible values of acceptance probability p and a realistic range of input error ϵ, with fixed $k\epsilon_{\text{out}} = 10^{-9}$. Whenever $\log_{10}(\Delta n_U + 1) > 0$, upper bounds on code length from majorization give tighter constraints than those from DPI, reaching $\Delta n_U = O(10^4)$ in the low p, low ϵ regime. **(b)** We show the trade-off relation given by bounds on the maximum achievable fidelity $F_{\max}(\rho)$ vs. target acceptance probability p, under an n-1 CSS code projection, where $\rho = \frac{3}{4}|H\rangle\langle H| + \frac{1}{8}\mathbb{1}$. For p above a given threshold (≈ 0.6) no perfect distillation is theoretically possible, even for $n \to \infty$ copies of the input state. Majorization (maj.) is shown to give stronger constraints than that of DPI.

Our collaborator Rhea Alexander answered this question in the affirmative in Fig. 4.7: Δn_U is positive over a wide range of parameter regimes. In particular, the low acceptance probability p and low input error ϵ regime of Fig. 4.7(a) produces differences in upper bounds $\Delta n_U := n_U^{DPI} - n_U^{maj}$ (the amount by which majorization "beats" DPI) of the order $\Delta n_U = O(10^4)$. We thus conclude that the constraints on CSS protocols stemming from majorization go beyond those from the DPI.

4.6 Concluding Remarks

In Part II of this thesis, we extended the statistical mechanics framework introduced in Ref. [27] for QCSI on qudits of odd prime dimensions to the technologically significant case of qubits. To achieve this, we made use of a Wigner representation first introduced in Ref. [72], wherein completely CSS-preserving channels correspond to stochastic transformations on phase spaces associated to qubit systems. These channels include CSS circuits, the subset of stabilizer cir-

cuits wherein CSS states play the role of stabilizer states. These circuits consist precisely of the gateset that can be encoded transversally using surface code constructions [8], and can be promoted to universal quantum computing with the injection of rebit magic states [73, 74]. Within this framework, we leveraged relative majorization to identify entropic constraints on completely CSS-preserving magic distillation protocols. These can be fine-tuned to a reference process singling out physics that is distinctive to a particular class of protocols, or the limitations of hardware on which distillation is carried out. Concretely, we applied these constraints to stabilizer code projection protocols, in terms of which all magic distillation protocols carried out by CSS circuits can be decomposed, and established fundamental trade-off relations between parameters governing their performance.

A natural extension of the research conducted in this part of the thesis would be applying our constraints to more sophisticated magic distillation protocols. For instance, we might ask how the use of intermediate Clifford corrections (similar to those deployed in the 5-1 Bravyi-Kitaev protocol [18]) between measurements of stabilizer generators might affect the trade-offs we found between resource cost and fidelity improvement. To this end, a significant limitation of the statistical mechanics framework we developed to identify such constraints is its restriction to completely CSS-preserving protocols in the qubit case. This originated from the demand for a Wigner representation that represents magic distillation stochastically while also respecting the sequential and parallel composition of processes. As the second requirement was largely introduced in order to simplify the extraction of lower bounds on overhead via results such as Theorem 30, which was complicated in this chapter by the possibility of non-trivial dependence on overhead in the reference output state, future work could consider extending our framework to qubit Wigner representations (such as that developed in Ref. [92]) that do not respect the sequential and parallel composition of processes, but thereby become capable of representing the entire stabilizer subtheory stochastically.

We have also obtained a set of monotones $\{M_\alpha\}$ for completely CSS-preserving magic distillation, each of which forms a convex optimization problem. We speculate that an analogous monotone can be constructed for any resource theory for which the free operations are a subset of operations that completely preserve Wigner positivity. From the perspective of quantum optics experiments, wherein Gaussian operations and probabilistic randomness are readily available, it may be of interest to consider the case of continuous variable systems under the set

of Gaussian operations and statistical mixtures [49]. Since the individual α-Rényi divergences on quasiprobability distributions were seen in Section 4.5.2 to typically produce stronger constraints than the corresponding constraints given by α-Rényi divergences on quantum states, it would be interesting to see how well these quasiprobability distribution-based monotones perform relative to known state-based counterparts.

On a technical note regarding majorization theory, we point to an interesting direction for further study. The Wigner representation of Ref. [72] recovers the covariance over symplectic affine transformations on qubit phase spaces, a property shared by Gross's Wigner function on odd-dimensional systems. This added structure on the phase space was ignored by our analysis, but could be utilised to tighten the obtained bounds in future work. In particular, as explained in the discussion of Ref. [27], the stochastic majorization used in our analysis is only a special case of G-majorization, where G can be taken as a subgroup of the stochastic group such as the symplectic group. It can then be shown [93, 94, 95] that one should expect to obtain a set of finite lower bound constraints on distillation, which will be tighter than stochastic majorization constraints.

Part III

Infinitesimal reference frames suffice to determine the asymmetry properties of a quantum system

Chapter 5

Single Entropic Condition for the Resource Theory of Asymmetry

5.1 Introduction

5.1.1 Symmetry constraints on general quantum channels

Identifying constraints imposed by symmetry on the evolution of a system has broad applications throughout physics. When finding exact solutions to the laws of motion becomes too complex, one can often make non-trivial inferences by appealing to symmetry principles. In classical mechanics, the most prominent example of such an appeal is Noether's Theorem, which enables one to infer conserved quantities from (differentiable) symmetries in the laws of motion, and vice versa [28]. However, Noether's Theorem is restricted to closed system dynamics, and can only be extended to *unitary evolution* in quantum mechanics [29, 30]. How symmetry principles constrain the general evolution of a quantum system via *quantum channels* therefore remains a crucial question.

The evolution of a quantum system possesses symmetry with respect to a group G, represented by unitary channels on that system, when it is restricted to the subset of quantum channels that commute with all group actions, known as *G-covariant channels*. This definition of symmetric dynamics is best understood by example. Consider a (finite-dimensional) quantum system constrained to channels that commute with every possible time evolution of that system (representing the group $\mathsf{U}(1)$). Since this is equivalent to restricting channels to those that can occur

at any time to exactly the same effect, and therefore require no clocks to perform, it makes sense to say that the system evolution possesses time symmetry (formally, is time-covariant).

In general, symmetry with respect to a group G naturally induces the structure of a *resource theory for asymmetry*, or asymmetry theory for short, on the dynamics of quantum systems, wherein G-covariant channels constitute the free operations while all channels that break symmetry become resources. Furthermore, the pre-order on quantum states defined by convertibility under G-covariant channels in asymmetry theory provides a rigorous basis for quantifying the extent to which a quantum system breaks a symmetry [30, 33, 34]. Taking the example of time symmetry again, only states that are not time-symmetric can act as clocks, and a state ρ is at least as good a clock as another state σ if ρ can be converted into σ without bringing in an additional clock. As we have seen, this is equivalent to whether ρ can be converted into σ via a time-covariant channel. The pre-order on quantum states established by asymmetry theory is ultimately the pre-order on how well they encode a group element.

5.1.2 Asymmetry as an information-theoretic resource

Asymmetry is an example of *unspeakable* information [7, 96], which is contrasted with *speakable* information capable of being encoded in any physical system (paradigmatically, spoken over the phone) that constitutes the typical interest of abstract information theory[1]. Why asymmetry must be unspeakable becomes clear once we ask what sorts of information-processing tasks require symmetry-breaking as a resource. Consider a scenario where Alice and Bob are in spaceships so far apart that they have no shared stars. If Alice would like to align her Cartesian frame with Bob's, there is no way for her to simply describe an alignment direction over the phone. Instead, Alice must share a physical system with Bob that points out the desired direction for one axis in their Cartesian frames, which means the system she sends cannot be rotationally symmetric. In this scenario, asymmetry with respect to the group of rotations becomes a resource without which Alice and Bob cannot do anything that requires them to have a shared spatial orientation.

In general, the task of *reference frame alignment*, which encompasses clock synchronisation and global positioning as well as the example of Cartesian frame alignment just discussed, requires the transfer of a physical system that carries a non-trivial action with respect to a symmetry

[1] For instance, the Shannon-Hartley theorem holds equally well for bits communicated by a person saying "yes" or "no", a coin showing heads or tails, or a switch being turned on or off.

group G via a G-covariant quantum channel [7, 96], thereby establishing a shared reference frame with respect to G amongst parties that previously had none. Asymmetry theory can therefore also be regarded as the *resource theory of quantum reference frames* with respect to a symmetry group G. The lack of a quantum reference frame for a degree of freedom is mathematically equivalent to imposing a superselection rule [7] – for instance, lacking a shared phase reference in quantum optics prevents one experimenter from establishing coherences between states with different total photon numbers from the perspective of another experimenter.

Beyond quantifying the quality of reference frames [7, 31, 97, 98, 99, 100], asymmetry is a vital resource for many tasks in quantum metrology [101, 102], such as bounding the error of estimating optical phase shifts. More recently, asymmetry has also been identified as a resource in quantum computing, especially for (partially) circumventing the Eastin-Knill no-go theorem [57] against quantum error-correcting codes that transversally encode and are therefore covariant with respect to a continuous group of gates [103]. As a result, bounds on the correctable error of such codes [104] and approximately covariant codes [57, 105, 103, 106, 107, 108, 109] have been constructed. Beyond symmetry-constrained dynamics [30, 110, 111], asymmetry theory also finds application in thermodynamics [35, 112, 113], measurement theory [114, 115, 116, 117, 118], macroscopic coherence [119], and quantum speed-limits [120].

Of particular interest for this part of the thesis are two current strands of quantum gravity research that draw on tools of asymmetry theory. One strand concerns how transformations between inertial reference frames established by relativity should be updated in light of their ultimately quantum nature, and the extent to which foundational aspects of relativity such as the weak equivalence principle can be recovered afterwards [121, 122]. Another strand identifies single-shot entropies that can capture fundamental relationships between quantum information and the geometry of spacetime in the context of the AdS/CFT duality [107, 108]. Notably, very recent work from this latter strand has identified the conditional max-entropy as the right information measure for the causal development of a spatial region [123]. This part of the thesis builds on research in asymmetry theory [36] that uses the dual form of the conditional max-entropy to quantify the quality of quantum reference frames, and so combines two very active lines of research on foundational topics from quantum gravity.

5.1.3 The search for asymmetry monotones

The central concern of asymmetry theory is finding a complete set of resource monotones, known as *asymmetry monotones*, capable of capturing the pre-order induced on quantum states by G-covariant channels, which would isolate all the dynamical properties governing the behaviour of quantum systems under a symmetry constraint and fully address the fundamental question:

> *When is it possible to transform a state ρ of input system A into a state σ of output system B via a quantum channel that is covariant with respect to a symmetry group G?*

Finding asymmetry monotones for general quantum dynamics (mixed states evolving via channels) turned out to be surprisingly complex and highly counterintuitive. Taking the particular example of rotational symmetry, one might guess from closed system dynamics governed by Noether's Theorem that expectation values (and higher moments[2]) of angular momenta would always be good asymmetry monotones. However, this turns out to be false in general – in fact, expectation values of angular momenta can be *arbitrarily increased* by rotationally covariant channels (such as the optimal universal cloner [124]) from a spin-j_A system to a spin-j_B system by simply taking j_B sufficiently higher than j_A [110]. This example shows why we need information-theoretic monotones that go beyond traditional Noetherian conserved quantities to quantify asymmetry as a resource – while the expectation values of angular momenta can increase under rotationally covariant channels, the degree to which a quantum system can encode a direction in space does not.

Significant progress has been made in recent years in the identification of asymmetry monotones, including relative entropy measures [31, 32], the skew-Fisher information [30, 34, 125] as well as the purity of coherence [35]. Recently, Ref. [36] identified the first complete set of asymmetry monotones, which are expressed in terms of correlations between the input/output quantum system S and a spontaneously emerging quantum reference frame system R locally in a state that transforms non-trivially under the group action. These correlations are measured via the *single-shot quantum conditional min-entropy* $H_{\min}(R|S)$, a central quantity in quantum encryption [126] whose dual form has also been identified in very recent research

[2]The extension of Noether's Theorem to quantum mechanics states that a necessary (and, between pure states, sufficient) condition for the existence of a G-covariant unitary state transition is the conservation of the moments $\langle J_k^n \rangle$ of every generator J_k of the group representation $U(g) := \exp\left(i \sum_k g_k J_k\right)$ [29, 30].

in quantum gravity as the right information measure for the causal development of a spatial region [123]. However, the physical significance of these monotones remains obscure, and the complete set of conditions require $H_{\min}(R|A)$ to not decrease at *every* possible reference frame state η_R, which leads to an infinite number of conditions to check. This opens the question of whether one can reduce these monotones to a simpler, ideally finite, set of conditions for asymmetry theory, and whether such reduction can shed light on the asymmetry properties of quantum systems.

5.1.4 Chapter summary

The rest of this chapter proceeds as follows. In Section 5.2, which forms the main technical background for this chapter, we formally define asymmetry theory and review the first complete set of asymmetry monotones identified by Ref. [36], focussing on the spontaneous emergence of quantum reference frames in these monotones. We then examine a warm-up example in Section 5.3 of time-symmetric state transitions in a non-degenerate qubit, where we find that the original infinite set of asymmetry monotones can be reduced to just two, which are evaluated at reference frames that, contrary to intuition from the theory of quantum reference frames, are infinitesimally close to time-stationary. Guided by the warm-up example, we identify generic reference frame redundancies in sections 5.4 and 5.5 applicable to all symmetry groups and systems. In particular, we find that one can restrict to reference frames forming any closed surface enclosing the maximally mixed state. By taking this surface arbitrarily close to the maximally mixed state, which is always symmetric, we obtain a single necessary and sufficient entropic condition for G-covariant state conversion. Though we find that evaluating this condition poses significant technical difficulties, it can nevertheless be interpreted as generalising an insight from our warm-up example: one does not need reference frames that perfectly encode a symmetry group to characterise the asymmetry properties of a quantum system – in fact, reference frames that are infinitesimally close to being symmetric suffice.

5.2 Technical background

5.2.1 Asymmetry theory of compact symmetry groups

We now formalise the resource theory of asymmetry with respect to a *compact* group G.

Definition 39 (Resource Theory of G-Asymmetry)**.** *Let G be a compact group that, given any quantum system S, has a unitary representation $\{U_S(g)|g \in G\}$ on $\mathcal{H}_S$. Furthermore, let $\mathcal{U}_g^S[\cdot] \coloneqq U_S(g)[\cdot]U_S^\dagger(g)$ denote the unitary channel induced by $U_S(g)$. We say that a quantum channel $\Phi_{A\to B}$ from input system A to output system B is G-covariant when performing this channel before any group action is operationally equivalent to performing this channel after that group action, i.e.*

$$\forall g \in G,\ \rho_A \in \mathcal{D}_A :\ \mathcal{U}_g^B \circ \Phi_{A\to B}(\rho_A) = \Phi_{A\to B} \circ \mathcal{U}_g^A(\rho_A). \tag{5.1}$$

The resource theory of G-asymmetry is defined by taking the set of G-covariant channels as the free operations $\mathcal{R}_G$. A state τ of system S therefore belongs to the set of free states $\mathcal{F}_G$ if and only if it is invariant under all group actions, i.e.

$$\tau \in \mathcal{F}_G \iff \forall g \in G : \mathcal{U}_g(\tau) = \tau. \tag{5.2}$$

We will also find it helpful to define the *G-twirling* operation on a bounded operator O, which constructs a symmetric operator from O_S as follows[3].

Definition 40 (G-twirling)**.** *The G-twirl averages the output of all group actions on a bounded operator O_S of quantum system S as*

$$\mathcal{G}(O_S) \coloneqq \int_G dg\, \mathcal{U}_g^S(O_S). \tag{5.3}$$

When G is finite, $\int_G dg$ is naturally replaced by $\frac{1}{|G|}\sum_{g\in G}$. Though we will use the Lie group notation exclusively from this point forwards, all results apply equally well to finite groups. Throughout this part of the thesis, we will use $\rho \xrightarrow{G} \sigma$ to denote the existence of a G-covariant channel capable of transforming ρ into σ.

[3] In the next chapter, we will see that the G-twirl operator is in fact the projector onto the subspace of $\mathcal{B}(S)$ that is invariant under all group actions, or equivalently the subspace spanned by all trivial representation spaces of G on $\mathcal{B}(S)$.

5.2.1.1 Example: resource theory of time asymmetry ("clockiness") in systems with periodic Hamiltonians

Consider a quantum system S with a Hamiltonian H_S such that there exists a finite minimum length of time $T_S > 0$ at which $e^{-iH_S T_S} = I$, which is known as the period of H_S. Then the time evolution of this system forms the unitary representation $\left\{e^{-iH_S t} \middle| \, t \in [0, T_S)\right\}$ for the group $\mathsf{U}(1)$ on $\mathcal{H}_S$. Any free state τ_S in the time asymmetry theory of system S must therefore be *time-stationary*, i.e.

$$\forall t \in [0, T_S) : e^{-iH_S t} \tau_S e^{iH_S t} = \tau_S \iff [H_S, \tau_S] = 0 \tag{5.4}$$

so τ_S must be a *mixture of energy eigenstates* for H_S. Free operations in the time asymmetry theory of systems with periodic Hamiltonians, i.e. their time-covariant channels, must therefore take mixtures of energy eigenstates on the input system to mixtures of energy eigenstates on the output system, making time asymmetry theory in such systems equivalent to a *resource theory of coherence in energy*.

5.2.2 Quantum reference frames

Asymmetry is ultimately a *relational* quantity. For instance, the direction in space encoded by a rotationally asymmetric system (e.g. a spin-$\frac{1}{2}$ particle in a pure state) is not defined relative to "absolute space" but another physical system (e.g. a set of gyroscopes); similarly, the phase encoded in a time-asymmetric system (e.g. an atom in coherent superposition between two eigenstates of different energies) is not defined relative to "absolute time" but a physical clock [7]. The systems with respect to which unspeakable asymmetry information is defined, such as clocks, gyroscopes and metre sticks, are known as *quantum reference frames*. In this section, we briefly review how the asymmetry properties of a quantum system are established through correlations with such reference frames, rather than in any absolute sense.

Under a global symmetry described by a group G, one can nevertheless effectively perform *any* channel on a subsystem to arbitrary precision by executing that channel *relative* to a quantum reference frame encoding the group elements [7, 127, 125, 128, 129, 130]. Concretely, we say that a quantum reference frame η_R of system R encodes G as the set of states $\{\eta_R(g) | \mathcal{U}_g^R(\eta_R)\}$, known as the G-*orbit* of η_R. This reference frame is said to be *complete*[4] when it becomes a *pure*

[4]Formally, the G-orbit of a complete reference frame $|\eta_R\rangle$, $\{|\eta_R(g)\rangle \coloneqq U_R(g) |\eta_R\rangle \, | g \in G\}$, constitutes a left-

state $|\eta_R\rangle$ whose G-orbit encodes different elements of G with *perfect distinguishability* so that

$$\forall g, g' \in G : \langle \eta_R(g) | \eta_R(g') \rangle = \delta(g^{-1}g'). \tag{5.5}$$

One can produce a globally-symmetric version of *any* bounded operator $O_S \in \mathcal{B}(S)$ of system S by taking its *bipartite G-twirl* with the reference frame η_R on system R, which is defined as

$$\mathcal{G}(\eta_R \otimes O_S) \equiv \mathcal{G}_\eta(O_S) \coloneqq \int_G dg\, \mathcal{U}_g^R(\eta_R) \otimes \mathcal{U}_g^S(O_S). \tag{5.6}$$

Given a state τ_S on the system S, the bipartite G-twirl induces *classical correlations* between R and S whose strength depend on the extent to which η_R and τ_S break symmetry. To take the simplest non-trivial example, let us consider the G-twirl over two qubits with respect to the cyclic group $\mathbb{Z}_2$, which is represented on the Hilbert space of either qubit by unitaries $\{\mathbb{1}, Z\}$ generated by the Pauli Z-operator. The bipartite G-twirl on a reference qubit R in state η_R and system qubit S in state τ_S is therefore

$$\mathcal{G}(\eta_R \otimes \tau_S) = \frac{1}{2}\eta_R \otimes \tau_S + \frac{1}{2} Z(\eta_R) Z \otimes Z(\tau_S) Z. \tag{5.7}$$

At one extreme, we have that η_R or τ_S is symmetric under this representation of $\mathbb{Z}_2$ – for instance, if we have $\tau_S = |0\rangle\langle 0|$. In this case, we obtain

$$\mathcal{G}(\eta_R \otimes |0\rangle\langle 0|) = \left[\frac{1}{2}\eta_R + \frac{1}{2} Z(\eta_R) Z\right] \otimes |0\rangle\langle 0| = \mathcal{G}(\eta_R) \otimes |0\rangle\langle 0|\,, \tag{5.8}$$

from which we see that the bipartite G-twirl generates *no* correlations between R and S if one of them carries a symmetric state (and destroys any asymmetry carried by the other). At the other extreme, we consider the case where $\eta_R = \tau_S = |+\rangle\langle +|$, which is maximally symmetry-breaking because its G-orbit, $\{\mathbb{1}\,|+\rangle\langle +|\,\mathbb{1} = |+\rangle\langle +|\,, Z\,|+\rangle\langle +|\,Z = |-\rangle\langle -|\}$, encodes this representation of $\mathbb{Z}_2$ with perfect distinguishability. Consequently, we find that the bipartite G-twirl when both R and S are in the state $|+\rangle\langle +|$ yields

$$\mathcal{G}(|+\rangle\langle +|_R \otimes |+\rangle\langle +|_S) = \frac{1}{2}\,|+\rangle\langle +| \otimes |+\rangle\langle +| + \frac{1}{2}\,|-\rangle\langle -| \otimes |-\rangle\langle -|\,, \tag{5.9}$$

which is a maximally classically correlated state between R and S.

regular representation of G.

As η_R tends towards a complete reference frame, one finds, given any two bounded operators $T_S, O_S \in \mathcal{B}(S)$ of quantum system S, that

$$\begin{aligned}\mathcal{G}_\eta(T_S)\mathcal{G}_\eta(O_S) &= \left[\int_G dg\, \mathcal{U}_g^R(\eta_R)\otimes\mathcal{U}_g^S(T_S)\right]\left[\int_G dg'\, \mathcal{U}_{g'}^R(\eta_R)\otimes\mathcal{U}_{g'}^S(O_S)\right] \\ &= \int_G dg \int_G dg'\ \eta_R(g)\eta_R(g')\otimes\mathcal{U}_g^S(T_S)\mathcal{U}_{g'}^S(O_S) \\ &\xrightarrow{\text{complete ref.}} \int_G dg\ |\eta_R(g)\rangle\langle\eta_R(g)|\otimes\left[U_S(g)(T_S)U_S^\dagger(g)U_S(g)(O_S)U_S^\dagger(g)\right] \\ &= \int_G dg\, \mathcal{U}_g^R(|\eta_R\rangle\langle\eta_R|)\otimes\mathcal{U}_g^S(T_SO_S) = \mathcal{G}_\eta(T_SO_S). \end{aligned} \tag{5.10}$$

Therefore, given any measurement described by POVM elements $\{E_i\}$, any channel described by Kraus operators $\{K_j\}$ and any state τ on subsystem S, one can always construct a globally symmetric measurement described by POVM elements $\{\mathcal{G}_\eta(E_i)\}$, a globally symmetric channel described by Kraus operators $\{\mathcal{G}_\eta(K_j)\}$ and a globally symmetric state $\mathcal{G}_\eta(\tau)$ on the joint system RS with which, in the limit of a complete reference frame, the quantum statistics of the Born rule on subsystem S can be perfectly reproduced, i.e.

$$\begin{aligned}\sum_j \mathrm{Tr}\left[\mathcal{G}_\eta(E_i)\mathcal{G}_\eta(K_j)\mathcal{G}_\eta(\tau)\mathcal{G}_\eta\left(K_j^\dagger\right)\right] &\xrightarrow{\text{complete ref.}} \sum_j \mathrm{Tr}\left[\mathcal{G}_\eta\left(E_iK_j(\tau)K_j^\dagger\right)\right] \\ &= \sum_j \mathrm{Tr}\left[E_iK_j(\tau)K_j^\dagger\right]. \end{aligned} \tag{5.11}$$

In other words, the bipartite G-twirl with a reference frame provides a *globally symmetric* setting wherein *any* subsystem channel can be executed *relative* to this reference frame to *arbitrary* precision as the state of that reference frame becomes more complete.

The preparation of a state τ_S on system S relative to a reference frame η_R tends towards a classical-quantum state where R behaves as a classical "register" for symmetry group elements as η_R becomes more complete. Formally, one sees that

$$\mathcal{G}(\eta_R\otimes\tau_S) \xrightarrow{\text{complete ref.}} \int_G dg\ |\eta_R(g)\rangle\langle\eta_R(g)|\otimes\mathcal{U}_g^S(\tau_S). \tag{5.12}$$

For this reason, complete reference frames are interpreted as *classical*. Eq. (5.11) suggests that the asymmetry properties of a quantum system only become completely meaningful in the limit of classical reference frames, which can be seen as an instance of the Correspondence Principle.

5.2.3 Complete entropic monotones for asymmetry theory

Recently, Ref. [36] identified the first complete set of asymmetry monotones in terms of the *single-shot conditional min-entropy* $H_{\min}(R|S)$ [126].

Definition 41. *The single-shot condition min-entropy $H_{\min}(R|S)_\Omega$ of a state Ω_{RS} on a bipartite system RS is defined via an infimum over positive-semidefinite operators X_S on $\mathcal{H}_S$ as*

$$H_{\min}(R|S)_\Omega := -\log \inf_{X_S \geq 0} \{\mathrm{Tr}[X_S] : I_R \otimes X_S \geq \Omega_{RS}\}. \tag{5.13}$$

This complete set of asymmetry monotones is given by the single-shot conditional min-entropy evaluated on the input/output state prepared relative to every possible quantum reference frame on an external system R that *spontaneously* emerges.

Theorem 42 ([36]). *Let A and B be input and output quantum systems whose Hilbert spaces carry unitary representations of a compact symmetry group G. Furthermore, let*

$$H_\eta(\rho) \coloneqq H_{\min}(R|S)_{\mathcal{G}(\eta_R \otimes \tau_S)} \tag{5.14}$$

denote the single-shot conditional min-entropy on the bipartite G-twirl of a state τ_S on system S and a state η_R on a reference system R for G, whose Hilbert space carries a unitary representation of G that is complex conjugate to that carried by the Hilbert space of the output system B, i.e. $U_R(g) = U_B(g)^*$ for all $g \in G$ (which also implies $\dim(\mathcal{H}_R) = \dim(\mathcal{H}_B)$). We then have $\rho_A \xrightarrow{G} \sigma_B$ if and only if

$$\Delta H_\eta \coloneqq H_\eta(\sigma) - H_\eta(\rho) \geq 0 \tag{5.15}$$

for all states η_R of the reference system R.

Theorem 42 explicitly shows asymmetry to be a *relational* quantity that is not intrinsic to a quantum state, but instead lies in correlations with quantum reference frames, which explains why asymmetry must be unspeakable and cannot be fully captured by Hermitian observables on that state alone. However, in contrast to what one might expect, the asymmetry properties of a quantum system can be completely defined relative to incomplete reference frames, as the reference system in Theorem 42 does not have to be infinitely big, so its reference frames will

typically not encode elements of a continuous symmetry group with perfect distinguishability.

In the classical limit at which the quantum reference frame for G on system R becomes complete, the asymmetry monotone H_η tends to the *optimal guessing probability* $p_{\text{opt}}(g)$ [131] for the group element g held in the classical register R given system S, i.e.

$$H_\eta(\tau) \xrightarrow{\text{complete ref.}} -\log p_{\text{opt}}(g). \tag{5.16}$$

In particular, when G is the group of time-translation symmetry, $p_{\text{opt}}(t)$ is the optimal probability for guessing the time t on a Page-Wootters clock [132] given system S. Indeed, the formalism of quantum reference frames is a generalisation for the Page-Wootters construction of time as a relational quantity emerging from correlations with a physical clock, which was created to understand the appearance of temporal evolution in a universe that must remain globally time-symmetric under the Wheeler-deWitt equation [132] central to quantum gravity.

The monotonicity of H_η under G-covariant transformations on system S can then be loosely interpreted as a degradation in the ability to deduce the state of a classical register R for the symmetry group G from measurements on a covariantly-correlated system S, because the strength of those correlations must decrease as the state of S becomes more symmetric under G-covariant transformations.

5.3 Warm-up example: infinitesimally symmetric reference states

Though Theorem 42 showed that the asymmetry properties of a quantum system can be fully captured by incomplete quantum reference frames, one nevertheless expects intuitively that the more complete a reference frame η_R is (in the sense of encoding any two group elements by states at a larger distance), the more powerful the asymmetry monotone H_η generated by η_R will be (in the sense of ruling out a larger subset of G-covariantly inaccessible states). In this section, however, we present a startling counter-example where the entire structure of G-covariant state transitions can be obtained from just two reference frames that are infinitesimally close to completely symmetric, which further demonstrates that the complete monotones of Theorem 42 can be highly redundant.

We consider time-covariant transformations from a qubit S with Hamiltonian $H_S = Z$ to itself, which forms the symmetry group $\mathsf{U}(1)$ represented on $\mathcal{H}_S$ as $\{\mathcal{U}_t := e^{-iZt} | t \in [0, 2\pi)\}$. In this

case, the reference system R is a qubit with Hamiltonian $H_R = -Z$. The set of symmetric states are statistical mixtures of $|0\rangle$ and $|1\rangle$, i.e. the Z-axis of the Bloch sphere. Given a fixed input state ρ, we can define

$$\mathcal{T}_\eta(\rho) := \{\sigma : H_\eta(\sigma) \geq H_\eta(\rho)\}. \tag{5.17}$$

as the set of states classed by a *single* reference frame η_R as time-covariantly accessible. We emphasise that $\sigma \in \mathcal{T}_\eta(\rho)$ constitutes a necessary, but not sufficient, condition on $\rho \xrightarrow{\mathsf{U}(1)} \sigma$. Concretely, letting $\mathcal{T}(\rho)$ denote the full set of G-covariantly accessble states from ρ, we have

$$\mathcal{T}(\rho) = \bigcap_{\eta_R \in \mathcal{D}(R)} \mathcal{T}_\eta(\rho). \tag{5.18}$$

We can now formalise the question of which asymmetry monotones H_η are the most powerful by asking which $\mathcal{T}_\eta(\rho)$ best approximate $\mathcal{T}(\rho)$, for which a closed-form expression is known [133]. From the formalism of quantum reference frames, the natural answer would appear to be $|+\rangle$ (or any state on its orbit), since given any two times t_1 and t_2, states encoding these times on the orbit of $|+\rangle$ are at least as distinguishable as states encoding these times on the orbit of any other reference frame; more formally, we have

$$\forall t_1, t_2, \eta'_R : \mathrm{Tr}[\mathcal{U}_{t_1}(|+\rangle\langle+|), \mathcal{U}_{t_2}(|+\rangle\langle+|)] \geq \mathrm{Tr}\left[\mathcal{U}_{t_1}(\eta'_R)\mathcal{U}_{t_2}(\eta'_R)\right]. \tag{5.19}$$

In another way, $|+\rangle$ is the "best" clock state that one can find for the qubit because, as a uniform superposition over energy eigenstates, it can encode a single bit of data about the parameter t, which is the maximum allowed by the Holevo bound [37].

In Fig. 5.1, we compare $\mathcal{T}_\eta(\rho)$ (blue region) generated by a pure reference frame η_R (blue dot) at different choices of polar angle θ (and zero azimuthal angle[5]) against $\mathcal{T}(\rho)$ (region bound by black lines) for a fixed input state $\rho = \frac{1}{2}\left(\mathbb{1} + \frac{X+Z}{2}\right)$ (black dot). We see that the constraint due to $|+\rangle$ ($\theta = \frac{\pi}{2}$ on the Bloch sphere) actually yields a poor approximation for the true set of time-covariantly accessible states. On the other hand, choosing a symmetric reference frame ($\theta = \pi$) is even worse, as it does not constrain accessible states at all. Unexpectedly, the more we let reference frames approach the set of symmetric states *without* actually becoming symmetric ($\theta = 0.4\pi \to 0.01\pi$ and $\theta = 0.6\pi \to 0.99\pi$), the better the approximations they generate to the

[5]Changing the azimuthal angle of a reference frame η_R for time-covariant transitions in this non-degenerate qubit leaves $\mathcal{T}_\eta(\rho)$ constant.

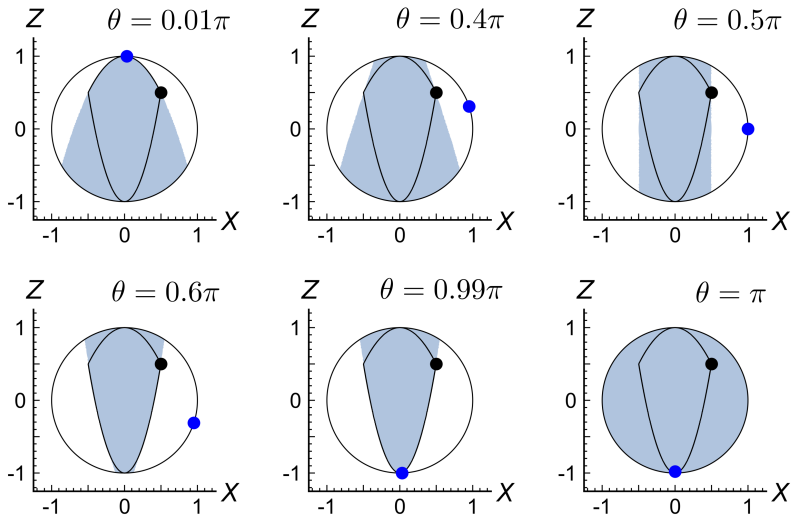


Figure 5.1: **"Less is more": near–symmetric reference frames are optimal**. We consider a qubit system under time-translation symmetry. Given a reference frame η_R (blue dot), the shaded region $\mathcal{T}_\eta(\rho)$ corresponds to potential states in the Bloch sphere which η_R classes as accessible from the state ρ (black dot) under a time-covariant transformation. The black curve marks the boundary of all states that are covariantly accessible from ρ, and is obtained from the intersection of all regions $\mathcal{T}_\eta(\rho)$. Surprisingly, the high coherence state $\eta_R = |+\rangle\langle+|$ gives a weak bound, while in contrast taking η_R very close to $|0\rangle\langle 0|$ or $|1\rangle\langle 1|$ provides complete constraints. Note that changing the azimuthal angle of η_R leaves $\mathcal{T}_\eta(\rho)$ constant.

set of time-covariantly accessible states become, until we find that

$$\mathcal{T}_{\eta(0.01\pi)}(\rho) \cap \mathcal{T}_{\eta(0.99\pi)}(\rho) \approx \mathcal{T}(\rho). \tag{5.20}$$

In other words, the sets of states classed as time-covariantly accessible from ρ by just *two* reference frames that are arbitrarily close to, without becoming, the time-symmetric states $|0\rangle$ and $|1\rangle$ approximate the full set of time-covariantly accessible states from ρ arbitrarily well.

These results are quite surprising. The formalism of quantum reference frames intuitively leads us to assume that, as the "best clock", the reference frame $|+\rangle$ would provide the most information on the time-asymmetry properties of an input state ρ, as expressed by which states ρ would be resourceful enough to transform into under time-covariant transformations. Instead, the constraints improve as reference frames approach the set of time-symmetric states, even though symmetric reference frames themselves provide no constraints at all! The findings of Fig. 5.1 can be generalised to all input states as follows (see proof in Appendix C.4.1).

Lemma 43. *In spherical polar co-ordinates on the Bloch sphere, there exists a time-covariant transformation ρ to σ in a qubit with Hamiltonian $\propto Z$ if and only if*

$$\partial_\theta^2(\Delta H_\eta)|_{\theta=0} \geq 0 \text{ and } \partial_\theta^2(\Delta H_\eta)|_{\theta=\pi} \geq 0. \tag{5.21}$$

Alternatively, parameterising a state ω of the non-degenerate qubit in the energy eigenbasis as

$$\omega = \begin{pmatrix} p_\omega & c_\omega \\ c_\omega^* & 1-p_\omega \end{pmatrix}, \tag{5.22}$$

where p_ω is a probability and c_ω is a complex number such that $|c_\omega|^2 \leq p_\omega(1-p_\omega)$, we find that there exists a time-covariant transition from ρ to σ if just *two* entropic conditions at reference frames with Bloch vectors (x_1, y_1, z_1) and (x_2, y_2, z_2) satisfying

$$0 < \frac{\sqrt{x_1^2+y_1^2}}{2z_1} \leq \min\left\{\frac{1-p_\rho}{|c_\rho|}, \frac{1-p_\sigma}{|c_\sigma|}\right\} \text{ and } 0 > \frac{\sqrt{x_2^2+y_2^2}}{2z_2} \geq \max\left\{-\frac{p_\rho}{|c_\rho|}, -\frac{p_\sigma}{|c_\sigma|}\right\} \tag{5.23}$$

respectively (see proof in Appendix C.4.1) are met. As these reference frames tend towards time-symmetric states at $x_1 = y_1 = 0$ and $x_2 = y_2 = 0$ respectively, they certify the existence of a time-covariant transition between an increasing range of input and output states. Informally, the structure of time-covariant transitions in a non-degenerate qubit is determined by asymmetry monotones H_η at two reference frames that are "infinitesimally close" to symmetric.

This simple example suggests that the dynamical constraints enforced by the asymmetry monotones H_η have a non-trivial and counter-intuitive dependence on the reference frame η_R, and can be extremely redundant. As we have seen, Lemma 43 reduces infinitely many monotone conditions to evaluate at all reference frames to just two curvature conditions to evaluate near symmetric states. This raises the question of whether similar results carry over to more general situations, and to what extent we can reduce the set of reference frames to determine the minimal relational data needed to specify the asymmetry of arbitrary quantum systems for general symmetry group G.

5.4 Sufficient surfaces of reference frames

In this section, we establish some basic redundancies in reference frames generating the asymmetry monotones of Theorem 42 due to the group structure of the symmetry constraint, and conclude that any surface of reference frame enclosing the maximally mixed state generate a sufficient set of monotones. All results apply to transformations between arbitrary systems under a general symmetry group G.

5.4.1 Basic reference frame redundancies

In the warm-up example, we observed that asymmetry monotones generated by time-symmetric reference frames seem to impose no constraints on time-covariant dynamics in a non-degenerate qubit. This observation turns out to be true of all symmetry groups (see proof in Appendix C.1.2).

Lemma 44. *Given any input state ρ_S of system S and reference frame η_R of system R,*

$$H_{\mathcal{G}(\eta)}(\rho) = H_\eta(\mathcal{G}(\rho)) = H_{\mathcal{G}(\eta)}(\mathcal{G}(\rho)) = -\log \|\mathcal{G}(\eta)\|_\infty, \tag{5.24}$$

where $\|\cdot\|_\infty$ is the Schatten-∞ norm defined in Eq. (2.8). Therefore, if η is symmetric so we have $\eta = \mathcal{G}(\eta)$, then $\Delta H_\eta = 0$.

Furthermore, any two reference frames related by a G-covariant unitary channel generate equivalent asymmetry monotones. In the warm-up example, this meant we only had to consider reference frames at the same azimuthal angle in the Bloch sphere.

Lemma 45. *If there exists a G-covariant unitary channel $\mathcal{U}^R$ on the reference system R such that $\mathcal{U}^R(\eta_R) = \eta'_R$, then $\Delta H_\eta \geq 0$ if and only if $\Delta H_{\eta'} \geq 0$ because $\Delta H_\eta = \Delta H_{\eta'}$.*

The invariance of entropic conditions under G-covariant unitary channels on the reference system is a consequence the following lemma, which we prove in Appendix C.1.1.

Lemma 46. *Let $\mathcal{U}^R$ and $\mathcal{V}^S$ be unitary channels on the reference and input systems respectively such that $\mathcal{U}^R \otimes \mathcal{V}^S$ is G-covariant (on the joint system RS). Then given any reference frame η of system R and input state τ of system S, we have that*

$$H_{\mathcal{U}^R(\eta)}\left[\mathcal{V}^S(\tau)\right] = H_\eta(\tau). \tag{5.25}$$

For Abelian groups, one need only consider reference frames with all *modes of asymmetry* (introduced in the next chapter) in common with the input state (see Appendix D.3.3.1 for proof).

5.4.2 Necessary and sufficient surfaces of reference frame states

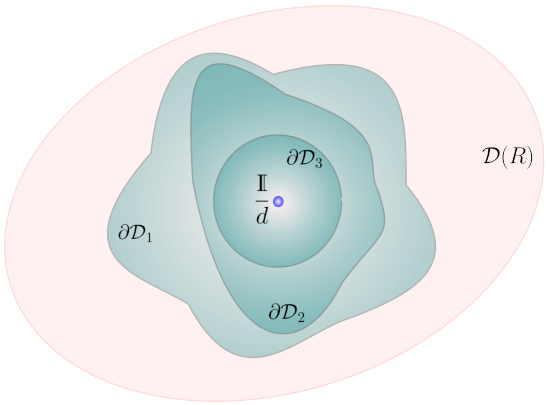


Figure 5.2: **(Sufficient surfaces of reference frames)**. There is extensive freedom in the choice of sufficient reference states. According to Theorem 47, any surface in the reference system state space $\mathcal{D}(R)$ that encloses the maximally mixed state (blue dot) is a sufficient set of reference frames – the three surfaces $\partial\mathcal{D}_1$, $\partial\mathcal{D}_2$ and $\partial\mathcal{D}_3$ provide the same information.

The entropic conditions of Theorem 42 turn out to possess a more extensive kind of redundancy, in particular that one need only consider reference frames on any surface enclosing the maximally mixed state (see proof in Appendix C.2.2). This is essentially because the entropic conditions turn out to be invariant under partial depolarisation of the reference system.

Theorem 47. *(Sufficient surfaces of states). Let ρ_A and σ_B be states of an input system A and output system B respectively, and let G be a compact symmetry group. Furthermore, let $\partial\mathcal{D}$ be any closed surface in the state space $\mathcal{D}(R)$ of the reference system R for G-covariant state transitions from A to B that encloses the maximally mixed reference state $\frac{I}{d}$, where $d = \dim(\mathcal{H}_R)$. Then $\rho \xrightarrow{G} \sigma$ if and only if*

$$\forall \eta \in \partial\mathcal{D} : \Delta H_\eta \geq 0. \tag{5.26}$$

Combined with the reference frame redundancies identified in the previous subsection, we conclude that only a subset of $\partial\mathcal{D}$ will produce non-trivial constraints – namely, the intersection of $\partial\mathcal{D}$ with asymmetric states quotiented by the sub-group of G-covariant unitary channels on the reference system alone.

5.5 Infinitesimal reference frames and a single entropic minimality condition for asymmetry theory

Theorem 47 showed any surface enclosing the maximally mixed state constitutes a sufficient set of reference frames. In particular, we are therefore free to choose a surface that bounds an arbitrarily small neighbourhood around the maximally mixed state, and so restrict to reference frames that are infinitesimally close to completely symmetric. This indicates that the counter-intuitive features we highlighted in the case of time-covariant transitions within a non-degenerate qubit are in fact generic and appear in all systems and symmetry groups. Moreover, by shrinking the neighbourhood bounded by a sufficient surface of reference frames towards the maximally mixed state, we arrive at a single necessary and sufficient condition for G-covariant state conversion.

Theorem 48. *Let ρ_A and σ_B be states of an input system A and output system B respectively, and let G be a compact symmetry group. Then $\rho_A \xrightarrow{G} \sigma$ if and only if ΔH_η attains a local minimum at the maximally mixed reference state $\eta_R = \frac{I}{d}$.*

Proof. We first note that $\Delta H_\eta = 0$ whenever η_R is symmetric due to Lemma 44. Since the maximally mixed state is symmetric for all G, we have $\Delta H_\eta = 0$ when $\eta_R = \frac{I}{d}$. If we assume $\rho_A \xrightarrow{G} \sigma_B$, then Theorem 42 implies that ΔH_η has a global minimum at $\eta_R = \frac{I}{d}$, which must therefore be a local minimum as well. Conversely, if we assume ΔH_η has a local minimum at $\eta_R = \frac{I}{d}$, then there exists a neighbourhood $\mathcal{D}$ around $\eta_R = \frac{I}{d}$ in which $\Delta H_\eta \geq \Delta H_{I/d} = 0$. Because H_η is continuous in η_R (see Lemma 52 in the next chapter), we conclude that $\Delta H_\eta \geq 0$ on $\partial\mathcal{D}$ as well, and so obtain from Theorem 47 that $\rho_A \xrightarrow{G} \sigma_B$. Therefore, $\rho_A \xrightarrow{G} \sigma_B$ if and only if $\eta_R = \frac{I}{d}$ is a local minimum of ΔH_η. □

This result is quite surprising in light of the formalism for quantum reference frames. Because arbitrary channels can only be executed perfectly relative to a classical reference frame for the symmetry group G, one might have expected that the asymmetry properties of a quantum system are best captured by correlations with reference frames that encode different group actions as distinguishably as possible. However, Theorem 48 shows that the opposite is the case – the asymmetry resources of a state can be completely determined from correlations with a reference frame that is infinitesimally close to transforming trivially under all group actions,

and so encoding different group actions completely indistinguishably.

As Theorem 48 is a local minimum condition, we might hope that the problem of G-covariant state conversion can (typically) be resolved by checking whether ΔH_η has a critical point at the maximally mixed state and conducting a second partial derivative test. This would imply that the asymmetry properties of a system might be described by a form of quantum Fisher information [134], a key quantity in asymmetry theory [125, 135] related to the curvature of divergences from which conditional min-etropies are constructed [136].

Unfortunately, ΔH_η often has a conic singularity at the maximally mixed state (see Appendix C.3 for proof). To be concrete, let us decompose every reference frame η in terms of an orthogonal basis for Hermitian operators $\{I, X_1, \ldots, X_{d^2-1}\}$ with $\|X_k\|_\infty = 1$ as

$$\eta(\boldsymbol{x}) \coloneqq \frac{1}{d}\left(I + \sum_{k=1}^{d^2-1} x_k X_k\right), \tag{5.27}$$

where d is the reference system dimension and $x_k \in \mathbb{R}$ provide coordinates for the state. We can then describe all directions out of the maximally mixed state by considering a unit sphere of states around it at co-ordinates $\mathcal{S} \coloneqq \left\{\boldsymbol{x} \in \mathbb{R}^{d^2-1} \middle| \left\|\sum_{k=1}^{d^2-1} x_k X_k\right\|_\infty = 1\right\}$. In the case of a qubit, this is simply the surface of the Bloch sphere.

Lemma 49. *In a sufficiently small neighbourhood $\epsilon > 0$ around the maximally mixed state, ΔH_η changes linearly along any direction out of the maximally mixed state. Thus one can write*

$$\forall \boldsymbol{x} \in \mathcal{S} : \Delta H_{\eta(\epsilon \boldsymbol{x})} = f(\boldsymbol{x})\epsilon + O(\epsilon^2), \tag{5.28}$$

for some function $f : \mathcal{S} \to \mathbb{R}$.

More broadly, the above lemma shows that ΔH_η will have a conical singularity at the maximally mixed state unless it becomes completely linear there. In the example of time-covariant transitions in a non-degenerate qubit, this conical singularity always appears when neither ρ nor σ is symmetric unless ρ and σ can be converted into each other via time evolution under the Hamiltonian (see the end of Appendix C.4). A conical singularity in ΔH_η can therefore appear at the maximally mixed state much more often than stated in the Lemma.

We remark that though ΔH_η is typically not differentiable at the maximally mixed state, it

may still have nice differentiability properties at points on a sufficient surface of reference frames where $\Delta H_\eta = 0$. In that case, if ΔH_η has a local minimum, as determined by the second derivative test, at every location where $\Delta H_\eta = 0$, then a G-covariant transition can occur. A successful example of this strategy reducing infinitely many asymmetry monotone conditions on a sufficient surface of reference frames to finitely many curvature conditions is time-covariant transitions in a non-degenerate qubit with Hamiltonian proportional to Z. One sufficient surface of reference frames for such transitions is the arc from $\theta = 0$ to $\theta = \pi$ at constant ϕ on the surface of the Bloch sphere. To see this, we can begin from the surface of the Bloch sphere as a larger sufficient surface of reference frames in accordance with Theorem 47, and then reduce this to the arc from $\theta = 0$ to $\theta = \pi$ by applying Lemma 45 to time evolution under the Hamiltonian. By Lemma 44, we know that $\Delta H_\eta = 0$ when the polar angle θ of the reference frame η takes on the values $\theta = 0$ or $\theta = \pi$ because η becomes the symmetric energy eigenstate $|0\rangle$ or $|1\rangle$ at those polar angles respectively. Appendix C.4.1 shows that ΔH_η is infinitely differentiable with respect to θ on this sufficient surface at $\theta = 0$ and $\theta = \pi$. We have already seen in section 5.3 that demanding $\partial_\theta^2 \Delta H_\eta|_{\theta=0,\pi} \geq 0$ is sufficient to determine whether a time-covariant transition can occur.

Chapter 6

Sufficient depolarisation for G-covariant state conversion

6.1 Introduction

In principle, the local minimality condition derived in the previous chapter (Theorem 48) for ΔH_η at the maximally mixed state provides a complete description of the asymmetry properties of quantum states. However, even in simple cases such as time-symmetric transitions in a non-degenerate qubit, solving the minimisation problem defining ΔH_η is fairly non-trivial, and we have seen that standard tests for local minima are typically not applicable for ΔH_η at the maximally mixed state. Computing this condition therefore presents a significant technical challenge that we must leave for future study.

In this chapter, we adopt a more physical perspective by allowing a certain amount of noise into state conversion, and attempt to extract finite sets of complete conditions from Theorem 42 under this weakened demand. In particular, we focus on depolarisation noise because the maximally mixed state is symmetric under arbitrary group actions, so a G-covariant state conversion can always take place once the output state has been sufficiently depolarised. More concretely, we depolarise our output state σ by some probability p to

$$\sigma_p \coloneqq (1-p)\sigma + p\frac{I}{d}, \tag{6.1}$$

where d is the dimension of output system, and pose the question

What is the minimal amount of depolarisation noise p needed to make σ accessible from the initial state ρ via a G-covariant channel?

We will take two approaches to addressing this problem, one based on coarse-graining a surface of sufficient reference frames (Theorem 47) and the entropic conditions they generate into a finite set, and the other based on decomposing quantum states into *modes of asymmetry* [137] that are never mixed under G-covariant channels.

6.1.1 Chapter summary

The rest of this chapter proceeds as follows. In section 6.2, we coarse-grain the surface of sufficient reference frames into a finite set known as an *ϵ-covering*, then upper-bound the depolarisation needed to make ΔH_η non-negative in the neighbourhood around reference frames in the covering based on its values at those reference frames. Section 6.3.2 reviews modes of asymmetry, the main technical background for this chapter. In section 6.4, we identify sufficient conditions for G-covariant state transitions defined on the modes of asymmetry present in the desired input and output states. These conditions require finding a family of G-covariant channels with members each attempting to generate as much fidelity as possible relative to a distinct reference frame. By taking this family to be measure-and-prepare channels based on the Pretty Good Measurement [138] and carefully choosing the state prepared by each channel of this family, we identify a sufficient depolarisation condition that becomes independent of the reference frame for each mode of asymmetry in the output system.

6.2 Coarse-graining conditions

We first introduce the notion of an *ϵ-ball*, which formalizes the notion of an arbitrarily small neighbourhood around a quantum state.

Definition 50 (ϵ-ball)**.** *Let η be a (possibly subnormalized) quantum state of system R. We define the ϵ-ball $\mathcal{B}_\epsilon(\eta)$ as the subset of subnormalized states that are ϵ-close to ρ as measured by the trace distance, so we can write*

$$\mathcal{B}_\epsilon(\eta) \coloneqq \{\tilde{\eta} \in \mathcal{D}_{\leq}(R) | D(\eta, \tilde{\eta}) \leq \epsilon\}. \tag{6.2}$$

With this definition in hand, we can formalize what we mean by an ϵ-covering of sufficient reference frames in the following Lemma (see Appendix D.1.2 for proof), and upper-bound the number of states in the covering.

Lemma 51. *For every $\epsilon > 0$, there exists an ϵ-covering of sufficient reference frames, i.e. a finite set of reference frames $\mathcal{N}_\epsilon$ such that every state η on a sufficient surface of reference frames $\partial\mathcal{D}$ lies within an ϵ-ball around some state in $\mathcal{N}_\epsilon$. The cardinality of $\mathcal{N}_\epsilon$ is upper-bounded as*

$$|\mathcal{N}_\epsilon| \leq \left(1 + \frac{1}{\epsilon}\right)^{d^2-1}. \tag{6.3}$$

The geometric intuition behind an ϵ-covering of sufficient reference frames is conveyed by the illustration of Lemma 51 in Fig. 6.1.

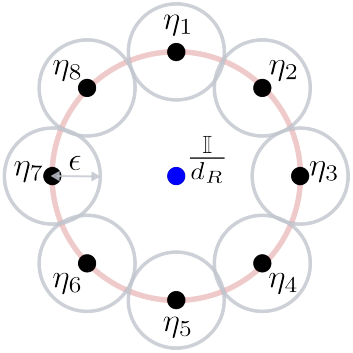


Figure 6.1: **An ϵ-covering of sufficient reference frames.** The pale pink circle shows a surface of sufficient reference frames enclosing the maximally mixed state, while the black dots form ϵ-covering $\mathcal{N}_\epsilon$ such that every state on the surface is within an ϵ-ball of a state in $\mathcal{N}_\epsilon$.

Throughout this section, we will find it easier to work with the *exponentiation* of the asymmetry monotone H_η given by $F_\eta(\tau) := 2^{-H_\eta(\tau)}$, which provides equivalent constraints on the existence of a G-covariant transition from ρ to σ because $\Delta F_\eta(\rho,\sigma) := F_\eta(\rho) - F_\eta(\sigma) \geq 0$ implies $\Delta H_\eta(\rho,\sigma) := H_\eta(\sigma) - H_\eta(\rho) \geq 0$ and vice versa (see the end Appendix C.1 for details).

We first show that the function ΔF_η is continuous in η, and upper-bound how much ΔF_η can vary within an ϵ-ball around η, in the following Lemma (see Appendix D.1.1 for proof).

Lemma 52. *Given any input state ρ of system A, output state σ of system B, reference frame η of system R with $\dim(R) = \dim(B) := d$ and state $\tilde{\eta} \in \mathcal{B}_\epsilon(\eta)$, we have that*

$$|\Delta F_{\tilde{\eta}}(\rho,\sigma) - \Delta F_\eta(\rho,\sigma)| \leq 2d\epsilon := r(\epsilon). \tag{6.4}$$

We highlight that the continuity of $\Delta F_\eta(\rho,\sigma)$ in reference frame η implies the continuity of $\Delta H_\eta(\rho,\sigma)$ in η as well (see the end of Appendix D.1.1 for details).

Lemma 52 shows that, if $\Delta F_{\eta_k}(\rho,\sigma) \geq r(\epsilon)$ on the single reference frame η_k in an ϵ-covering, then $\Delta F_\eta(\rho,\sigma) \geq 0$ for *all* reference frames in the ϵ-ball around η. While $\Delta F_{\eta_k}(\rho,\sigma) < 0$ means no G-covariant transformation is possible, if $r(\epsilon) \geq \Delta F_{\eta_k}(\rho,\sigma) \geq 0$, then we can attempt to boost $\Delta F_\eta(\rho,\sigma)$ to $\Delta F_\eta(\rho,\sigma_p) = \Delta F_\eta(\sigma,\sigma_p) + \Delta F_\eta(\rho,\sigma)$ by depolarising σ (we are guaranteed $\Delta F_\eta(\sigma,\sigma_p) \geq 0$ as partial depolarisation is G-covariant) until $\Delta(\rho,\sigma_p) \geq r(\epsilon)$. This line of thinking leads to the following result (see Appendix D.1.3 for proof).

Theorem 53. *Given any $\epsilon > 0$, there is a finite set of reference frames $\mathcal{N} := \{\eta_k\}_{k=1}^N$ with $|\mathcal{N}| \leq (1+\frac{1}{\varepsilon})^{d^2-1}$, where d is the dimension of the reference system, such that:*

- *if $\Delta F_{\eta_k}(\rho,\sigma) < 0$ for any $\eta_k \in \mathcal{N}$ then $\rho \to \sigma$ is forbidden under G-covariant channels.*
- *if $\Delta F_{\eta_k}(\rho,\sigma) \geq r(\epsilon)$ for all $\eta_k \in \mathcal{N}$ then $\rho \to \sigma$ under a G-covariant quantum channel.*
- *For each $\eta_k \in \mathcal{N}$ where $0 \leq \Delta F_{\eta_k}(\rho,\sigma) < r(\epsilon)$ we obtain the lower bound*

$$p \geq \frac{r(\epsilon) - \Delta F_{\eta_k}(\rho,\sigma)}{\Delta F_{\eta_k}\left(\sigma, \frac{I}{d}\right)} \tag{6.5}$$

on the amount of depolarisation required to ensure $\rho \to \sigma_p$ under a G-covariant channel.

Because $r(\epsilon)$ scales with output system dimension, generating sufficient depolarisation conditions from a reasonable number of reference frames may only be practical in low-dimensional systems. Therefore, instead of developing this line further here, we take a different approach by exploiting the *modes of asymmetry* within a quantum system to identify sufficient depolarisation conditions that are independent of reference frame.

6.3 Technical background: modes of asymmetry

We begin with a brief technical overview on modes of asymmetry. Brilliant and much more detailed exposition of this material can be found at e.g. Ref [125, 137].

6.3.1 The irreducible tensor operator (ITO) basis

The modes of asymmetry in a quantum system S are constructed in terms of an *irreducible tensor operator (ITO) basis* for the representation of a finite or compact Lie group G on the set of

bounded operators $\mathcal{B}(S)$. In this section, we introduce the definition of an ITO basis alongside some of its basic properties.

Given any unitary operator U on $\mathcal{H}_S$, its associated super-operator $\mathcal{U}(\cdot) \coloneqq U(\cdot)U^\dagger$ preserves this inner product, so $\mathcal{U}$ can be thought of as a unitary acting on $\mathcal{B}(S)$. Therefore, a unitary representation $g \to U(g)$ of a group G on $\mathcal{H}_S$ induces a unitary representation $g \to \mathcal{U}_g \coloneqq U(g)(\cdot)U^\dagger(g)$ of G on $\mathcal{B}(S)$ because we have

$$\forall g_1, g_2 \in G : \mathcal{U}_{g_1}\mathcal{U}_{g_2} = \mathcal{U}_{g_1 g_2}. \tag{6.6}$$

An ITO basis is any orthonormal basis for $\mathcal{B}(S)$ that places $\mathcal{U}_g$ into a block-diagonal form, wherein every block is an irrep of G, and every copy of an irrep is represented by identical matrix components.

Definition 54 (Irreducible Tensor Operator (ITO) basis)**.** *Let $g \to \mathcal{U}_g$ be a unitary representation of a compact group G on the bounded operators $\mathcal{B}(S)$ of a quantum system S. Then $\mathcal{U}_g$ can be decomposed as follows into distinct finite-dimensional irreducible representations (irreps) labelled by λ,*

$$\mathcal{U}_g = \bigoplus_\lambda \left(\mathcal{U}_g^{(\lambda)}\right)^{\oplus \alpha_\lambda} \tag{6.7}$$

where $\mathcal{U}_g^{(\lambda)}$ is the λ-irrep of G and occurs α_λ times in the irrep decomposition of $\mathcal{U}_g$. We say that a basis $\{X_j^{(\lambda,\alpha)}\}$ for $\mathcal{B}(S)$ is an irreducible tensor operator (ITO) basis for the representation $\mathcal{U}_g$ if its elements are orthonormal and satisfy

$$\mathcal{U}_g\left(X_j^{(\lambda,\alpha)}\right) = \sum_{i=1}^{d_\lambda} U_{ij}^{(\lambda)}(g) X_i^{(\lambda,\alpha)}, \tag{6.8}$$

where $U_{ij}^{(\lambda)}(g)$ are complex matrix components for $\mathcal{U}_g^{(\lambda)}$, d_λ is the dimension of the λ-irrep, and the index $\alpha \in \{1, \ldots, \alpha_\lambda\}$ accounts for the multiplicity of the λ-irrep.

We highlight a couple of important requirements that are implicit in the ITO basis Given any two representations $\mathcal{U}_g^A$ on $\mathcal{B}(A)$ and $\mathcal{U}_g^B$ on $\mathcal{B}(B)$ of the same group G, where A and B could be distinct quantum systems, the ITO basis elements for any irrep λ common to both representations must be chosen such that λ is represented by the *same* matrix elements $U_{ij}^{(\lambda)}(g)$ in both

representations. Furthermore, two irreps are considered equivalent under the ITO formalism if and only if they are related by a unitary similarity transformation, as ITO basis elements must be orthonormal.

Before presenting some examples of ITO bases, we note that the irreps in $\mathcal{U}_g$ occur in pairs related by complex conjugation. Given an ITO basis $\{X_j^{(\lambda,\alpha)}\}$ for the representation $\mathcal{U}_g$ of a compact G on the bounded operators $\mathcal{B}(S)$ for a quantum system S, one can identify another orthonormal basis for $\mathcal{B}(S)$ by taking the Hermitian conjugate of the basis elements to obtain $\{X_j^{(\lambda,\alpha)\dagger}\}$. Because elements in the ITO basis that span the β-copy of the μ-irrep, $\{X_j^{(\lambda=\mu,\alpha=\beta)}\}_{j=1,\ldots,d_\mu}$, transform solely amongst each other under the group action according to $U^{(\mu)}(g)$, one can show that their Hermitian conjugates, $\{X_j^{(\lambda=\mu,\alpha=\beta)\dagger}\}_{j=1,\ldots,d_\mu}$, transform solely amongst each other under the group action according to *complex conjugate* of $U^{(\mu)}(g)$ as

$$\mathcal{U}_g\left(X_j^{(\mu,\beta)^\dagger}\right) = \left(\mathcal{U}_g\left(X_j^{(\mu,\beta)}\right)\right)^\dagger = \sum_{i=1}^{d_\mu}\left(U_{ij}^{(\mu)}(g)\right)^* X_i^{(\mu,\beta)^\dagger}. \tag{6.9}$$

It can be shown (see Appendix D.3.1) that $U^{(\mu)}(g)^*$ is also an irrep of G, and that the irreps $U^{(\mu_1)}(g)^*$ and $U^{(\mu_2)}(g)^*$ are equivalent if and only if $U^{(\mu_1)}(g)$ and $U^{(\mu_2)}(g)$ are equivalent. This implies $\{X_j^{(\lambda,\alpha)\dagger}\}$ constitutes another ITO basis for $\mathcal{B}(S)$, wherein $\{X_j^{(\mu,\beta)\dagger}\}_{j=1,\ldots,d_\lambda}$ constitute an orthonormal basis for a single copy of the unique irrep appearing in $g \to \mathcal{U}_g$ that is equivalent to $U^{(\mu)}(g)^*$, which we denote by μ^*.

6.3.1.1 Example: ITO basis for $\mathsf{U}(1)$ represented as time evolution by ladder Hamiltonian

Consider a quantum system S of dimension d whose Hamiltonian is a "ladder" of equally-spaced energy levels, $H_S := \sum_{n=0}^{d-1} n\Delta\,|E_n\rangle\langle E_n|$. The time evolution of system S, given by $\{\mathcal{U}_t := e^{-iHt}(\cdot)e^{iHt} | t \in [0, \frac{2\pi}{\mathrm{lcm}(1,\ldots,d-1)})\}$ wherein $\mathrm{lcm}(\cdot)$ is the lowest common multiple function, is a unitary representation of the group $\mathsf{U}(1)$ on $\mathcal{B}(S)$.

Because $\mathsf{U}(1)$ is Abelian, its (complex) irreps must be one-dimensional [139]. If we evolve any operator of the form $|E_j\rangle\langle E_k|$ with time, we find that

$$e^{-iHt}\,|E_j\rangle\langle E_k|\,e^{iHt} = e^{-i(j-k)\Delta t}\,|E_j\rangle\langle E_k|\,, \tag{6.10}$$

which implies that $|E_j\rangle\langle E_k|$ spans the subspace for a one-dimensional irrep we can label by

$\lambda \coloneqq (j-k)\Delta$. Since the set of such operators, $\{|E_j\rangle\langle E_k| \,|j,k=0,\ldots,d-1\}$, is a orthonormal basis for $\mathcal{B}(S)$, it constitutes an ITO basis for the representation $t \to \mathcal{U}_t$ of $\mathsf{U}(1)$ on $\mathcal{B}(S)$.

The distinct irreps appearing in this representation are then determined by the range of values $(j-k)$ can take, i.e. we have $\lambda \in \{-(d-1)\Delta,\ldots,-\Delta,0,\Delta,\ldots,(d-1)\Delta\}$. The complex conjugate of irrep λ is $-\lambda$, because while $|E_j\rangle\langle E_k|$ transforms under $\mathcal{U}_t$ by picking up a factor of $e^{-i\lambda t}$, its Hermitian conjugate $|E_k\rangle\langle E_j|$ transforms under $\mathcal{U}_t$ by picking up a factor of $e^{i\lambda t}$.

6.3.1.2 Example: ITO basis for $\mathsf{SU}(2)$ represented by arbitrary unitary evolution of a qubit

The possible unitary evolution a single qubit, given by $\{\mathcal{U}_{\hat{\boldsymbol{n}},\theta}|\theta \in [0,\pi), \hat{\boldsymbol{n}} \in \mathbb{R}^3 \text{ s. t. } |\hat{\boldsymbol{n}}| = 1\}$ where $\mathcal{U}_{\hat{\boldsymbol{n}},\theta} \coloneqq e^{-i\theta\hat{n}\cdot\boldsymbol{S}}(\cdot)e^{i\theta\hat{n}\cdot\boldsymbol{S}}$ for $\boldsymbol{S}$ being the vector of Pauli operators $\boldsymbol{S} \coloneqq (X,Y,Z)^T$, is a unitary representation of the group $\mathsf{SU}(2)$ on $\mathcal{B}(1)$. In the Bloch sphere, $\mathcal{U}_{\hat{\boldsymbol{n}},\theta}$ rotates the Bloch vector $\boldsymbol{r}$ for the quantum state $\rho = \frac{\mathbb{1}}{2} + \frac{1}{2}\boldsymbol{r}\cdot\boldsymbol{S}$ about the $\hat{\boldsymbol{n}}$ direction through angle θ.

Consider the orthonormal basis for $\mathcal{B}(1)$ formed by the Pauli operators $\frac{1}{\sqrt{2}}\{\mathbb{1},X,Y,Z\}$. Because $\mathcal{U}(\mathbb{1}) = \mathbb{1}$ for any unitary super-operator on $\mathcal{B}(1)$, we see that $\frac{\mathbb{1}}{\sqrt{2}}$ is the basis for a *trivial irrep* of $\mathsf{SU}(2)$. From the Bloch sphere description, we can see that $\mathcal{U}_{\hat{\boldsymbol{n}},\theta}$ is represented in the basis $\frac{1}{\sqrt{2}}\{X,Y,Z\}$ exactly as the rotation $R(\hat{\boldsymbol{n}},\theta)$ about the $\hat{\boldsymbol{n}}$ direction for angle θ would be represented in the basis provided by three orthonormal vectors $\hat{\boldsymbol{x}}$, $\hat{\boldsymbol{y}}$ and $\hat{\boldsymbol{z}}$ in 3D-space. Therefore, just as the representation for the group of rotations $\mathsf{SO}(3)$ in 3D space is irreducible, $\frac{1}{\sqrt{2}}\{X,Y,Z\}$ must also provide the basis for a three-dimensional irrep of $\mathsf{SU}(2)$.

Therefore, $\frac{1}{\sqrt{2}}\{\mathbb{1},X,Y,Z\}$ is an ITO basis for $\mathcal{B}(1)$ under the representation $\mathcal{U}_{\theta,\hat{\boldsymbol{n}}}$ for $\mathsf{SU}(2)$. Both irreps in this example are complex conjugate to themselves.

6.3.2 Constructing modes of asymmetry

The ITO basis is used to decompose any bounded operator into *modes of asymmetry*.

Definition 55 (Modes of asymmetry). *Let $\left\{X_j^{(\lambda,\alpha)}\right\}$ be an ITO basis for the representation $g \to \mathcal{U}_g$ of a compact group G on the bounded operators $\mathcal{B}(S)$ of a quantum system S. Then we can decompose any operator $O \in \mathcal{B}(S)$ into modes of asymmetry $\{O_j^\lambda\}$ as*

$$O = \sum_\lambda \sum_{j=1}^{\dim(\lambda)} \left(\sum_\alpha \left\langle X_j^{(\lambda,\alpha)}, O \right\rangle X_j^{(\lambda,\alpha)} \right) \coloneqq \sum_\lambda \sum_{j=1}^{\dim(\lambda)} O_j^{(\lambda)}. \tag{6.11}$$

We will denote the set of modes of asymmetry present in a quantum state ρ by $\mathrm{modes}(\rho)$, and similarly the set of modes of asymmetry that could be carried by any bounded operator in $\mathcal{B}(S)$ for some quantum system S by $\mathrm{modes}(S)$.

A simple way of identifying the (μ, k) mode of asymmetry for a bounded operator $O \in \mathcal{B}(S)$ for some quantum system S in the ITO basis $\left\{X_j^{(\lambda,\alpha)}\right\}$ is provided by the projector

$$\mathcal{P}_k^{(\mu)} \coloneqq \int_G dg \, \dim(\mu) u_{kk}^{(\mu^*)}(g) \, \mathcal{U}_g \tag{6.12}$$

on $\mathcal{B}(S)$, where $u_{kk}^{(\mu^*)}$ is the (k,k) matrix component (in the given ITO basis) for the μ^* irrep occurring in the representation $g \to \mathcal{U}_g$ on $\mathcal{B}(S)$. Concretely, we have that (see Appendix D.2.1 for proof).

$$\mathcal{P}_k^{(\mu)}(O) = O_k^{(\mu)}. \tag{6.13}$$

We conclude by remarking that distinct modes of asymmetry are orthogonal, i.e. given any two bounded operators A and B in $\mathcal{B}(S)$, we have that

$$\langle A_j^{(\lambda)}, B_k^{(\mu)} \rangle = \delta_{\lambda,\mu}\delta_{j,k}\langle A_j^{(\lambda)}, B_k^{(\mu)} \rangle, \tag{6.14}$$

(see Appendix D.2.2 for proof), which further implies

$$\langle A, B \rangle = \sum_{\lambda,j} \langle A_j^{(\lambda)}, B_j^{(\lambda)} \rangle. \tag{6.15}$$

6.3.2.1 Example: modes of asymmetry for $\mathsf{U}(1)$ represented by time evolution of a qubit

Consider a qubit with two energy eigenstates $|E_0\rangle, |E_1\rangle$ at distinct energies $0, \Delta$ respectively. In Section 6.3.1.1, we identified $\{|E_0\rangle\langle E_0|, |E_1\rangle\langle E_1|, |E_0\rangle\langle E_1|, |E_1\rangle\langle E_0|\}$ as an ITO basis for $\mathsf{U}(1)$ as represented by the time evolution of this qubit, from which we build the modes of asymmetry

$$\begin{aligned}
\lambda = 0 : \rho^{(0)} &= \rho_{00} |E_0\rangle\langle E_0| + |E_1\rangle\langle E_1| && \implies \mathcal{U}_t(\rho^{(0)}) = \rho^{(0)}. \\
\lambda = \Delta : \rho^{(1)} &= \rho_{10} |E_1\rangle\langle E_0| && \implies \mathcal{U}_t(\rho^{(1)}) = e^{-i\Delta t}\rho^{(\Delta)} \\
\lambda = -\Delta : \rho^{(-1)} &= \rho_{01} |E_0\rangle\langle E_1| && \implies \mathcal{U}_t(\rho^{(-\Delta)}) = e^{i\Delta t}\rho^{(-\Delta)}.
\end{aligned}$$

In the energy eigenbasis, the modes of asymmetry for a qubit state ρ can simply be read off its density matrix for state ρ as

$$\rho = \begin{pmatrix} \rho_{00} & \rho_{01} \\ \rho_{10} & \rho_{11} \end{pmatrix}, \tag{6.16}$$

with the $\lambda = 0$ mode in blue, the $\lambda = \Delta$ mode in green, and the $\lambda = -\Delta$ mode in red.

6.3.3 Properties of the modes of asymmetry

What makes modes of asymmetry such a useful construction for analysing state transitions under a symmetry constraint is that they cannot be mixed under G-covariant channels.

Theorem 56 (Lemma 8 of Ref. [125]). *Let $\Phi_{A\to B}$ be a covariant channel with respect to a compact group G from input system A to output system B. Then given any state ρ and mode of asymmetry (λ, j) of system A, we have that*

$$\Phi_{A\to B}\left(\rho_j^{(\lambda)}\right) = \Phi_{A\to B}(\rho)_j^{(\lambda)}. \tag{6.17}$$

This theorem implies that G-covariant channels eliminate modes of asymmetry that can be present in the input state but not the output state. As a simple example, given two qubits A and B with different gaps Δ_A and Δ_B between their two energy levels, the only mode of asymmetry they share with respect to time-translation symmetry is $\lambda = 0$. Consequently, time-symmetric transformations from A to B reduce to preparations of time-symmetric (i.e. free) states on B. More generally, if $\rho_A \xrightarrow{G} \sigma_B$ for a compact group G, then $\mathrm{Modes}(\sigma_B) \subseteq \mathrm{Modes}(\rho_A)$.

Generalising from Example 6.3.1.2, we can see that, no matter which symmetry group G we choose, its representation $\mathcal{U}_g$ on the bounded operators $\mathcal{B}(S)$ of a quantum system S necessarily contains the trivial irrep because I transforms trivially under all group actions, i.e. $\mathcal{U}_g(I) = I$ for all g. Throughout the rest of this chapter, we assign the label $\lambda = 0$ to the mode of asymmetry in the subspace spanned by trivial irreps, which we will call the *trivial mode of asymmetry*.

Noting that the trivial irrep is its own complex conjugate, we see from Eq. (6.12) that the projector for the trivial mode of asymmetry of an operator reduces to the G-twirl. In particular,

given any operator $O_S \in \mathcal{B}(S)$ of a quantum system S, we can write

$$O_S^{(0)} = \mathcal{G}(O_S). \tag{6.18}$$

6.4 Modal conditions

We now derive simple and closed-form sufficient depolarisation conditions, one for each mode of asymmetry, on G-covariant state conversions for compact symmetry groups whose multi-dimensional irreps have no multiplicity. These include all compact Abelian groups (as their complex irreps are always one-dimensional [139]) as well as important representations of non-Abelian groups such as $\mathsf{SU}(2)$ on a spin-j system. For such representations, one can always choose (see Appendix D.3.2 for proof) ITO bases satisfying the following two conditions, which will turn out to be extremely useful throughout this section.

Lemma 57. *Let ρ be the input state for a G-covariant state transition from a quantum system S, and consider the representation $\mathcal{U}_g$ of a compact group G on $\mathcal{B}(S)$ whose multidimensional irreps have no multiplicity. Then there exists an ITO basis $\{X_j^{(\lambda,\alpha)}\}$ for $\mathcal{U}_g$ such that*

1. *corresponding basis elements of complex conjugate irreps are Hermitian conjugates of each other. Thus one can write*

$$X_j^{(\lambda^*,\alpha)} = X_j^{(\lambda,\alpha)\dagger}. \tag{6.19}$$

2. *in every irrep λ, the input state ρ has at most a single non-zero mode of asymmetry, which we label $(\lambda, 0)$. Thus one can write*

$$\rho_j^{(\lambda)} = \delta_{j,0}\rho_0^{(\lambda)}. \tag{6.20}$$

Throughout the rest of this section, we assume ITO bases satisfying the conditions of Lemma 57. As concrete examples, we highlight that ITO bases presented in sections 6.3.1.1 and 6.3.1.2 both satisfy Equation 6.19.

We begin with the following Lemma from Ref. [36], which exploits the convexity of asymmetry theory to frame the search for sufficient conditions on G-covariant state conversion in terms of a family of G-covariant channels $\{\Phi^\eta\}$ indexed by every state η of the reference system.

Lemma 58 ([36]). *Let ρ, σ, and η be quantum states on input system A, output system B, and reference system R respectively, where $\dim(\mathcal{H}_R) = \dim(\mathcal{H}_B)$. We have $\rho \xrightarrow{G} \sigma$ if we can find a family of G-covariant channels $\{\Phi^\eta | \eta \in \mathcal{D}(R)\}$ from system A to B such that*

$$\forall \eta : \langle \eta, \Phi^\eta(\rho) \rangle \geq \langle \eta, \sigma \rangle. \tag{6.21}$$

The figure below provides an illustration of how Lemma 58 generates sufficient conditions for G-covariant state conversion.

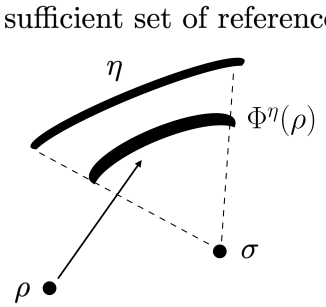


Figure 6.2: **Graphical interpretation of Lemma 58.** If, for a complete set of reference frame states $\{\eta\}$, we construct some family of covariant protocols that transform from a state ρ to $\Phi^\eta(\rho)$ that has a higher overlap with η than σ has with η, then it is possible to transform from ρ to σ under a G–covariant channel.

We saw in Section 6.3.2 that distinct modes of asymmetry are orthogonal, and in Theorem 56 that G-covariant channels do not mix distinct modes of asymmetry in a quantum state. These facts enable us to split the sufficient conditions in Lemma 58 by modes of asymmetry.

Corollary 3. *Let ρ, σ and η be quantum states of systems A, B and R respectively, where $\dim(\mathcal{H}_R) = \dim(\mathcal{H}_B)$. We have that $\rho \xrightarrow{G} \sigma$ if for any reference frame η, we can find a G-covariant channel Φ^η from system A to B such that*

$$\sum_{\lambda,j} \left\langle \eta_j^{(\lambda)}, \Phi^\eta(\rho)_j^{(\lambda)} - \sigma_j^{(\lambda)} \right\rangle = \sum_{\lambda,j} \left\langle \eta_j^{(\lambda)}, \Phi^\eta\left(\rho_j^{(\lambda)}\right) - \sigma_j^{(\lambda)} \right\rangle \geq 0. \tag{6.22}$$

6.4.1 Measure-and-prepare channels from the Pretty Good Measurement (PGM)

In this subsection, we extract slightly more practical sufficient conditions from Corollary 3 by choosing Φ^η to be a family of measure-and-prepare channels based on the Pretty Good

Measurement (PGM) [138], which are G-covariant by construction. The PGM is a measurement described by POVM elements $\{M_{\text{pgm}}(g)|g \in G\}$ defined for an input state ρ as

$$M_{\text{pgm}}(g) \coloneqq \mathcal{G}(\rho)^{-\frac{1}{2}}\mathcal{U}_g(\rho)\mathcal{G}(\rho)^{-\frac{1}{2}} = \mathcal{U}_g(\overline{\rho}), \tag{6.23}$$

where we have introduced the notation $\overline{\rho} \coloneqq \mathcal{G}(\rho)^{-\frac{1}{2}}\rho\mathcal{G}(\rho)^{-\frac{1}{2}}$. Concretely, we choose Φ^η to be

$$\Phi^\eta_{\text{pgm}}(\rho) \coloneqq \int_G dg \ \text{Tr}[M_{\text{pgm}}(g)\rho]\mathcal{U}_g(\tau(\eta)) \tag{6.24}$$

for some state $\tau(\eta)$ dependent upon η. This channel can be interpreted as performing the PGM and preparing the state $\mathcal{U}_g(\tau(\eta))$ upon obtaining the g, and is G-covariant because given any group element g', we have that

$$\begin{aligned}
\Phi^\eta_{\text{pgm}} \circ \mathcal{U}_{g'}(\rho) &= \int_G dg \ \text{Tr}\Big[M_{\text{pgm}}(g)U(g')\rho U^\dagger(g')\Big]\mathcal{U}_g(\tau(\eta)) \\
&= \int_G dg \ \text{Tr}\Big[U^\dagger(g')M_{\text{pgm}}(g)U(g')\rho\Big]\mathcal{U}_g(\tau(\eta)) \\
&= \int_G dg \ \text{Tr}\big[M_{\text{pgm}}(g'^{-1}g)\rho\big]\mathcal{U}_g(\tau(\eta)) \\
&= \int_G dg \ \text{Tr}[M_{\text{pgm}}(g)\rho]\mathcal{U}_{g'g}(\tau(\eta)) \\
&= \int_G dg \ \text{Tr}[M_{\text{pgm}}(g)\rho]\mathcal{U}_{g'} \circ \mathcal{U}_g(\tau(\eta)) \\
&= \mathcal{U}_{g'}\left[\int_G dg \ \text{Tr}[M_{\text{pgm}}(g)\rho]\mathcal{U}_g(\tau(\eta))\right] = \mathcal{U}_{g'} \circ \Phi^\eta_{\text{pgm}}(\rho).
\end{aligned} \tag{6.25}$$

By restricting Φ^η to PGM measure-and-prepare channels Φ^η_{pgm} in Corollary 3, we replace the search for a G-covariant channel at every reference frame η with the following easier search for a state $\tau(\eta)$ (see proof in Appendix D.3.3).

Lemma 59. *Let ρ, σ and η be quantum states of systems A, B and R respectively, where we have $\dim(\mathcal{H}_R) = \dim(\mathcal{H}_B)$. Given a symmetry group G whose multidimensional irreps have no multiplicity, we have $\rho \xrightarrow{G} \sigma$ if for any reference frame η, we can find a state $\tau(\eta)$ such that*

$$\sum_{\lambda,j} \left\langle \eta_j^{(\lambda)}, f_j^{(\lambda)}(\rho)\tau(\eta)_j^{(\lambda)} - \sigma_j^{(\lambda)} \right\rangle \geq 0, \tag{6.26}$$

where we have introduced $f_j^{(\lambda)}(\rho) \coloneqq \frac{1}{d_\lambda}\left\langle \rho_j^{(\lambda^*)}, \overline{\rho}_j^{(\lambda^*)} \right\rangle$.

We note in passing that the quantity $f_j^{(\lambda)}(\rho)$ can be related to the Sandwiched α–Rényi divergence $D_\alpha(\rho||\sigma)$ for two states ρ, τ of a quantum system, which is defined as [16, 91]

$$D_\alpha(\rho||\tau) := \frac{1}{\alpha - 1} \log \mathrm{Tr}\left[\tau^{\frac{1-\alpha}{2\alpha}} \rho \tau^{\frac{1-\alpha}{2\alpha}}\right]^\alpha, \tag{6.27}$$

whenever the support of ρ lies in the support of τ, and is infinite otherwise. Focussing on $\alpha = 2$ and extending the domain of the first argument to all linear operators as

$$D_2(X||\tau) := \log \mathrm{Tr}\left[\left(\tau^{-\frac{1}{4}} X \tau^{-\frac{1}{4}}\right)^\dagger \left(\tau^{-\frac{1}{4}} X \tau^{-\frac{1}{4}}\right)\right], \tag{6.28}$$

we can relate $f_j^{(\lambda)}(\rho)$ to the $\alpha = 2$ Sandwiched Rényi entropy as

$$\log f_j^{(\lambda)}(\rho) = D_2\left(\rho_j^{(\lambda)} \middle|\middle| \mathcal{G}(\rho)\right) - \log d_\lambda. \tag{6.29}$$

6.4.2 General sufficient conditions for G-covariant state conversion

As they stand, the sufficient conditions of Lemma 59 seem no easier to apply than the complete entropic conditions provided by Theorem 42, as Lemma 59 similarly has one condition for each reference frame and therefore an infinite set of conditions overall. In this subsection, we show that it is possible to choose $\tau(\eta)$ such that the dependence on η in Lemma 59 disappears, and we extract modal sufficient conditions for G-covariant state conversion that depend purely on the properties of the input and output states.

We begin with the following Lemma, for which a proof can be found in Appendix D.3.5.

Lemma 60. *Let ρ, σ and η be quantum states of systems A, B and R respectively, where we have $\dim(\mathcal{H}_R) = \dim(\mathcal{H}_B)$. Furthermore, let us define components for states η, $\tau(\eta)$ and σ in the ITO basis $\{X_j^{(\lambda,\alpha)}\}$ of system B as $n_j^{(\lambda,\alpha)} := \langle X_j^{(\lambda,\alpha)}, \eta\rangle$, $t_j^{(\lambda,\alpha)} := \langle X_j^{(\lambda,\alpha)}, \tau(\eta)\rangle$ and $s_j^{(\lambda,\alpha)} := \langle X_j^{(\lambda,\alpha)}, \sigma\rangle$ respectively. Given a symmetry group G whose multidimensional irreps have no multiplicity, we have that $\rho \xrightarrow{G} \sigma$ if for any reference frame η, we can find a state $\tau(\eta)$ of system B such that*

$$\forall(\lambda, \alpha, j) : \begin{cases} f_j^{(\lambda)}(\rho) t_j^{(\lambda,\alpha)} = s_j^{(\lambda,\alpha)} \text{ or} \\ f_j^{(\lambda)}(\rho)\left|t_j^{(\lambda,\alpha)}\right| \geq \left|s_j^{(\lambda,\alpha)}\right| \text{ and } \arg(t_j^{(\lambda,\alpha)}) = \arg(n_j^{(\lambda,\alpha)}). \end{cases} \tag{6.30}$$

The main proof idea behind this Lemma is effectively to make the sufficient conditions in Lemma 59 stricter by insisting that they apply for individual modes of asymmetry, i.e. demanding that for every reference frame η, we find a state $\tau(\eta)$ such that

$$\forall(\lambda, j) \in \text{modes(B)} : \left\langle \eta_j^{(\lambda)}, f_j^{(\lambda)}(\rho)\tau(\eta)_j^{(\lambda)} - \sigma_j^{(\lambda)} \right\rangle \geq 0, \tag{6.31}$$

and then showing that each such condition is met if $\tau(\eta)$ has ITO basis components satisfying the conditions stated in Eq. (6.30).

The sufficient conditions in Lemma 60 are nearly η-independent – the only place where η appears is in the arguments of ITO basis components for $\tau(\eta)$ in the second case of Eq. (6.30). By choosing the magnitudes of ITO basis components on $\tau(\eta)$ to ensure that it remains a valid quantum state *regardless* of the arguments on those components, we arrive at the following Theorem, for which a proof is provided in Appendix D.3.6.

Theorem 61. *Let ρ and σ be two states of systems A and B respectively. Given a symmetry group G whose multidimensional irreps have no multiplicity, we have $\rho \xrightarrow{G} \sigma$ if*

$$\forall \lambda \neq 0 : \begin{cases} \mu_{\min} n^{-1} f_j^{(\lambda)}(\rho) \geq g_j^{(\lambda)}(\sigma) & \text{for } j = 0 \\ \sigma_j^{(\lambda)} = 0 & \text{for } j \neq 0. \end{cases} \tag{6.32}$$

where we have defined $g_j^{(\lambda)}(\sigma) \coloneqq \sum_\alpha \left| \langle X_j^{(\lambda,\alpha)}, \sigma \rangle \right|$ for the ITO basis $\{X_j^{(\lambda,\alpha)}\}$ of system B, $\mu_{\min}$ as the smallest eigenvalue of $\mathcal{G}(\sigma)$, and n as the number of distinct non-trivial irreps appearing in the representation of G on $\mathcal{B}(B)$.

Because we have made a choice of ITO basis such that $\rho_j^{(\lambda)} = \delta_{j,0}\rho_0^{(\lambda)}$, the condition in Eq. (6.32) on the non-trivial mode of asymmetry ($\lambda \neq 0, j \neq 0$) is a manifestation of the fact that $\text{Modes}(\sigma) \subseteq \text{Modes}(\rho)$ is a necessary condition for a G-covariant state transition from ρ to σ, as discussed in section 6.3.3.

6.4.3 Sufficient depolarisation conditions for G-covariant state transitions

We now substitute the output state σ in Theorem 61 by σ_p to identify the minimum depolarisation sufficient to guarantee a G-covariant state transition from the input state ρ. As the maximally mixed state is symmetric and therefore only contains the trivial mode of asymmetry

by Eq. (6.18), we can decompose σ_p into modes of asymmetry as follows

$$\sigma_p = \left(p\frac{I}{d} + (1-p)\mathcal{G}(\sigma)\right) + (1-p)\sum_{\lambda\neq 0,j}\sigma_j^{(\lambda)}. \tag{6.33}$$

Denoting the smallest eigenvalue of $\mathcal{G}(\sigma)$ by $\mu_{\min}$ as previously done, we see from Eq. (6.33) that the smallest eigenvalue of $\mathcal{G}(\sigma_p)$ is $(1-p)\mu_{\min} + \frac{p}{d}$ and that the $(\lambda \neq 0, j)$ non-trivial mode of asymmetry for σ_p is simply $(1-p)\sigma_j^{(\lambda)}$, which further implies $g_j^{(\lambda)}(\sigma_p) = (1-p)g_j^{(\lambda)}(\sigma)$ when $\lambda \neq 0$. Using these to substitute σ_p for σ in Eq. (6.32), we obtain the following theorem.

Theorem 62. *Let ρ and σ be two states of systems A and B respectively, and $p \in [0,1]$ be a depolarisation probability. Given a symmetry group G whose multidimensional irreps have no multiplicity, there exists a G-covariant channel transforming ρ to $\sigma_p \coloneqq (1-p)\sigma + p\frac{I}{d}$, where d is the dimension of the output system B, if*

$$\forall \lambda \neq 0 : \begin{cases} n^{-1}\left(\mu_{\min} + \frac{p}{d(1-p)}\right) f_j^{(\lambda)}(\rho) \geq g_j^{(\lambda)}(\sigma) & \text{for } j = 0 \\ (1-p)\sigma_j^{(\lambda)} = 0 & \text{for } j \neq 0 \end{cases} \tag{6.34}$$

where we have defined $g_j^{(\lambda)}(\sigma) \coloneqq \sum_\alpha \left|\langle X_j^{(\lambda,\alpha)}, \sigma\rangle\right|$ for the ITO basis $\{X_j^{(\lambda,\alpha)}\}$ on $\mathcal{B}(B)$, $\mu_{\min}$ as the smallest eigenvalue of $\mathcal{G}(\sigma)$, and n as the number of distinct non-trivial irreps appearing in the representation of G on $\mathcal{B}(B)$.

In the case where $\mu_{\min} = 0$, we note that Theorems 61 and 62 can be slightly strengthened by applying them to an output system B' formed by truncating $\mathcal{H}_B$ to the support of $\mathcal{G}(\sigma)$ instead of B, because such a truncation has no impact on whether $\rho \xrightarrow{G} \sigma$ is possible (see Appendix D.4 for proof). We now examine the performance of these sufficient depolarisation conditions in a couple of simple examples.

6.4.3.1 Example: SU(2) represented by unitary channels of a qubit

In Example 6.3.1.2, the representation of SU(2) by unitary qubit channels on $\mathcal{B}(1)$ decomposes into a single copy of the trivial irrep, spanned by the ITO basis element $\frac{\mathbb{1}}{\sqrt{2}}$, and a single copy of a three-dimensional irrep spanned by ITO basis elements $\frac{1}{\sqrt{2}}\{X, Y, Z\}$, which we will label by $\lambda = 3$. We now apply Theorem 62 to identify the minimum depolarisation needed to transform an input state ρ to an output state σ of the same qubit under a unitarily covariant channel.

We first choose an ITO basis that satisfies the requirements of Lemma 57. All elements of the ITO basis presented in the last paragraph are Hermitian. Because each irrep in this representation of SU(2) was shown to be its own complex conjugate in Example 6.3.1.2, this basis already satisfies the first requirement of Lemma 57. All that remains is to "rotate" the ITO basis for the $\lambda = 3$ irrep until its first element points along the Bloch vector of the input state. Formally, we can denote the Bloch vector of the input state ρ by $\boldsymbol{r}_\rho$, and select three pairwise orthogonal vectors $\{\hat{\boldsymbol{r}}_\rho, \hat{\boldsymbol{n}}_1, \hat{\boldsymbol{n}}_2\}$ in 3D space, where $\hat{\boldsymbol{r}}_\rho$ is a unit vector in the direction of $\boldsymbol{r}_\rho$. One can straightforwardly verify that

$$\frac{\mathbb{1}}{\sqrt{2}} \cup \frac{1}{\sqrt{2}}\{\hat{\boldsymbol{r}}_\rho \cdot \boldsymbol{S}, \hat{\boldsymbol{n}}_1 \cdot \boldsymbol{S}, \hat{\boldsymbol{n}}_2 \cdot \boldsymbol{S}\}, \tag{6.35}$$

where $\boldsymbol{S} \coloneqq (X, Y, Z)^T$ is the vector of non-trivial Pauli operators, is an ITO basis that satisfies both conditions laid out in Lemma 57.

We now apply Theorem 62 in our chosen ITO basis, and use $\boldsymbol{r}_\sigma$ to denote the Bloch vector of the output state σ. There are now two possibilities:

Case I: $\boldsymbol{r}_\sigma$ is not parallel to $\boldsymbol{r}_\rho$. The $(\lambda = 3, j = 1, 2)$ conditions in Eq. (6.34) force the minimum sufficient depolarisation to be 1.

Case II: $\boldsymbol{r}_\sigma$ is parallel to $\boldsymbol{r}_\rho$. In this case we can write

$$\boldsymbol{r}_\sigma = \gamma \boldsymbol{r}_\rho \tag{6.36}$$

for some real number γ. This immediately satisfies the $(\lambda = 3, j = 1, 2)$ conditions in Eq. (6.34), so we turn to the $(\lambda = 3, j = 0)$ condition. Noting that $\mathcal{G}(\rho) = \frac{\mathbb{1}}{2}$, we calculate

$$f^3_{\hat{\boldsymbol{r}}_\rho}(\rho) = \frac{1}{3} \operatorname{Tr}\left[\sqrt{2}\mathbb{1}\frac{1}{2}(\boldsymbol{r}_\rho \cdot \boldsymbol{S})\sqrt{2}\mathbb{1}\frac{1}{2}(\boldsymbol{r}_\rho \cdot \boldsymbol{S})\right] = \frac{1}{3}\boldsymbol{r}_\rho^2. \tag{6.37}$$

Since $\mathcal{G}(\sigma) = \frac{\mathbb{1}}{2}$, we further have $\mu_{\min} = \frac{1}{2}$. Finally noting that Eq. (6.36) implies $g^3_{\hat{\boldsymbol{r}}_\rho} = \frac{|\gamma \boldsymbol{r}_\rho|}{\sqrt{2}}$, we substitute all these values into the $(\lambda = 3, j = 0)$ condition of Eq. (6.34) to obtain the following lower bound on sufficient depolarisation:

$$\frac{|\boldsymbol{r}_\rho|}{3\sqrt{2}} \geq (1 - p)|\gamma| \iff p \geq \frac{|\boldsymbol{r}_\rho|}{|\gamma| 3\sqrt{2}}. \tag{6.38}$$

Ref. [140] established the following necessary and sufficient condition on the existence of a unitarily covariant transition in a qubit from an input state ρ with Bloch vector $\boldsymbol{r}_\rho$ to an output state σ with Bloch vector $\boldsymbol{r}_\sigma$:

$$\boldsymbol{r}_\sigma = \gamma \boldsymbol{r}_\rho \text{ where } \gamma \in \left[-\frac{1}{3}, 1\right]. \tag{6.39}$$

Combining cases I and II (for which we note that $\sigma_p = \frac{\mathbb{1}}{2} + \frac{1}{2}(1-p)\gamma \boldsymbol{r}_\rho \cdot \boldsymbol{S}$), we see that under Theorem 62, ρ can be transformed into the depolarised output state σ_p with Bloch vector $\boldsymbol{r}_{\sigma_p}$ if

$$\boldsymbol{r}_{\sigma_p} = \gamma_p \boldsymbol{r}_\rho \text{ where } \gamma_p \in \left[-\frac{|\boldsymbol{r}_\rho|}{3\sqrt{2}}, \frac{|\boldsymbol{r}_\rho|}{3\sqrt{2}}\right]. \tag{6.40}$$

As $|\boldsymbol{r}_\rho| \leq 1$, this falls within the full range given by Eq. (6.39) and offers particularly good coverage when the input state is pure.

6.4.3.2 Example: $\mathsf{U}(1)$ represented by time evolution of a (non-degenerate) qubit

Letting $|E_0\rangle$ and $|E_1\rangle$ respectively denote the ground and excited states of a non-degenerate qubit, we saw in Example 6.3.2.1 that the representation of $\mathsf{U}(1)$ on $\mathcal{B}(1)$ as the time evolution of this qubit decomposes into two copies of the trivial irrep, spanned by ITO basis elements $|E_0\rangle\langle E_0|$ and $|E_1\rangle\langle E_1|$, one copy of the irrep labelled $\lambda = \Delta$, spanned by the ITO basis element $|E_1\rangle\langle E_0|$, and one copy of its complex conjugate labelled $\lambda = -\Delta$, spanned by the ITO basis element $|E_0\rangle\langle E_1|$. This choice of ITO basis already satisfies both requirements in Lemma 57.

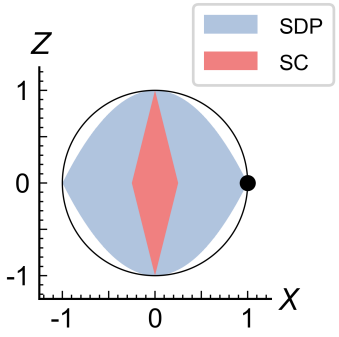


Figure 6.3: **General depolarisation conditions for state transitions covariant with time evolution in a (non-degenerate) qubit.** The black dot shows the initial qubit state $\rho = |+\rangle\langle+|$. The large blue shaded region (SDP) defines the full set of output qubit states that can be reached under G-covariant channels as determined in Ref [133], while the smaller pink shaded region (SC) overlapping this shows the region given by the conditions of Theorem 62.

Given an input state ρ and desired output state σ in our non-degenerate qubit, we now apply

Theorem 62 in our chosen ITO basis to find a depolarisation sufficient for ρ to transform into σ under a time-symmetric channel. Parameterising states of the non-degenerate qubit in the energy eigenbasis as shown in Eq. (5.22), we can write $\mathcal{G}(\rho) = p_\rho |E_0\rangle\langle E_0| + (1-p_\rho)|E_1\rangle\langle E_1|$ and recall from Example 6.3.2.1 that $\rho^{(\Delta)} = c_\rho |E_1\rangle\langle E_0| = \rho^{(\Delta)\dagger}$. We can then calculate that

$$f^{(\Delta)} = \frac{|c_\rho|^2}{\sqrt{(1-p_\rho)p_\rho}} = f^{(-\Delta)}. \tag{6.41}$$

as well as

$$g^{(\Delta)}(\sigma) = |\mathrm{Tr}[|E_1\rangle\langle E_0|\,\sigma]| = |c_\sigma| = |\mathrm{Tr}[|E_0\rangle\langle E_1|\,\sigma]| = g^{(-\Delta)}(\sigma). \tag{6.42}$$

Substituting all this back into Eq. (6.34), we obtain

$$\frac{1}{2}\left(\frac{|c_\rho|^2}{\sqrt{(1-p_\rho)p_\rho}}\right) \geq \frac{|c_\sigma|}{\min\{p_\sigma, 1-p_\sigma\} + \frac{p}{2(1-p)}} \tag{6.43}$$

whose performance we plot for the example input state $|+\rangle$ in Figure 6.3.

6.4.4 Sufficient depolarisation conditions for identical input and output systems

When the input and output systems are the same, we can use the fact that the identity channel is now G-covariant to strengthen Theorem 61. Concretely, one can, without loss of generality, replace the covariant channel Φ^η in Lemma 58 with $(1-q)\mathsf{id} + q\Phi^\eta$, where q is a probability and id is the identity channel. The sufficient conditions of Lemma 58 then become identifying a family of G-covariant channels $\{\Phi^\eta\}$ indexed by every reference frame η such that

$$(1-q)\langle\eta,\rho\rangle + q\langle\eta,\Phi^\eta(\rho)\rangle \geq \langle\eta,\sigma\rangle \text{ for all } \eta \text{ and any } q\in[0,1], \tag{6.44}$$

which can be rearranged as

$$\langle\eta, q\Phi^\eta(\rho)\rangle \geq \langle\eta,(\sigma-(1-q)\rho)\rangle \text{ for all } \eta \text{ and any } q\in[0,1]. \tag{6.45}$$

Defining

$$\sigma(q) := \sigma - (1-q)\rho \tag{6.46}$$

for any $q \in [0,1]$, we can express Eq. (6.45) more compactly as

$$\langle \eta, q\Phi^{\eta}(\rho)\rangle \geq \langle \eta, \sigma(q)\rangle \text{ for all } \eta \text{ and any } q \in [0,1], \tag{6.47}$$

which can be obtained from the original sufficient conditions in Lemma 58 by the substitution $\Phi^{\eta} \to q\Phi^{\eta}$ and $\sigma \to \sigma(q)$. By pursuing the impact of this substitution across the proof of Theorem 61, we can specialise it to the case where the input and output systems are identical (see Appendix D.3.7 for proof).

Theorem 63. *Let ρ and σ be states of system S and let $p \in [0,1]$ be a depolarisation probability. Given a symmetry group G whose multidimensional irreps have no multiplicity, we have that $\rho \xrightarrow{G} \sigma_p \coloneqq (1-p)\sigma + p\frac{I}{d}$ if $\sigma_p = \rho$ or if there exists a probability $q \in (q^*, 1]$ such that*

$$\forall \lambda \neq 0 : \begin{cases} \mu_{\min}(q) n^{-1} f_j^{(\lambda)}(\rho) \geq g_j^{(\lambda)}(\sigma_p(q)) & \text{for } j = 0 \\ (1-p)\sigma_j^{(\lambda)} = 0 & \text{for } j \neq 0. \end{cases} \tag{6.48}$$

where we have defined $\sigma_p(q) \coloneqq \sigma_p - (1-q)\rho$, $q^* \coloneqq \min\{q \in [0,1] : \mathcal{G}(\sigma_p(q)) \geq 0\}$, $g_j^{(\lambda)}(\sigma_p(q)) \coloneqq \sum_\alpha \left|\langle X_j^{(\lambda,\alpha)}, \sigma_p(q)\rangle\right|$ for the ITO basis $\{X_j^{(\lambda,\alpha)}\}$ on $\mathcal{B}(S)$, $\mu_{\min}(q)$ as the smallest eigenvalue of $\mathcal{G}(\sigma_p(q))$, and n as the number of distinct non-trivial irreps in the representation of G on $\mathcal{B}(S)$.

Once again, we examine the performance of these sufficient conditions for simple examples.

6.4.4.1 Example: SU(2) represented by unitary evolution of a qubit

Using the ITO basis and calculations developed in Example 6.4.3.1, we once again find two possibilities.

Case I: $\boldsymbol{r}_\sigma$ is not parallel to $\boldsymbol{r}_\rho$. The $(\lambda = 3, j = 1, 2)$ conditions in Eq. (6.34) force the minimum sufficient depolarisation to be 1.

Case II: $\boldsymbol{r}_\sigma$ is parallel to $\boldsymbol{r}_\rho$. In this case, we can write

$$\boldsymbol{r}_\sigma = \gamma \boldsymbol{r}_\rho \tag{6.49}$$

for some real number γ. This immediately satisfies the $(\lambda = 3, j = 1, 2)$ conditions in Eq. (6.34),

so we turn to the $(\lambda = 3, j = 0)$ condition, which becomes

$$\exists q \in [q^*, 1] : \mu_{\min}(q) \left(\frac{1}{3} \boldsymbol{r}_\rho^2 \right) \geq \frac{1}{\sqrt{2}} |(1-p)\gamma \boldsymbol{r}_\rho - (1-q)\boldsymbol{r}_\rho|. \tag{6.50}$$

Since $\mathcal{G}(\tau) = \frac{\mathbb{1}}{2}$ for any state τ, we have that $\mathcal{G}(\sigma_p(q)) = q\frac{\mathbb{1}}{2}$, so $\mu_{\min}(q) = \frac{q}{2}$, which implies $q^* = 0$. Substituting this into Eq. (6.50), we obtain the sufficient condition

$$\exists q \in [0, 1] : q \left(\frac{|\boldsymbol{r}_\rho|}{3\sqrt{2}} \right) \geq |(1-p)\gamma - (1-q)|. \tag{6.51}$$

We note that

$$\boldsymbol{r}_\sigma = \gamma \boldsymbol{r}_\rho \to \sigma_p = \frac{\mathbb{1}}{2} + \frac{1}{2}(1-p)\gamma \boldsymbol{r}_\rho \cdot \boldsymbol{S}. \tag{6.52}$$

We first consider the scenario where $\gamma \geq 0$. When $(1-p)\gamma < 1$, we can always choose a probability q such that $(1-q) = (1-p)\gamma$, which immediately satisfies Eq. (6.51). At $(1-p)\gamma = 1$ exactly, we see by Eq. (6.52) that we simply have $\sigma_p = \rho$. However, when $(1-p)\gamma > 1$, we cannot find any probability satisfying Eq. (6.51), because if such a probability q were to exist, it would always make the part on which the absolute is taken on the right hand side of Eq. (6.51) positive, which implies

$$\begin{aligned} & q \left(\frac{|\boldsymbol{r}_\rho|}{3\sqrt{2}} \right) > (1-p)\gamma - (1-q) \\ \implies & q \left(\frac{|\boldsymbol{r}_\rho|}{3\sqrt{2}} \right) + (1-q) > (1-p)\gamma \\ \implies & q + (1-q) > q \left(\frac{|\boldsymbol{r}_\rho|}{3\sqrt{2}} \right) + (1-q) \geq (1-p)\gamma \\ \implies & 1 > (1-p)\gamma \end{aligned} \tag{6.53}$$

contradicting the initial assumption of $(1-p)\gamma > 1$. We therefore conclude by Theorem 63 that when $\gamma \geq 0$, we have $\rho \xrightarrow{G} \sigma_p$ if $(1-p)\gamma \leq 1$.

We next consider the case where $\gamma < 0$, which implies that the part inside the absolute on the

right hand side of Eq. (6.51) must be non-positive, so we obtain the sufficient condition

$$\begin{aligned}\exists q \in [0,1] : & q\left(\frac{|\boldsymbol{r}_\rho|}{3\sqrt{2}}\right) \geq (1-q) - (1-p)\gamma = (1-q) + (1-p)|\gamma| \\ \implies & q\left(\frac{|\boldsymbol{r}_\rho|}{3\sqrt{2}} + 1\right) \geq 1 + (1-p)|\gamma|\end{aligned} \tag{6.54}$$

Choosing $q = 1$ to maximise the LHS, we obtain the sufficient condition

$$\frac{|\boldsymbol{r}_\rho|}{3\sqrt{2}} \geq (1-p)|\gamma|. \tag{6.55}$$

Putting everything together, Theorem 63 certifies that $\rho \xrightarrow{\mathsf{SU}(2)} \sigma_p$ if

$$\boldsymbol{r}_\sigma = \gamma \boldsymbol{r}_\rho \text{ and } (1-p)|\gamma| \leq \begin{cases} 1 & \text{if } \gamma \geq 0 \\ \frac{|\boldsymbol{r}_\rho|}{3\sqrt{2}} & \text{otherwise.} \end{cases} \tag{6.56}$$

By Eq. (6.52), the conditions in Eq. (6.56) are equivalent to saying that ρ can be transformed into the depolarised output state σ_p with Bloch vector $\boldsymbol{r}_{\sigma_p}$ in a unitarily covariant way if

$$\boldsymbol{r}_{\sigma_p} = \gamma_p \boldsymbol{r}_\rho \text{ where } |\gamma| \in \left[-\frac{|\boldsymbol{r}_\rho|}{3\sqrt{2}}, 1\right]. \tag{6.57}$$

We see that Eq. (6.57) outperforms the range of accessible states identified by depolarisation conditions that do not account for the input and output systems being the same in Eq. (6.40), because we recovered all partially depolarised ρ as unitarily accessible from ρ.

6.4.4.2 Example: $\mathsf{U}(1)$ represented by time evolution of a (non-degenerate) qubit

In Figure 6.4, we plot the range of time-covariantly accessible states from a fixed input state ρ according to the sufficient conditions in Theorem 63. Comparison to 6.3 similarly indicates that Theorem 63 provides more powerful sufficient conditions than Theorem 62.

6.4.5 Concluding Remarks

In Part III of this thesis, we have shown that the first complete set of asymmetry monotones identified by Ref. [36], given by the single-shot conditional min-entropy evaluated on the input/output state prepared relative to every possible state on a spontaneously emerging ref-

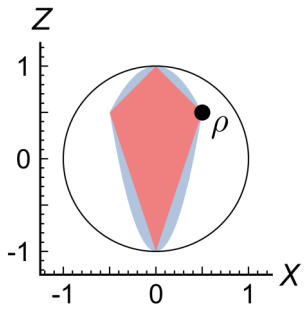


Figure 6.4: **Depolarisation conditions for state transitions covariant with time evolution in a (non-degenerate) qubit, accounting for the input and output system being the same.** The black dot shows an initial qubit state ρ with Bloch vector $\boldsymbol{r} := (\frac{1}{2}, 0, \frac{1}{2})$. The large blue shaded region (SDP) defines the full set of output qubit states that can be reached under time-symmetric channels as determined in Ref [133]. The smaller pink shaded region (SC) overlapping this shows the region given by the conditions given in Theorem 63. (Note that despite appearances the boundary of the SDP region is not linear, but curves outwards.)

erence system, are highly redundant. In general, one need only consider reference frames on any surface enclosing the maximally mixed state, so we could reduce the original infinite set of monotone conditions on G-covariant state transitions to a single local minimality condition.

The fact that all asymmetry properties of a quantum system can be captured relative to reference frames with arbitrarily small modes of asymmetry suggests a deeper description in terms of differential geometry, at least when allowing a small amount of noise in transitions would smooth out conical singularity issues such as those discussed at the end of Chapter 5. We expect this description to take the form of a Fisher-like information [125, 30, 34]. In particular, it would be interesting to see whether one can replace $H_{\min}(R|A)$ with the conditional von-Neumann entropy $H(R|A)$, which would allow explicit analytic computations.

Applying the single minimality condition for asymmetry obtained in Chapter 5 presents great technical challenges. Therefore, in Chapter 6, we allow noise to affect transitions and thereby extract two *finite* sets of sufficient conditions on how far the output state should be depolarised to make a G-covariant transition possible. The two sets of sufficient depolarisation conditions respectively rely on ϵ-smoothing of the reference frames and the separation of modes of asymmetry under G-covariant transitions. However, while the ϵ-smoothing conditions may only be practical for low-dimensional systems, the modal conditions are restricted to (representations of) symmetry groups whose multidimensional irreps have no multiplicity.

Nevertheless, the modal sufficient depolarisation conditions are very easy to compute and, in simple examples, provide decent coverage of the entire set of G-covariantly accessible states

from a fixed initial state ρ. By considering different families of G-covariant channels than the PGM measure-and-prepare schemes used in Chapter 6, we anticipate that the modal sufficient depolarisation conditions can almost certainly be improved upon. In particular it would be exciting to see whether one can make choices that lift the restriction to (representations of) groups with multiplicity-free multidimensional irreps.

Beyond this, a range of other interesting questions exist. For example we have not exploited the duality relations [131, 141] between $H_{\min}(R|A)$ and the single-shot quantum conditional min-entropy $H_{\max}(R|S)$, where S is a purifying system for the state Ω_{RA}. For example, for the case of time-translation symmetry the joint purified state admits two notable forms. The first is an energetic form:

$$\Omega_{RAS} = \sum_E \sqrt{p(E)}\, |\varphi(E)\rangle_{RA} \otimes |E\rangle_S \tag{6.58}$$

obtained from considering $\Omega_{RA} = \mathcal{G}(\eta_R \otimes \rho_A) = \sum_E \Pi(E)(\eta_R \otimes \rho_A)\Pi(E)$ as an ensemble of states over energy sectors, $\Pi(E)$ being the projector onto the energy E subspace of RA. While the second is a temporal form, given by

$$\Omega_{RAS} = \int dt \;\; |\phi(t)\rangle_{RAS}\,, \tag{6.59}$$

with $|\phi(t)\rangle_{RAS}$ being a purification of $\mathcal{U}_t(\eta_R) \otimes \mathcal{U}_t(\rho_A)$. It would be of interest to explore these two forms and also their connection to entropic uncertainty relations. In particular, as $H_{\max}$ has been identified as the correct information measure for the causal development of a spatial region [123], we hope that exploiting its duality relations with $H_{\min}$ might allow the characterisation of quantum reference frames in terms of $H_{\min}$ continued in this work to feed into future research on quantum gravity.

Finally, it would also be valuable to see how the explicit conditions given by Theorem 62 and Theorem 63 could be used in concrete settings, such as for covariant quantum error-correcting codes [105, 103, 106], thermodynamics [36] or metrology [101]. Moreover, the method of constructing these conditions can certainly be improved upon by using more detailed covariant protocols.

Part IV

Concluding Materials

Chapter 7

Summary of results

This thesis has built on and contributed to the characterisation, within a resource-theoretic framework, of properties in quantum systems that can be harnessed for information-processing. We have focussed in particular on properties described by generalised quantum entropies in the resource theories of magic and asymmetry, wherein we have considered speakable and unspeakable resources in tasks that are narrow and practical as well as general and foundational in scope. The theme of resource theories unify the two research directions making up this thesis, which are respectively the identification of constraints on magic distillation arising from its representation as statistical mechanics on quasiprobability distributions, and the identification of finite sets of conditions on when a state transition can occur under a (compact) symmetry constraint. We now summarise the main results from each part of the thesis in turn.

Part I of this thesis concerns magic distillation, a central component of fault-tolerant quantum computing via magic state injection that was first experimentally demonstrated in December 2024 [19]. In Chapter 3, our main result extends a statistical mechanics framework developed in Ref. [27] for magic distillation on qudits of odd prime dimension to the technologically more important case of qubits. This framework recasts qubit computing by state injection in a phase-space setting that respects the sequential and parallel composition of processes, wherein stochastic transformations representing fault-tolerant channels process quasiprobability distributions representing quantum states, and induce a pre-order on such distributions that is captured by a dense subset of α-Réyni entropies. In Chapter 4, we apply these abstract resource-theoretic tools to establish fundamental bounds on the performance of real-world magic distillation protocols, in particular on those based on projection onto CSS codes. Our

main result consists of analytic trade-off relations between parameters governing the performance of such protocols, including input and output error rates, success probability and code parameters. These trade-off relations may be instructive for designing protocols with optimised parameters.

Part II of this thesis concerns restrictions imposed by compact symmetries on quantum state transitions. In Chapter 5, we show that the first complete but infinite set of asymmetry measures [36], given by quantum conditional min-entropies capturing the strength of correlations between the input/output state and all possible states of a reference system, possess extensive redundancies. This work continues the characterisation of quantum reference frames in terms of single-shot entropies begun in Ref. [36], and we hope that this combination of tools of active interest in recent quantum gravity research can contribute to future work in that area. Our main result establishes that one need only consider entropic measures produced by reference frames that are arbitrarily close to symmetric, from which we obtain a single minimality condition that completely determines whether a state transition can occur under a given symmetry constraint. As this condition is rather technically intractable, we weaken the demand for exact state transitions in Chapter 6. Our main result is a finite set of closed-form sufficient conditions on the depolarisation of the output state needed to ensure a state transition that respects any symmetry group whose multidimensional irreps have no multiplicity. All Abelian groups as well as important representations of non-Abelian groups such as $\mathsf{SU}(2)$ on a spin-j system fall under the scope of this result.

Appendices

Appendix A

Appendices to Chapter 3

A.1 Characterisation of stabilizer groups (proof of theorem 5)

We begin with the characterisation of stabilizer groups provided by the following Lemma.

Lemma 64. *A subgroup* S *of the* n*-qubit Pauli group* P_n *is a stabilizer group if and only if* $-I \notin \mathsf{S}$*, which further implies that the order of* S *is related to the dimension of its stabilized subspace* $\mathcal{H}_{\mathsf{S}}$ *as* $|\mathsf{S}| \dim(\mathcal{H}_{\mathsf{S}}) = 2^n$.

Proof. We begin by assuming that $-I \notin S$, from which we will prove that S is a stabilizer group. To do so, we first need to define the operator

$$\Pi \coloneqq \frac{1}{|\mathsf{S}|} \sum_{P \in \mathsf{S}} P. \tag{A.1}$$

The unitarity of Pauli operators, as well as the group property of S, allow us show that

$$\Pi^\dagger = \frac{1}{|\mathsf{S}|} \sum_{P \in \mathsf{S}} P^\dagger = \frac{1}{|\mathsf{S}|} \sum_{P \in \mathsf{S}} P^{-1} = \Pi. \tag{A.2}$$

The group property of S further enables us to show that

$$\Pi^2 = \frac{1}{|\mathsf{S}|^2} \sum_{P \in \mathsf{S}} P \left(\sum_{P' \in \mathsf{S}} P' \right) = \frac{1}{|\mathsf{S}|^2} \sum_{P \in \mathsf{S}} \left(\sum_{P'' \in \mathsf{S}} P'' \right) = \frac{1}{|\mathsf{S}|^2} |\mathsf{S}| \left(\sum_{P'' \in \mathsf{S}} P'' \right) = \Pi, \tag{A.3}$$

where $P'' \coloneqq PP'$. When we put these two equations together, we can conclude that Π is a *projective operator*, and can therefore define a subspace $\mathcal{H}_\Pi$ in the Hilbert space of n qubits, $\mathcal{H}_n$,

which forms the space that Π projects onto. The dimension of this subspace is

$$\dim(\mathcal{H}_\Pi) = \mathrm{Tr}[\Pi] = \frac{1}{|\mathsf{S}|}\sum_{P\in\mathsf{S}} \mathrm{Tr}[P]. \tag{A.4}$$

Because X, Y and Z are traceless, every Pauli operator *except* for $\pm I$ and $\pm iI$ is also traceless. Since S is a group, I must be a member of S. However, since $(\pm iI)^2 = -I$, the assumption $-I \notin \mathsf{S}$ implies $\pm iI \notin \mathsf{S}$. Therefore, the only member of S that can contribute to the above sum is I, and we conclude that

$$\dim(\mathcal{H}_\Pi) = \frac{1}{|\mathsf{S}|}\mathrm{Tr}[I] = \frac{2^n}{|\mathsf{S}|}, \tag{A.5}$$

which implies $\mathcal{H}_\Pi$ cannot be empty. Now given any vector $|\psi\rangle \in \mathcal{H}_\Pi$ and any element $P \in \mathsf{S}$, we have that

$$\Pi|\psi\rangle = P\Pi|\psi\rangle = \left(\frac{1}{|\mathsf{S}|}P\sum_{P'\in\mathsf{S}} P'\right)|\psi\rangle = \left(\frac{1}{|\mathsf{S}|}\sum_{P''\in\mathsf{S}} P''\right)|\psi\rangle = \Pi|\psi\rangle = |\psi\rangle\,, \tag{A.6}$$

where once again $P'' \coloneqq PP'$, and we used the group property of S in the second equality. From the above equation, we see that any $|\psi\rangle \in \mathcal{H}_\Pi$ lies in $\mathcal{H}_\mathsf{S}$, and therefore $\mathcal{H}_\Pi \subseteqq \mathcal{H}_\mathsf{S}$. We thus conclude that $\mathcal{H}_\mathsf{S}$ is not empty, so S is a stabilizer group.

We now assume that S is a stabilizer group, from which we will prove that $-I \notin \mathsf{S}$. Since $\mathcal{H}_\mathsf{S}$, the Hilbert space stabilized by S, is not empty in this case, we can consider a *non-zero* vector $|\psi\rangle \in \mathcal{H}_\mathsf{S}$. If $-I \in \mathsf{S}$, then $-I|\psi\rangle = |\psi\rangle$, which produces the contradiction that $|\psi\rangle$ must be the zero vector. We therefore conclude that $-I \notin \mathsf{S}$.

We have now proven that S is a stabilizer group if and only if $-I \notin \mathsf{S}$, and will conclude by showing why this implies the relationship between the order S and the dimension of its stabilized subspace given in the lemma. Let us consider an arbitrary vector $|\psi\rangle \in \mathcal{H}_\mathsf{S}$. Then

$$\Pi|\psi\rangle = \frac{1}{|\mathsf{S}|}\sum_{P\in\mathsf{S}} P|\psi\rangle = \frac{1}{|\mathsf{S}|}\sum_{P\in\mathsf{S}} |\psi\rangle = |\psi\rangle\,, \tag{A.7}$$

which implies $|\psi\rangle \in \mathcal{H}_\Pi$ for any $|\psi\rangle \in \mathcal{H}_\mathsf{S}$, and therefore $\mathcal{H}_\mathsf{S} \subseteqq \mathcal{H}_\Pi$. Since we had already shown that S is a stabilizer group if and only if $-I \notin \mathsf{S}$, and furthermore that $-I \notin \mathsf{S}$ implies $\mathcal{H}_\Pi \subseteqq \mathcal{H}_\mathsf{S}$, we conclude that $\mathcal{H}_\Pi = \mathcal{H}_\mathsf{S}$. By substituting this equality into Eq. (A.5), we immediately

conclude that $|\mathsf{S}| \dim(\mathcal{H}_\mathsf{S}) = 2^n$. □

We next verify some more basic properties of stabilizer groups.

Lemma 65. *Every stabilizer group is an Abelian group of Pauli observables.*

Proof. Let S be a stabilizer group. Because Pauli operators that are not an observable cannot stabilize the non-zero state vectors that exist by definition in the stabilized subspace of S, every element of S must be a Pauli observable. We now prove that S is Abelian by contradiction. If S were not Abelian, then by Lemma 3, it must contain two non-trivial Pauli observables P and P' that anti-commute, i.e. $PP' = -P'P$, which implies $(PP')PP' = (-P'P)PP' = -I \in \mathsf{S}$ by the group property of S and the fact that every non-trivial Pauli observable is of order 2. Since this contradicts S being a stabilizer group by Lemma 64, we conclude that S must be Abelian. □

A finite Abelian group S generated by a set of m commuting, independent and non-trivial generators $G \coloneqq \{g_1, \ldots, g_m\}$, each of which has order 2, can be characterised as

$$\mathsf{S} = \{g_1^{b_1} \ldots g_m^{b_m} | b_1, \ldots, b_m \in \{0, 1\}\}. \tag{A.8}$$

Because each b_i is a bit and all the generators commute, this characterisation is equivalent to the set for products of elements in all possible subsets of G, so we can write

$$\mathsf{S} = \left\{ \prod_{g_i \in X} g_i \middle| X \in P(G) \right\}, \tag{A.9}$$

where $P(G)$ is the power set (i.e. set of all subsets) of G.

With the above lemmas and discussion in hand, we are now in a position to prove Theorem 5.

Theorem 5. *A subgroup S of the n-qubit Pauli group P_n is a stabilizer group if and only if it is an Abelian group generated by a set of $n \geq m \geq 0$ commuting, independent and non-trivial Pauli observables such that $-I \notin \mathsf{S}$ (with $m = 0$ corresponding to S being the trivial group). The stabilized subspace $\mathcal{H}_\mathsf{S}$ of S then has dimension $\dim(\mathcal{H}_\mathsf{S}) = 2^{n-m}$.*

Proof. We begin by assuming S is a stabilizer group and proving the forward direction. By Lemma 65, we immediately conclude that S is either trivial or generated from a set of $m \geq 1$ independent and non-trivial Pauli observables $G \coloneqq \{g_1, \ldots, g_m\}$ that commute. Since every non-trivial Pauli observable has order 2, we can alternatively characterise S as shown in

Eq. (A.8) and Eq. (A.9). By Lemma 64, we know that $-I \notin \mathsf{S}$, so the characterisation of S given in Eq. (A.9) demonstrates that $-I$ is not a product of elements from any subset in G. Furthermore, the characterisation given of S given in Eq. (A.8) also implies $|\mathsf{S}| = 2^m$. Together with Lemma 64, we arrive at $\dim(\mathcal{H}_{\mathsf{S}}) = 2^{n-m}$, which places an upper bound $n \geq m$ on m.

In the reverse direction, we have that S is either trivial or we can begin immediately with the alternative characterisation of S in Eq. (A.9), from which we see that $-I$ not being the product of elements in any subset of G implies $-I \notin \mathsf{S}$. By Lemma 64, S is a stabilizer group. □

A.2 Properties of the Wigner representation of qubit states

We begin with a few remarks on the displacement operators defined in Eq. (3.11). Because $XZ = -ZX$, these displacement operators satisfy the commutation relation

$$\begin{aligned} D_{\boldsymbol{u}} D_{\boldsymbol{v}} &= \bigotimes_{i=1}^{n} Z^{(u_z)_i} X^{(u_x)_i} Z^{(v_z)_i} X^{(v_x)_i} \\ &= \bigotimes_{i=1}^{n} (-1)^{(u_x)_i (v_z)_i} Z^{(u_z)_i} Z^{(v_z)_i} X^{(u_x)_i} X^{(v_x)_i} \\ &= \bigotimes_{i=1}^{n} (-1)^{((u_x)_i (v_z)_i + (u_z)_i (v_x)_i)} Z^{(v_z)_i} X^{(v_x)_i} Z^{(u_z)_i} X^{(u_x)_i} = (-1)^{[\boldsymbol{u}\cdot\boldsymbol{v}]} D_{\boldsymbol{v}} D_{\boldsymbol{u}}. \end{aligned} \tag{A.10}$$

Furthermore, the displacement operators are unitary:

$$D_{\boldsymbol{u}} D_{\boldsymbol{u}}^\dagger = \bigotimes_{i=1}^{n} Z^{(u_z)_i} X^{(u_x)_i} X^{(u_x)_i} Z^{(u_z)_i} = I = \bigotimes_{i=1}^{n} X^{(u_x)_i} Z^{(u_z)_i} Z^{(u_z)_i} X^{(u_x)_i} = D_{\boldsymbol{u}}^\dagger D_{\boldsymbol{u}}. \tag{A.11}$$

A.2.1 Properties of qubit phase-point operators (proof of 12)

We note that $A_{\mathbf{0}}$ can also be cast into the form $A_{\mathbf{0}} = D_{\mathbf{0}} A_{\mathbf{0}} D_{\mathbf{0}}^\dagger$ as $D_{\mathbf{0}} = I$. Therefore, as a consequence of Eq. (A.10), the phase-point operators we have defined can much more usefully be re-expressed as

$$A_{\boldsymbol{u}} = \frac{1}{2^n} \sum_{\boldsymbol{v} \in \mathcal{P}_n} D_{\boldsymbol{u}} D_{\boldsymbol{v}} D_{\boldsymbol{u}}^\dagger = \frac{1}{2^n} \sum_{\boldsymbol{v} \in \mathcal{P}_n} (-1)^{[\boldsymbol{u},\boldsymbol{v}]} D_{\boldsymbol{v}} D_{\boldsymbol{u}} D_{\boldsymbol{v}}^\dagger = \frac{1}{2^n} \sum_{\boldsymbol{v} \in \mathcal{P}_n} (-1)^{[\boldsymbol{u},\boldsymbol{v}]} D_{\boldsymbol{v}}. \tag{A.12}$$

Lemma 12. *The phase point operators $\{A_{\boldsymbol{u}}\}_{\boldsymbol{u} \in \mathcal{P}_n}$ have the following properties:*

(A1) *Factorizability: $A_{\boldsymbol{u}_X \oplus \boldsymbol{u}_Y} = A_{\boldsymbol{u}_X} \otimes A_{\boldsymbol{u}_Y}$ on a bipartite system XY, where $\boldsymbol{u}_X$ and $\boldsymbol{u}_Y$ are respectively points in the phase spaces of subsystems X and Y.*

(A2) *Orthogonality:* $\langle A_{\boldsymbol{u}}, A_{\boldsymbol{v}} \rangle = 2^n \delta_{\boldsymbol{u},\boldsymbol{v}}$,

(A3) *Unit trace:* $\mathrm{Tr}[A_{\boldsymbol{u}}] = 1$,

(A4) *Completeness:* $\sum_{\boldsymbol{u} \in \mathcal{P}_n} A_{\boldsymbol{u}} = 2^n I$.

Proof. We prove each property listed by the Lemma in turn.

Proof of Property (A1).

From the definition of $D_{\boldsymbol{u}}$ in Eq. (3.11), it is clear that

$$D_{\boldsymbol{u}} = D_{\boldsymbol{u}_X} \otimes D_{\boldsymbol{u}_Y}. \tag{A.13}$$

Let n_X and n_Y be the numbers of qubits in subsystems X and Y respectively. Then the zero phase point operator on the bipartite system, $A_{\mathbf{0}}$, is

$$\begin{aligned} A_{\mathbf{0}} \coloneqq \frac{1}{2^{n_X+n_Y}} \sum_{\boldsymbol{u} \in \mathcal{P}_{XY}} D_{\boldsymbol{u}} = \frac{1}{2^{n_X+n_Y}} \sum_{\boldsymbol{u}_X \in \mathcal{P}_X, \boldsymbol{u}_Y \in \mathcal{P}_Y} D_{\boldsymbol{u}_X \oplus \boldsymbol{u}_Y} = \frac{1}{2^{n_X+n_Y}} \sum_{\boldsymbol{u}_X \in \mathcal{P}_X} \sum_{\boldsymbol{u}_Y \in \mathcal{P}_Y} D_{\boldsymbol{u}_X} \otimes D_{\boldsymbol{u}_Y} \\ = A_{\mathbf{0}_X} \otimes A_{\mathbf{0}_Y}, \end{aligned} \tag{A.14}$$

which in turn implies that any phase point operator $A_{\boldsymbol{u}} \coloneqq A_{\boldsymbol{u}_X \oplus \boldsymbol{u}_Y}$ for some $\boldsymbol{u}_X \in \mathcal{P}_X$ and $\boldsymbol{u}_Y \in \mathcal{P}_Y$ is

$$\begin{aligned} A_{\boldsymbol{u}} \coloneqq A_{\boldsymbol{u}_X \oplus \boldsymbol{u}_Y} = D_{\boldsymbol{u}_X \oplus \boldsymbol{u}_Y} A_{\mathbf{0}_X \oplus \mathbf{0}_Y} D^\dagger_{\boldsymbol{u}_X \oplus \boldsymbol{u}_Y} = \left(D_{\boldsymbol{u}_X} A_{\mathbf{0}_X} D^\dagger_{\boldsymbol{u}_X} \right) \otimes \left(D_{\boldsymbol{u}_Y} A_{\mathbf{0}_Y} D^\dagger_{\boldsymbol{u}_Y} \right) \\ = A_{\boldsymbol{u}_X} \otimes A_{\boldsymbol{u}_Y}, \end{aligned} \tag{A.15}$$

as claimed.

Property **(A1)** enables us to break down any n-qubit phase-point operator $A_{\boldsymbol{u}}$ as a tensor product of single-qubit phase-point operators,

$$A_{\boldsymbol{u}} = \bigotimes_{i=1}^{n} A_{\boldsymbol{u}_j}, \quad \boldsymbol{u} = \bigoplus_{i=1}^{n} \boldsymbol{u}_j, \tag{A.16}$$

where $\boldsymbol{u}_j \in \mathbb{Z}_2 \times \mathbb{Z}_2$ is a co-ordinate in the phase space of the jth qubit *only*. It is therefore

instructive to calculate the single-qubit phase point operators, which are

$$\begin{aligned} A_{0,0} &= \frac{1}{2}(\mathbb{1} + X + Z + iY), \\ A_{0,1} &= \frac{1}{2}(\mathbb{1} - X + Z - iY), \\ A_{1,0} &= \frac{1}{2}(\mathbb{1} + X - Z - iY), \\ A_{1,1} &= \frac{1}{2}(\mathbb{1} - X - Z + iY). \end{aligned} \tag{A.17}$$

Proof of Property (A2). Let us first decompose $\boldsymbol{u}$ and $\boldsymbol{v}$ as $\boldsymbol{u} = \bigoplus_{i=1}^{n} \boldsymbol{u}_j$ and $\boldsymbol{v} = \bigoplus_{i=1}^{n} \boldsymbol{v}_j$, where $\boldsymbol{u}_j$ and $\boldsymbol{v}_j$ are phase point co-ordinates on the jth qubit only. By Eq. (A.17),

$$\left\langle A_{\boldsymbol{u}_j}, A_{\boldsymbol{v}_j} \right\rangle = 2\delta_{\boldsymbol{u}_j, \boldsymbol{v}_j}. \tag{A.18}$$

Therefore,

$$\langle A_{\boldsymbol{v}}, A_{\boldsymbol{u}} \rangle = \prod_{j=1}^{n} \left\langle A_{\boldsymbol{u}_j}, A_{\boldsymbol{v}_j} \right\rangle = \prod_{j=1}^{n} 2\delta_{\boldsymbol{v}_j, \boldsymbol{u}_j} = 2^n \delta_{\boldsymbol{u}, \boldsymbol{v}}. \tag{A.19}$$

as claimed.

Proof of Property (A3). Let us first decompose $\boldsymbol{u}$ as $\boldsymbol{u} = \bigoplus_{j=1}^{n} \boldsymbol{u}_j$, where $\boldsymbol{u}_j$ is a point in the phase space of the jth qubit. We see that $\mathrm{Tr}\left[A_{\boldsymbol{u}_j}\right] = 1$ from Eq. (A.17). Therefore,

$$\mathrm{Tr}[A_{\boldsymbol{u}}] = \prod_{j=1}^{n} \mathrm{Tr}\left[A_{\boldsymbol{u}_j}\right] = \prod_{j=1}^{n} 1 = 1, \tag{A.20}$$

as claimed.

Proof of Property (A4). We can directly calculate that

$$\sum_{\boldsymbol{u} \in \mathcal{P}_n} A_{\boldsymbol{u}} = \sum_{\boldsymbol{u}_1 \in \mathcal{P}_1}, \ldots, \sum_{\boldsymbol{u}_n \in \mathcal{P}_1} \left(\bigotimes_{j=1}^{n} A_{\boldsymbol{u}_j} \right) = \bigotimes_{j=1}^{n} \left(\sum_{\boldsymbol{u}_j \in \mathcal{P}_1} A_{\boldsymbol{u}_j} \right). \tag{A.21}$$

Using the explicit forms of single-qubit phase-point operators in Eq. (A.17), we calculate that

$$\sum_{\boldsymbol{v} \in \mathcal{P}_1} A_{\boldsymbol{v}} = 2\mathbb{1}, \tag{A.22}$$

from which the result immediately follows. □

A.2.2 Shared properties with Gross's Wigner representation (proof of Lemma 13)

Lemma 13. *The Wigner representation W_ρ for qubit states has the following properties*

(W1) *Informational completeness: Given any qubit state ρ, one can decompose it uniquely as $\rho = \sum_{\boldsymbol{u}\in\mathcal{P}_n} W_\rho(\boldsymbol{u})A_{\boldsymbol{u}}$ because $\{A_{\boldsymbol{u}}\}_{\boldsymbol{u}\in\mathcal{P}_n}$ form a complete orthogonal basis for $\mathcal{B}(n)$ under the Hilbert-Schmidt inner product.*

(W2) *Normalization: the function representing any qubit state ρ sums to 1 over phase-space, i.e. $\sum_{\boldsymbol{u}\in\mathcal{P}_n} W_\rho(\boldsymbol{u}) = 1$.*

Proof. We prove each property stated by the Lemma in turn.

Proof of Property (W1). There are $|P_n| = |\mathbb{Z}_2^n \times \mathbb{Z}_2^n| = 4^n$ phase point operators on n-qubits. Property (A2) thus implies $\{A_{\boldsymbol{u}}\}_{\boldsymbol{u}\in\mathcal{P}_n}$ forms an orthogonal complex basis for the complex vector space $\mathbb{C}^{2^n\times 2^n}$ of $2^n \times 2^n$ complex matrices under the Hilbert-Schmidt inner product, where each phase-point operator has a Hilbert-Schmidt norm of $\sqrt{2^n}$. Therefore, any n-qubit quantum state ρ can be uniquely decomposed as $\rho = \sum_{\boldsymbol{u}\in\mathcal{P}_n} W_\rho(\boldsymbol{u})A_{\boldsymbol{u}}$ for $W_\rho(\boldsymbol{u}) \coloneqq \frac{1}{2^n}\langle A_{\boldsymbol{u}}, \rho\rangle$.

Proof of Property (W2). Since any n-qubit state ρ has trace 1,

$$1 = \mathrm{Tr}[\rho] = \mathrm{Tr}\left[\sum_{\boldsymbol{u}\in\mathcal{P}_n} W_\rho(\boldsymbol{u})A_{\boldsymbol{u}}\right] = \sum_{\boldsymbol{u}\in\mathcal{P}_n} W_\rho(\boldsymbol{u})\,\mathrm{Tr}[A_{\boldsymbol{u}}] = \sum_{\boldsymbol{u}\in\mathcal{P}_n} W_\rho(\boldsymbol{u}), \tag{A.23}$$

where the last equality is established by Property (A3). □

A.2.3 Properties of rebit representation

Any n-qubit state ρ can be decomposed as

$$\rho = \left[\frac{1}{2}\left(\rho + \rho^T\right)\right] + i\left[\frac{-i}{2}\left(\rho - \rho^T\right)\right], \tag{A.24}$$

where the transposition is taken with respect to the computational basis. Because $\rho^* = \rho^T$ in any basis, we can identify

$$\rho^{(0)} \coloneqq \frac{1}{2}(\rho + \rho^T) = \mathfrak{Re}(\rho), \tag{A.25}$$

$$\rho^{(1)} = \frac{-i}{2}(\rho - \rho^T) = \mathfrak{Im}(\rho), \tag{A.26}$$

i.e., $\rho^{(0)}$ and $\rho^{(1)}$ are respectively the real and imaginary components of the density matrix of ρ in the computational basis.

With this notation in place, we now prove Lemma 14.

Lemma 14. *Given any n-qubit quantum state ρ,*

$$\mathfrak{Re}[W_\rho(\boldsymbol{u})] = W_{\mathfrak{Re}(\rho)}(\boldsymbol{u}) \tag{3.16}$$

$$\mathfrak{Im}[W_\rho(\boldsymbol{u})] = W_{\mathfrak{Im}(\rho)}(\boldsymbol{u}) \tag{3.17}$$

for all $\boldsymbol{u} \in \mathcal{P}_n$, where $\mathfrak{Re}(\rho)$ and $\mathfrak{Im}(\rho)$ are respectively the real and imaginary parts of the density matrix of ρ in the computational basis.

Proof. Adopting the identification $\rho^{(0)} = \mathfrak{Re}(\rho)$ and $\rho^{(1)} = \mathfrak{Im}(\rho)$, we can then decompose $W_\rho(\boldsymbol{u})$ as

$$W_\rho(\boldsymbol{u}) = \frac{1}{2^n}\left\langle A_{\boldsymbol{u}}, \rho^{(0)}\right\rangle + i\frac{1}{2^n}\left\langle A_{\boldsymbol{u}}, \rho^{(1)}\right\rangle. \tag{A.27}$$

We recall that $A_{\boldsymbol{u}}$ is real for all $\boldsymbol{u} \in \mathcal{P}_n$. Furthermore both $\rho^{(0)}$ and $\rho^{(1)}$ are real by construction. We then conclude that $\left\langle A_{\boldsymbol{u}}, \rho^{(0)}\right\rangle$ and $\left\langle A_{\boldsymbol{u}}, \rho^{(1)}\right\rangle$ are both real for all $\boldsymbol{u} \in \mathcal{P}_n$. Therefore, we can match the real/imaginary component of W_ρ with those of its density matrix in the computational basis

$$\mathfrak{Re}(W_\rho(\boldsymbol{u})) = \frac{1}{2^n}\left\langle A_{\boldsymbol{u}}, \rho^{(0)}\right\rangle = W_{\rho^{(0)}}(\boldsymbol{u}) \tag{A.28}$$

$$\mathfrak{Im}(W_\rho(\boldsymbol{u})) = \frac{1}{2^n}\left\langle A_{\boldsymbol{u}}, \rho^{(1)}\right\rangle = W_{\rho^{(1)}}(\boldsymbol{u}) \tag{A.29}$$

as claimed. □

Ref. [73] introduced an n-rebit Wigner representation $W_\rho^{(0)}$ defined as

$$W_\rho^{(0)}(\boldsymbol{u}) \coloneqq \frac{1}{2^n}\left\langle A_{\boldsymbol{u}}^{(0)}, \rho\right\rangle, \tag{A.30}$$

for all $\boldsymbol{u} \in \mathcal{P}_n$, where

$$A_{\boldsymbol{u}}^{(0)} \coloneqq \frac{1}{2^n}\sum_{\boldsymbol{v}\in\mathcal{P}_n^0}(-1)^{[\boldsymbol{u},\boldsymbol{v}]}D_{\boldsymbol{v}} \tag{A.31}$$

for the subset of phase space

$$\mathcal{P}_n^0 \coloneqq \{\boldsymbol{v} \,:\, \boldsymbol{v}_x \cdot \boldsymbol{v}_z = 0\}\,. \tag{A.32}$$

We now show that the rebit Wigner representation introduced in Ref. [73] has an extremely useful relationship to the full qubit Wigner representation we have shown.

Lemma 66. *W_ρ reduces to $W_\rho^{(0)}$ on all rebit states.*

Proof. The definition of qubit displacement operators in Eq. (3.11) implies the relations

$$\boldsymbol{u}_x \cdot \boldsymbol{u}_z = 0 \Leftrightarrow D_{\boldsymbol{u}}^\dagger = D_{\boldsymbol{u}}^T = D_{\boldsymbol{u}} \tag{A.33}$$

$$\boldsymbol{u}_x \cdot \boldsymbol{u}_z = 1 \Leftrightarrow D_{\boldsymbol{u}}^\dagger = D_{\boldsymbol{u}}^T = -D_{\boldsymbol{u}}, \tag{A.34}$$

We see from Eq. (A.33) that $\mathcal{P}_n^{(0)}$ is the subset of phase-space displacements corresponding to *real symmetric* displacement operators. The difference between W_ρ and $W_\rho^{(0)}$ thus comes down to the fact that $A_{\boldsymbol{u}}$ sums over all displacement operators while $A_{\boldsymbol{u}}^{(0)}$ only sums over real symmetric ones, which implies $A_{\boldsymbol{u}}^{(0)}$ is itself real symmetric.

We further observe from Eq. (A.34) that the complement of $\mathcal{P}_n^{(0)}$ in $\mathcal{P}_n$,

$$\mathcal{P}_n^{(1)} \coloneqq \{\boldsymbol{u} \,:\, \boldsymbol{u}_x \cdot \boldsymbol{u}_z = 1\}, \tag{A.35}$$

is the subset of phase-space displacements for *real anti-symmetric* displacement operators. Paralleling Eq. (A.31), we then introduce the set of real anti-symmetric phase-point operators

$$A_{\boldsymbol{u}}^{(1)} \coloneqq \frac{1}{2^n} \sum_{\boldsymbol{a} \in \mathcal{P}_n^{(1)}} (-1)^{[\boldsymbol{u},\boldsymbol{a}]} D_{\boldsymbol{a}} \text{ for any } \boldsymbol{u} \in \mathcal{P}_n, \tag{A.36}$$

By Eq. (A.31) and Eq. (A.36), each $A_{\boldsymbol{u}}$ splits up as

$$A_{\boldsymbol{u}} = A_{\boldsymbol{u}}^{(0)} + A_{\boldsymbol{u}}^{(1)}. \tag{A.37}$$

We can correspondingly split up the Wigner representation of ρ as

$$\begin{aligned} W_\rho(\boldsymbol{u}) &= \frac{1}{2^n}\langle A_{\boldsymbol{u}}, \rho\rangle = \frac{1}{2^n}\left\langle A_{\boldsymbol{u}}^{(0)} + A_{\boldsymbol{u}}^{(1)}, \rho\right\rangle \\ &= \frac{1}{2^n}\left\langle A_{\boldsymbol{u}}^{(0)}, \rho\right\rangle + \frac{1}{2^n}\left\langle A_{\boldsymbol{u}}^{(1)}, \rho\right\rangle \\ &=: W_\rho^{(0)}(\boldsymbol{u}) + W_\rho^{(1)}(\boldsymbol{u}), \end{aligned} \tag{A.38}$$

where we have defined

$$W_\rho^{(k)}(\boldsymbol{u}) := \frac{1}{2^n}\left\langle A_{\boldsymbol{u}}^{(k)}, \rho\right\rangle \tag{A.39}$$

We can now prove that the real and imaginary components of W_ρ coincide with $W_\rho^{(0)}$ and $W_\rho^{(1)}$ respectively. Because $A_{\boldsymbol{u}}^{(0)}$ and $A_{\boldsymbol{u}}^{(1)}$ are respectively real symmetric and real anti-symmetric, we have $A_{\boldsymbol{u}}^{(0)\dagger} = A_{\boldsymbol{u}}^{(0)T} = A_{\boldsymbol{u}}^{(0)}$ and $A_{\boldsymbol{u}}^{(1)\dagger} = A_{\boldsymbol{u}}^{(1)T} = -A_{\boldsymbol{u}}$. Thus for $k = 0, 1$, we obtain

$$\begin{aligned} \left[W_\rho^{(k)}(\boldsymbol{u})\right]^* &= \frac{1}{2^n}\left\langle A_{\boldsymbol{u}}^{(k)}, \rho\right\rangle^* \\ &= \frac{1}{2^n}\operatorname{Tr}\left[\left(A_{\boldsymbol{u}}^{(k)\dagger}\rho\right)^\dagger\right] \\ &= \frac{1}{2^n}\operatorname{Tr}\left[A_{\boldsymbol{u}}^{(k)}\rho\right] \\ &= \frac{1}{2^n}\operatorname{Tr}\left[(-1)^k A_{\boldsymbol{u}}^{(k)\dagger}\rho\right] \\ &= (-1)^k W_\rho^{(k)}(\boldsymbol{u}). \end{aligned} \tag{A.40}$$

This implies $W_\rho^{(0)}(\boldsymbol{u})$ is the real component of $W_\rho(\boldsymbol{u})$ while $W_\rho^{(1)}(\boldsymbol{u})$ is the imaginary component of $W_\rho(\boldsymbol{u})$. Thus given any n-qubit state ρ, we can write

$$\mathfrak{Re}(W_\rho(\boldsymbol{u})) = W_\rho^{(0)}(\boldsymbol{u}) \tag{A.41}$$

$$i\mathfrak{Im}(W_\rho(\boldsymbol{u})) = W_\rho^{(1)}(\boldsymbol{u}) \tag{A.42}$$

By combining these two equations with Lemma 14, we arrive at

$$W_{\mathfrak{Re}[\rho]}(\boldsymbol{u}) = W_{\boldsymbol{u}}^{(0)}(\rho). \tag{A.43}$$

when ρ is an n-rebit state, $\mathfrak{Re}(\rho) = \rho$, which implies $W_\rho(\boldsymbol{u}) = W_\rho^{(0)}(\boldsymbol{u})$. □

A.3 Properties of the Wigner representation for channels (Proof of Lemma 16)

Lemma 16. *The Wigner representation W_Φ of a quantum channel $\Phi : \mathcal{B}(n) \to \mathcal{B}(m)$ from n to m qubits has the following properties:*

(W3) *(Input-Output Relation). If $\sigma = \Phi(\rho)$, then $W_\sigma(\boldsymbol{v}) = \sum_{\boldsymbol{u}\in\mathcal{P}_n} W_\Phi(\boldsymbol{v}|\boldsymbol{u})W_\rho(\boldsymbol{u})$.*

(W4) *(Respects the sequential composition of channels). Given a channel $\Lambda : \mathcal{B}(n') \to \mathcal{B}(n)$ from n' to n qubits, we have that $W_{\Phi\circ\Lambda} = W_\Phi W_\Lambda$.*

(W5) *(Respects the parallel composition of channels). Given a channel $\Lambda : \mathcal{B}(n') \to \mathcal{B}(m')$ from n' to m' qubits, we have that $W_{\Phi\otimes\Lambda} = W_\Phi \otimes W_\Lambda$.*

(W6) *(Preserves normalization). Each column of W_Φ sums to 1, i.e. $\sum_{\boldsymbol{v}\in\mathcal{P}_m} W_\Phi(\boldsymbol{v}|\boldsymbol{u}) = 1$.*

Proof. We prove each of the properties listed by the Lemma in turn.

Proof of Property (W3). The factorizability (Property (A1)) of phase-point operators implies

$$W_\Phi(\boldsymbol{v}|\boldsymbol{u}) = \frac{2^n}{2^m}\left\langle A_{\boldsymbol{u}} \otimes A_{\boldsymbol{v}}, \mathcal{J}(\Phi)\right\rangle. \tag{A.44}$$

Using the identity $\Phi(X) = 2^n \operatorname{Tr}_{1,\ldots,n}\left[(X^T \otimes \mathbb{1}^{\otimes m})\mathcal{J}(\Phi)\right]$ for transposition taken with respect to the computational basis, and recalling that $A_{\boldsymbol{u}}$ is real in the computational basis,

$$W_\Phi(\boldsymbol{v}|\boldsymbol{u}) = \frac{1}{2^m}\left\langle A_{\boldsymbol{v}}, \Phi(A^*_{\boldsymbol{u}})\right\rangle = \frac{1}{2^m}\left\langle A_{\boldsymbol{v}}, \Phi(A_{\boldsymbol{u}})\right\rangle. \tag{A.45}$$

Therefore, if $\sigma = \Phi(\rho)$, we obtain

$$\begin{aligned} W_\sigma(\boldsymbol{v}) &= \frac{1}{2^m}\left\langle A_{\boldsymbol{v}}, \Phi(\rho)\right\rangle \\ &= \frac{1}{2^m}\operatorname{Tr}\left[\Phi\left(\sum_{\boldsymbol{u}\in\mathcal{P}_n} W_\rho(\boldsymbol{u})A_{\boldsymbol{u}}\right)A^\dagger_{\boldsymbol{v}}\right] \\ &= \frac{1}{2^m}\sum_{\boldsymbol{u}\in\mathcal{P}_n}\left\langle A_{\boldsymbol{v}}, \Phi\left(A_{\boldsymbol{u}}\right)\right\rangle W_\rho(\boldsymbol{u}) \\ &= \sum_{\boldsymbol{u}\in\mathcal{P}_n} W_\Phi(\boldsymbol{v}|\boldsymbol{u})W_\rho(\boldsymbol{u}). \end{aligned} \tag{A.46}$$

We thereby see that if Φ maps ρ to σ, then W_Φ is a matrix that maps W_ρ to W_σ, which justifies

regarding W_Φ as the representation of Φ on phase space.

Proof of Property (W4). Since $\{A_{\boldsymbol{x}}\}_{\boldsymbol{x}\in\mathcal{P}_n}$ constitute a complex orthogonal basis for $\mathcal{B}(n)$ under the Hilbert-Schmidt inner product, wherein each $A_{\boldsymbol{x}}$ has Hilbert-Schmidt or trace norm $\sqrt{2^n}$, we can decompose the action of Λ on an n'-qubit phase-point operator $A_{\boldsymbol{u}}$ uniquely as $\Phi(A_{\boldsymbol{u}}) = \frac{1}{2^n}\sum_{\boldsymbol{x}\in\mathcal{P}_n}\langle A_{\boldsymbol{x}}, \Lambda(A_{\boldsymbol{u}})\rangle A_{\boldsymbol{x}}$. We therefore obtain

$$\begin{aligned} W_{[\Phi\circ\Lambda]}(\boldsymbol{v}|\boldsymbol{u}) &= \frac{1}{2^m}\langle A_{\boldsymbol{v}}, \Phi\circ\Lambda(A_{\boldsymbol{u}})\rangle \\ &= \frac{1}{2^m}\left\langle A_{\boldsymbol{v}}, \Phi\left(\frac{1}{2^n}\sum_{\boldsymbol{x}\in\mathcal{P}_n}\langle A_{\boldsymbol{x}}, \Lambda(A_{\boldsymbol{u}})\rangle A_{\boldsymbol{x}}\right)\right\rangle \\ &= \sum_{\boldsymbol{x}\in\mathcal{P}_n}\frac{1}{2^m}\langle A_{\boldsymbol{v}}, \Phi(A_{\boldsymbol{x}})\rangle\frac{1}{2^n}\langle A_{\boldsymbol{x}}, \Lambda(A_{\boldsymbol{u}})\rangle \\ &= \sum_{\boldsymbol{x}\in\mathcal{P}_n} W_\Phi(\boldsymbol{v}|\boldsymbol{x})W_\Lambda(\boldsymbol{x}|\boldsymbol{u}), \end{aligned} \tag{A.47}$$

which is equivalent to multiplying the matrix representing Λ on phase space to that representing Φ as follows

$$W_{\Phi\circ\Lambda} = W_\Phi W_\Lambda. \tag{A.48}$$

Proof of Property (W5). Due to the factorizability (Property (A1)) of the phase-point operators,

$$\begin{aligned} W_{\Phi\otimes\Lambda}(\boldsymbol{x}\oplus\boldsymbol{y}|\boldsymbol{u}\oplus\boldsymbol{v}) &= \frac{1}{2^{(m'+m)}}\langle A_{\boldsymbol{x}}\otimes A_{\boldsymbol{y}}, \Phi\otimes\Lambda(A_{\boldsymbol{u}}\otimes A_{\boldsymbol{v}})\rangle \\ &= \frac{1}{2^m}\langle A_{\boldsymbol{x}}, \Phi(A_{\boldsymbol{u}})\rangle\frac{1}{2^{m'}}\langle A_{\boldsymbol{y}}, \Lambda(A_{\boldsymbol{v}})\rangle \\ &= W_\Phi(\boldsymbol{x}|\boldsymbol{u})W_\Lambda(\boldsymbol{y}|\boldsymbol{v}), \end{aligned} \tag{A.49}$$

which is equivalent to taking the tensor product of the representations for Λ and Φ on phase space as follows

$$W_{\Phi\otimes\Lambda} = W_\Phi \otimes W_\Lambda. \tag{A.50}$$

Proof of Property (W6). Property (A4) of the phase point operators states $\sum_{\boldsymbol{v}\in\mathcal{P}_m} A_{\boldsymbol{v}} = 2^m I$.

By applying this to the alternative formulation of W_Φ in Eq. (A.45), we see that

$$\begin{aligned}\sum_{\boldsymbol{v}\in\mathcal{P}_m} W_\Phi(\boldsymbol{v}|\boldsymbol{u}) &= \frac{1}{2^m}\sum_{\boldsymbol{v}\in\mathcal{P}_m}\langle A_{\boldsymbol{v}}, \Phi(A_{\boldsymbol{u}})\rangle \\ &= \frac{1}{2^m}\left\langle\left(\sum_{\boldsymbol{v}\in\mathcal{P}_m} A_{\boldsymbol{v}}\right), \Phi(A_{\boldsymbol{u}})\right\rangle \\ &= \frac{1}{2^m}\langle 2^m I\Phi(A_{\boldsymbol{u}})\rangle \\ &= \mathrm{Tr}[\Phi(A_{\boldsymbol{u}})]. \end{aligned} \tag{A.51}$$

Since $\mathrm{Tr}[A_{\boldsymbol{u}}] = 1$ by Property **(A3)** of the phase-point operators, and Φ is trace-preserving,

$$\sum_{\boldsymbol{v}\in\mathcal{P}_m} W_\Phi(\boldsymbol{v}|\boldsymbol{u}) = \mathrm{Tr}[A_{\boldsymbol{u}}] = 1. \tag{A.52}$$

which means every column of W_Φ sums up to 1. □

A.4 Completely CSS-preserving operations

A.4.1 Completely CSS-preserving unitaries

The group of CSS-preserving unitaries on n qubits [73] can be generated as

$$\mathsf{G}_+(n) \coloneqq \left\langle H^{\otimes n}, \mathrm{CNOT}(i,j), X_i, Z_i\right\rangle_{i,j=1,\ldots,n,i\neq j}. \tag{A.53}$$

wherein the collective Hadamard gate $H^{\otimes n}$ has the following conjugation relations with the other generators

$$H^{\otimes n}\,\mathrm{CNOT}(i,j) = \mathrm{CNOT}(j,i)H^{\otimes n} \tag{A.54}$$

$$H^{\otimes n}X(\boldsymbol{a}) = Z(\boldsymbol{a})H^{\otimes n}, \tag{A.55}$$

We begin by establishing the subset of $\mathsf{G}_+(n)$ that is *completely* CSS-preserving.

Lemma 67. *The group of completely CSS-preserving unitaries on n qubits is*

$$\mathsf{G}(n) \coloneqq \langle \mathrm{CNOT}(i,j), Z_i, X_i\rangle_{i,j=1,\ldots,n,i\neq j}. \tag{A.56}$$

Proof. Let U_+ be any CSS-preserving unitary on n qubits. The conjugation relations given by

Eq. (A.54) and Eq. (A.55), alongside the fact that $H^{\otimes n}$ is self-inverse, imply either $U_+ \in \mathsf{G}(n)$ or $U_+ = H^{\otimes n}U$ for some $U \in \mathsf{G}(n)$.

If $U_+ \in \mathsf{G}(n)$, then because $\mathsf{G}(n)$ is a subset of $\mathsf{G}_+(n')$ for all $n' \geq n$, U_+ must be *completely* CSS-preserving. If $U_+ \notin \mathsf{G}(n)$, then $U_+ = H^{\otimes n}U$ for some $U \in \mathsf{G}(n)$, which implies U_+ is not completely CSS-preserving as $H^{\otimes n}$ is not. Therefore, U_+ is completely CSS-preserving if and only if $U_+ \in \mathsf{G}(n)$. □

A.4.2 Completely CSS-preserving measurements

Throughout the rest of this Appendix, we extend, wherever necessary, the notion of being completely CSS-preserving to *trace-decreasing* operations – i.e. a trace-decreasing operation Φ from n to n' qubits is completely CSS-preserving if, given any CSS state ρ on $(m+n)$ qubits, $(\mathsf{id} \otimes \Phi)(\rho)$ is always a (possibly subnormalised) CSS state on $(m+n')$ qubits.

The projective measurement of any n-qubit Pauli observable S is carried out using projectors

$$P(\pm S) \coloneqq \frac{1}{2}(I \pm S) \tag{A.57}$$

corresponding to the ± 1 outcomes [1]. Post-selection for the ± 1 outcome is then carried out by the operation $\Pi(\pm S) \coloneqq P(\pm S)(\cdot)P(\pm S)$.

Lemma 68. *Post-selecting on any outcome in the projective measurement of a CSS observable is completely CSS-preserving.*

Proof. Let S be a CSS observable on n qubits and $|\psi\rangle$ be a CSS state on $m+n$ qubits for any $m \geq 0$. Let the stabilizer group for $|\psi\rangle$ be generated by a set of $m+n$ independent, commuting and non-trivial CSS observables as $\mathsf{S}(|\psi\rangle) = \langle S_1, \ldots, S_{m+n}\rangle$.

Post-selecting the ± 1 outcome in a projective measurement of S on the last n qubits of $|\psi\rangle$ yields the (possibly subnormalised) output state

$$|\phi_\pm\rangle \coloneqq \left[\mathbb{1}^{\otimes m} \otimes P(\pm S)\right]|\psi\rangle = \left[\frac{1}{2}\left(\mathbb{1}^{\otimes m+n} \pm \mathbb{1}^{\otimes m} \otimes S\right)\right]|\psi\rangle = P(\pm S')\,|\psi\rangle\,, \tag{A.58}$$

where we have defined the CSS observable $S' \coloneqq \mathbb{1}^{\otimes m} \otimes S$. There are now two possibilities:

(a) that S' commutes with every generator of $\mathsf{S}(|\psi\rangle)$, so S' or $-S'$ must stabilize $|\psi\rangle$. Therefore, either $|\phi_+\rangle = |\psi\rangle$ and $|\phi_-\rangle = 0$ or vice versa, so $|\phi_\pm\rangle$ are CSS states.

(b) that, without loss of generality, S' anti-commutes with just one CSS observable S_1 that generates $\mathsf{S}(|\psi\rangle)$. This follows from the fact that, in any set of $m+n$ CSS observables that generate $\mathsf{S}(|\psi\rangle)$, those that do *not* commute with S' must *all* be X or Z-type, so by picking one such generator and multiplying all others by it, we obtain another set of $m+n$ CSS observables generating $\mathsf{S}(|\psi\rangle)$ in which only one generator does not commute with S'. Then the states $|\phi_\pm\rangle$ have norm $\frac{1}{\sqrt{2}}$ and are stabilized respectively by $\langle \pm S, S_2, \ldots, S_{m+n}\rangle$ [1], so $|\phi_\pm\rangle$ are subnormalised CSS states.

Therefore, given any pure CSS state $|\psi\rangle$ on $m+n$ qubits, post-selecting the ± 1 outcome in the projective measurement of a CSS observable on the last n qubits of $|\psi\rangle$ always produces a (possibly subnormalised) CSS state. As all CSS states are mixtures of pure CSS states by construction, we can extend this conclusion to all CSS states. □

A.4.3 CSS circuits

In this section, we show that the subset of stabilizer operations covered by Lemma 19, which we referred to as *CSS circuits*, are completely CSS-preserving.

We recall that a CSS circuit is a statistical mixture of sequences of the following four *primitive CSS channels*

1. Introducing a CSS state on some qubits,
2. Performing a completely CSS-preserving unitary,
3. Projectively measuring a CSS observable (with the possibility of subsequently performing different sequences of CSS channels depending on outcome),
4. Discarding some qubits.

Each such sequence can be executed as a *binary tree* where the root node represents inputting qubits, the leaf nodes represent outputting qubits, and internal nodes represent primitive CSS operations. An illustrative example is provided in Fig. A.1, from which we see that a sequence of primitive CSS channels can produce outputs distinguished by the sequences of measurement outcomes leading up to them.

Different numbers of ancillary qubits may be introduced on different branches of a tree representing a sequence of primitive CSS channels. However, one can arbitrarily increase the

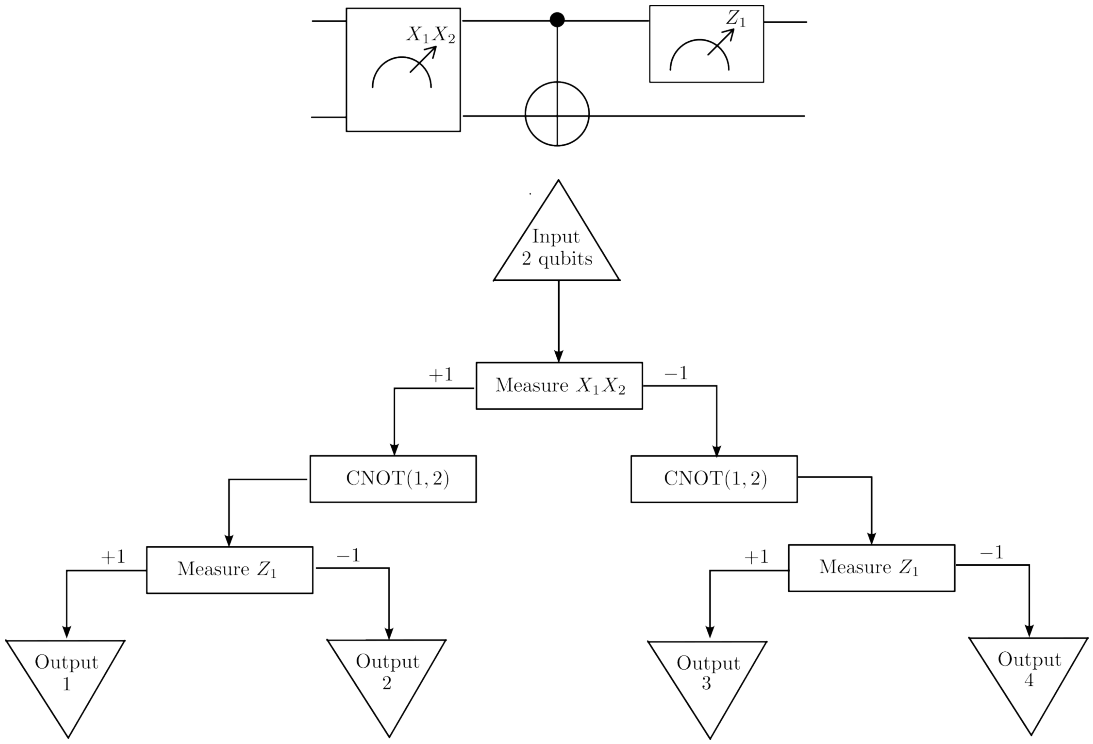


Figure A.1: The sequence of primitive CSS channels on the top can be executed as the binary tree on the bottom.

number of qubits introduced on any branch, without affecting what it does, by introducing the maximally mixed state on qubits that are immediately discarded just before the branch's output. Furthermore, different branches may have different lengths. However, one can arbitrarily lengthen any branch, without affecting what it does, by inserting identity channels just before the branch's output. Since introducing the maximally mixed state and the identity channel are both primitive CSS channels, we can, without loss of generality, only consider sequences of primitive CSS channels executed as binary trees where every branch has the same length and introduces the same number of ancillary qubits.

Lemma 69. *A CSS circuit Φ on n qubits is a statistical mixture of channels Φ_i representing sequences of primitive CSS channels, where the channel Φ_i representing the ith sequence in the mixture is a sum of (possibly trace-decreasing) operations $\Phi_{i,j}$ generating the jth distinguishable output of sequence Φ_i. Thus one can write*

$$\Phi(\rho) = \sum_i p_i \Phi_i, \text{ where } \Phi_i = \sum_j \Phi_{i,j} \text{ for } \Phi_{i,j} \coloneqq \mathrm{Tr}_{\mathcal{R}} \left[K_{i,j} (\rho \otimes \sigma_{i,j}) K_{i,j}^\dagger \right],$$
$$K_{i,j} \coloneqq \prod_{l=1}^{N} P(S_{(i,j),l}) U_{(i,j),l}, \tag{A.59}$$

where $\{p_i\}$ is a probability distribution, $\sigma_{i,j}$ is a CSS state on m ancillary qubits, $\mathcal{R}$ is a subset of the $(n+m)$ input and ancillary qubits, $U_{(i,j),l}$ is a completely CSS-preserving unitary and $P(S_{(i,j),l})$ projects onto the $+1$ eigenspace of the CSS observable $S_{(i,j),l}$. Moreover, $P(S_{(i,j),1}), \ldots, P(S_{(i,j),N})$ is the sequence of measurement outcomes that operationally distinguish the jth output of sequence Φ_i.

Proof. Without loss of generality, every branch of every sequence forming the mixture of Φ introduces the *same* number of ancillary qubits m and has the *same* length N. One can show that the jth branch of the ith sequence must generate the channel $\Phi_{i,j}$ by induction over the steps of the branch. □

We are now in a position to prove Lemma 19, which is reproduced below. To this end, it is convenient to define the unique 0-qubit state 1 as CSS.

Lemma 3. *Any CSS circuit is completely CSS-preserving.*

Proof. The decomposition of CSS circuits given in Lemma 69 implies that if performing completely CSS-preserving unitaries, (ii) conditioning on the $+1$ outcome in the projective measurement of a CSS observable, (iii) introducing a CSS state and (iv) discarding any number qubits are completely CSS-preserving, then all CSS circuits are completely CSS-preserving.

Now (i) is completely CSS-preserving by definition, we proved that (ii) is completely CSS-preserving in Lemma 68, and since the tensor product of two CSS states is always a CSS state, (iii) is completely CSS preserving.

Therefore, to prove that all CSS circuits are completely CSS-preserving, we just have to prove that discarding any number of qubits is completely CSS-preserving.

Consider discarding l qubits from n, where $n \geq l \geq 1$. As swapping the ith and jth qubit can be carried out by the completely CSS-preserving unitary $\mathrm{CNOT}(i,j)\,\mathrm{CNOT}(j,i)\,\mathrm{CNOT}(i,j)$, we need only consider discarding the *last* l qubits. Let $|\psi\rangle$ be a pure CSS state on $m+n$ qubits for any $m \geq 0$. Discarding the last l qubits of $|\psi\rangle$ then produces the state

$$\sigma := \mathsf{id} \otimes \mathrm{Tr}_{n-l+1,\ldots,n}[|\psi\rangle\langle\psi|] = \mathrm{Tr}_{m+n-l+1,\ldots,m+n}[|\psi\rangle\langle\psi|]. \tag{A.60}$$

As tracing out is unaffected by first performing a computational basis measurement on the last

l qubits, we have that

$$\sigma = \sum_{\boldsymbol{k}\in\{0,1\}^l} \mathrm{Tr}_{m+n-l+1,\ldots,m+n} \left[\mathbb{1}^{\otimes(m+n-l)} \otimes |\boldsymbol{k}\rangle\langle\boldsymbol{k}| \, (|\psi\rangle\langle\psi|) \mathbb{1}^{\otimes(m+n-l)} \otimes |\boldsymbol{k}\rangle\langle\boldsymbol{k}| \right]. \tag{A.61}$$

Since a computational basis measurement on the last l qubits means projectively measuring the CSS observable Z on each of those qubits, the trace-decreasing channel leading to any outcome of this computational basis measurement is completely CSS-preserving. Therefore, $\left(\mathbb{1}^{\otimes m+n-l} \otimes |\boldsymbol{k}\rangle\langle\boldsymbol{k}|\right) |\psi\rangle$ is a (possibly subnormalised) pure CSS state $\sqrt{p_{\boldsymbol{k}}} \, |\phi_{\boldsymbol{k}}\rangle \otimes |\boldsymbol{k}\rangle$, where $p_{\boldsymbol{k}}$ is the probability of getting the $|\boldsymbol{k}\rangle$ outcome in the computational basis measurement, and $|\phi_{\boldsymbol{k}}\rangle$ must be a (normalised) CSS state to keep the full state CSS. We thus obtain

$$\sigma = \sum_{\boldsymbol{k}\in\{0,1\}^l} \mathrm{Tr}_{m+n-l+1,\ldots,m+n} \left[p_{\boldsymbol{k}} \, |\phi_{\boldsymbol{k}}\rangle\langle\phi_{\boldsymbol{k}}| \otimes |\boldsymbol{k}\rangle\langle\boldsymbol{k}| \right] = \sum_{\boldsymbol{k}\in\{0,1\}^l} p_{\boldsymbol{k}} \, |\phi_{\boldsymbol{k}}\rangle\langle\phi_{\boldsymbol{k}}| \, , \tag{A.62}$$

which is a CSS state on $(m+n-l)$ qubits. Therefore, given any pure CSS state on $m+n$ qubits, discarding any $l \leq n$ of its final n qubits always produces a CSS state. As all CSS states are mixtures of pure CSS states by construction, we can extend this conclusion to all CSS states. □

A.4.3.1 Omission of the collective Hadamard gate

One can reasonably ask why we have excluded the collective Hadamard gate from the construction of CSS circuits. Our justification is that one can conjugate the collective Hadamard gate past any primitive CSS channel and leave another primitive CSS channel behind. This follows from the conjugation relations given by Eq. (A.54) and Eq. (A.55) for completely CSS-preserving unitaries and projective measurements of CSS observables, from the cyclic property of the trace for discarding qubits, and from

$$H^{\otimes n} \otimes \mathbb{1}^{\otimes m}(\rho \otimes \sigma) H^{\otimes n} \otimes \mathbb{1}^{\otimes m} = H^{\otimes(n+m)}(\rho \otimes (H^{\otimes m} \sigma H^{\otimes m})) H^{\otimes(n+m)} \tag{A.63}$$

for introducing a CSS state on m ancillary qubits to an n-qubit system, where we note that $H^{\otimes m}\sigma H^{\otimes m}$ is also a CSS state because $H^{\otimes m}$ is CSS-preserving on m qubits. Therefore, circuits from this wider subset are operationally equivalent to CSS circuits followed by the collective Hadamard gate conditioned upon obtaining certain outputs, and are therefore not more powerful as magic distillation protocols.

A.5 Sketch of qubit QCSI based on CSS circuits

For completeness, we sketch one such QCSI scheme presented in [73], which can be carried out using only rebits. This scheme is based an encoding of qubit states as rebit states first introduced in Ref. [142]. Explicitly, starting from a general n-qubit state

$$\rho = \sum_{\boldsymbol{v}\in\mathbb{Z}_n^2} r_{\boldsymbol{v}} e^{i\theta_{\boldsymbol{v}}} |\boldsymbol{v}\rangle , \tag{A.64}$$

the corresponding encoded state is

$$\overline{\rho} = \sum_{\boldsymbol{v}\in\mathbb{Z}_n^2} r_{\boldsymbol{v}} \cos(\theta_{\boldsymbol{v}}) |\boldsymbol{v}\rangle \otimes |R\rangle + r_{\boldsymbol{v}} \sin(\theta_{\boldsymbol{v}}) |\boldsymbol{v}\rangle \otimes |I\rangle . \tag{A.65}$$

where $|R\rangle \coloneqq |0\rangle$ and $|I\rangle \coloneqq |1\rangle$ are computational states of an *additional* ancilla rebit introduced to separately track the real and imaginary components of ρ.

In this encoded rebit setting, a channel Φ between qubits should be encoded as $\overline{\Phi}$ such that, if $\Phi(\rho) = \sigma$, then $\overline{\Phi}(\overline{\rho}) = \overline{\sigma}$ [142]. Because CSS circuits do not mix the real and imaginary parts of a state, we do not need to change them for the encoded setting [143], i.e. $\bar{\Phi} = \Phi$ for any CSS circuit Φ. Thus to achieve universal quantum computing, all we need to do is find rebit magic states that can promote completely CSS-preserving gates to a universal gateset encoded for this rebit setting, which can be achieved in two steps. First, injecting two-rebit magic state

$$|B\rangle \coloneqq \frac{1}{\sqrt{2}}(|0\rangle |+\rangle + |1\rangle |-\rangle), \tag{A.66}$$

allows us to perform encoded Hadamard gate as shown in Fig. A.2(a), as well as the encoded S-gate on the ith qubit as $\overline{S} = \mathrm{CZ}_{i,I/R}\,\mathrm{CNOT}_{i,I/R}$ [143], wherein the controlled-Z can be carried out as shown in Fig. A.2(b).

The magic state $|B\rangle$ thus enables us to promote CSS circuits to stabilizer circuits, and can be distilled using CSS circuits alone through a simple variation on Bell-state distillation protocols [143]. We can then reach universality by injecting $|H\rangle$, which can be distilled by CSS circuits such as CSS code projection protocols based on the 7-to-1 Steane or 24-to-1 Golay codes [20], as shown in Chapter 4.

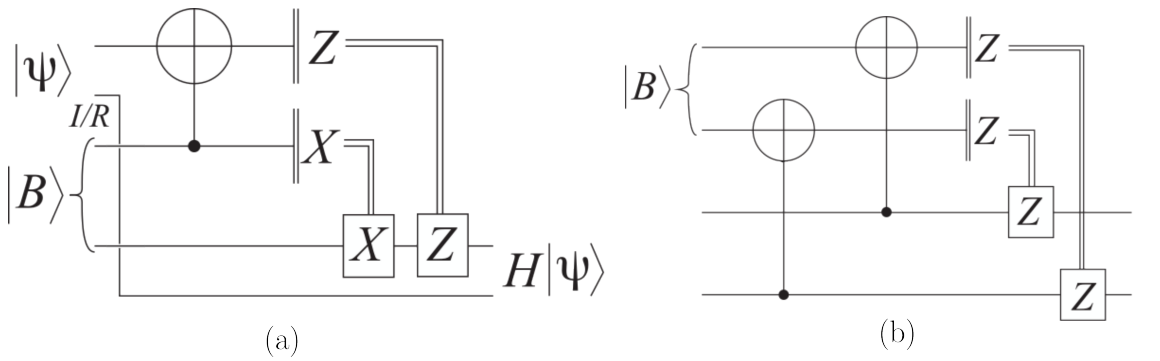


Figure A.2: **Promoting CSS to stabilizer computing in the encoded rebit setting through injection of the two-rebit magic state $|B\rangle$.** (a) the encoded Hadamard gate (b) the controlled-Z gate required for performing an encoded S-gate. [143].

A.6 Relative majorization of quasiprobability distributions (Proof of Theorem 25)

In this section, we prove Theorem 25, which we reproduce below for clarity.

Theorem 25. *Let $\boldsymbol{w} = (w_1, \ldots, w_N)$ and $\boldsymbol{w}' = (w'_1, \ldots, w'_{N'})$ be any two quasiprobability distributions and let $\boldsymbol{r} = (r_1, \ldots, r_N)$ and $\boldsymbol{r}' = (r'_1, \ldots, r'_{N'})$ be any two probability distributions. If $(\boldsymbol{w}, \boldsymbol{r}) \succ (\boldsymbol{w}', \boldsymbol{r}')$ such that $\mathrm{supp}(\boldsymbol{w}) \subseteqq \mathrm{supp}(\boldsymbol{r})$ and therefore $\mathrm{supp}(\boldsymbol{w}') \subseteqq \mathrm{supp}(\boldsymbol{r}')$, then*

$$D_\alpha(\boldsymbol{w}||\boldsymbol{r}) \geq D_\alpha(\boldsymbol{w}'||\boldsymbol{r}'), \tag{3.36}$$

or all $\alpha \in \mathcal{A} := \left\{ \frac{2a}{2b-1} \ : \ a, b \in \mathbb{N}^+, a \geq b \right\}$.

We begin by proving the following two lemmas.

Lemma 70. *Let $\boldsymbol{w} = (w_1, \ldots, w_N)$ and $\boldsymbol{w}' = (w'_1, \ldots, w'_{N'})$ be two quasiprobability distributions and let $\boldsymbol{r} = (r_1, \ldots, r_N)$ and $\boldsymbol{r}' = (r'_1, \ldots, r'_{N'})$ be two probability distributions. If $(\boldsymbol{w}, \boldsymbol{r}) \succ (\boldsymbol{w}', \boldsymbol{r}')$ and $\mathrm{supp}(\boldsymbol{w}) \subseteqq \mathrm{supp}(\boldsymbol{r})$, then $\mathrm{supp}(\boldsymbol{w}') \subseteqq \mathrm{supp}(\boldsymbol{r}')$.*

Proof. By the definition of relative majorization, there exists a stochastic map Λ such that $\Lambda \boldsymbol{w} = \boldsymbol{w}'$ and $\Lambda \boldsymbol{w}' = \boldsymbol{r}'$. If

$$r'_i = \sum_j \Lambda_{i,j} r_j = \sum_{j \in \mathrm{supp}(\boldsymbol{r})} \Lambda_{i,j} r_j = 0, \tag{A.67}$$

then as $r_j > 0$ for all $j \in \mathrm{supp}(\boldsymbol{r})$, we must have $\Lambda_{i,j} = 0$ for all $j \in \mathrm{supp}(\boldsymbol{r})$.

Since $\text{supp}(\boldsymbol{w}) \subseteqq \text{supp}(\boldsymbol{r})$, this implies $\Lambda_{i,j} = 0$ for all $j \in \text{supp}(\boldsymbol{w})$ as well, so

$$r'_i = 0 \implies w'_i = \sum_j \Lambda_{i,j} w_j = \sum_{j \in \text{supp}(\boldsymbol{w})} \Lambda_{i,j} w_j = 0. \tag{A.68}$$

We thus conclude that $\text{supp}(\boldsymbol{w}') \subseteqq \text{supp}(\boldsymbol{r}')$. □

Lemma 71. *If* $(\boldsymbol{w}, \boldsymbol{r}) \succ (\boldsymbol{w}', \boldsymbol{r}')$ *such that* $\text{supp}(\boldsymbol{w}) \subseteqq (\boldsymbol{r})$ *and therefore* $\text{supp}(\boldsymbol{w}) \subseteqq (\boldsymbol{r})$*, then* $(\tilde{\boldsymbol{w}}, \tilde{\boldsymbol{r}}) \succ (\tilde{\boldsymbol{w}}', \tilde{\boldsymbol{r}}')$*, where* $\tilde{\boldsymbol{w}}$ *and* $\tilde{\boldsymbol{r}}$ *are the parts of* $\boldsymbol{w}$ *and* $\boldsymbol{r}$ *on* $\text{supp}(\boldsymbol{r})$*, while* $\tilde{\boldsymbol{w}}'$ *and* $\tilde{\boldsymbol{r}}'$ *are the parts of* $\boldsymbol{w}$ *and* $\boldsymbol{r}$ *on* $\text{supp}(\boldsymbol{r}')$*.*

Proof. Let Ω and $\tilde{\Omega}$ denote the sample spaces of $\boldsymbol{r}$ and $\tilde{\boldsymbol{r}}$ respectively, and let $\tilde{N}$ denote the size of $\tilde{\Omega}$. We first define the stochastic map $\Lambda_+ : \tilde{\Omega} \to \Omega$ as

$$\Lambda_+ \coloneqq \begin{pmatrix} I_{\tilde{N} \times \tilde{N}} \\ \mathbf{0}_{(N-\tilde{N}),\tilde{N}} \end{pmatrix}, \tag{A.69}$$

which acts on $\tilde{\boldsymbol{r}}$ and $\tilde{\boldsymbol{w}}$ as

$$\Lambda_+ \tilde{\boldsymbol{w}} = \tilde{\boldsymbol{w}} \oplus \mathbf{0}_{N-\tilde{N}}, \tag{A.70}$$

$$\Lambda_+ \tilde{\boldsymbol{r}} = \tilde{\boldsymbol{r}} \oplus \mathbf{0}_{N-\tilde{N}}. \tag{A.71}$$

In words, Λ_+ pads $\tilde{\boldsymbol{w}}$ and $\tilde{\boldsymbol{r}}$ with 0s until they become vectors with N entries.

We next define the stochastic map $\Pi_{\boldsymbol{r}} : \Omega \to \Omega$ as a permutation of Ω that moves entries outside of $\text{supp}(\boldsymbol{r})$ to the end of the vector, while maintaining the ordering amongst entries within $\text{supp}(\boldsymbol{r})$, so we can write

$$\Pi_{\boldsymbol{r}} \boldsymbol{r} = \tilde{\boldsymbol{r}} \oplus \mathbf{0}_{N-\tilde{N}} \tag{A.72}$$

$$\Pi_{\boldsymbol{r}} \boldsymbol{w} = \tilde{\boldsymbol{w}} \oplus \mathbf{0}_{N-\tilde{N}} \tag{A.73}$$

as any entry outside $\text{supp}(\boldsymbol{r})$ is also outside $\text{supp}(\boldsymbol{w})$. Similarly, letting Ω' be the sample space of $\boldsymbol{r}'$, we define another permutation of Ω', $\Pi_{\boldsymbol{r}'} : \Omega' \to \Omega'$, that moves entries outside of $\text{supp}(\boldsymbol{r}')$ to the end of the vector, while maintaining the order amongst entries within $\text{supp}(\boldsymbol{r}')$, so we

can write

$$\Pi_{\boldsymbol{r}'}\boldsymbol{r}' = \tilde{\boldsymbol{r}}' \oplus \mathbf{0}_{N'-\tilde{N}'} \tag{A.74}$$

$$\Pi_{\boldsymbol{r}'}\boldsymbol{w}' = \tilde{\boldsymbol{w}}' \oplus \mathbf{0}_{N'-\tilde{N}'}. \tag{A.75}$$

As any entry outside $\mathrm{supp}(\boldsymbol{r}')$ is also outside $\mathrm{supp}(\boldsymbol{w}')$. Because they are permutations, $\Pi_{\boldsymbol{r}}$ and $\Pi_{\boldsymbol{r}'}$ have inverses that are also stochastic maps.

Finally, letting Ω' and $\tilde{\Omega}'$ denote the sample spaces of $\boldsymbol{r}'$ and $\tilde{\boldsymbol{r}}'$ respectively, we define a stochastic map $\Lambda_- : \Omega' \to \tilde{\Omega}'$ as

$$\Lambda_- \coloneqq \left(I_{\tilde{N}'\times\tilde{N}'} \middle| \Lambda'_{\tilde{N}',N'-\tilde{N}'} \right), \tag{A.76}$$

where $\Lambda'_{\tilde{N}',N'-\tilde{N}'}$ is an arbitrary stochastic matrix. Then Λ_- acts as

$$\Lambda_-(\tilde{\boldsymbol{r}}' \oplus \mathbf{0}_{N'-\tilde{N}'}) = \tilde{\boldsymbol{r}}', \tag{A.77}$$

$$\Lambda_-(\tilde{\boldsymbol{r}}' \oplus \mathbf{0}_{N'-\tilde{N}'}) = \tilde{\boldsymbol{w}}'. \tag{A.78}$$

i.e. removing the 0s at the end of $\tilde{\boldsymbol{r}}' \oplus \mathbf{0}_{N'-\tilde{N}'}$ and $\tilde{\boldsymbol{w}}' \oplus \mathbf{0}_{N'-\tilde{N}'}$.

By the definition of relative majorization, there exists a stochastic map Λ such that $\Lambda\boldsymbol{w} = \boldsymbol{w}'$ and $\Lambda\boldsymbol{w}' = \boldsymbol{r}'$. Since a sequence of stochastic maps is still stochastic, we can define a stochastic map $A : \tilde{\Omega} \to \tilde{\Omega}'$

$$A \coloneqq \Lambda_-\Pi_{\boldsymbol{r}'}\Lambda\Pi_{\boldsymbol{r}}^{-1}\Lambda_+. \tag{A.79}$$

that acts as $A\tilde{\boldsymbol{w}} = \tilde{\boldsymbol{w}}'$ and $A\tilde{\boldsymbol{r}} = \tilde{\boldsymbol{r}}'$. □

With these lemmas in place, we can turn to the following theorem established in Ref. [27].

Theorem 72 ([27]). *Let $\boldsymbol{w} = (w_1, \ldots, w_N)$ and $\boldsymbol{w}' = (w'_1, \ldots, w'_{N'})$ be two quasiprobability distributions and let $\boldsymbol{r} = (r_1, \ldots, r_N)$ and $\boldsymbol{r}' = (r'_1, \ldots, r'_{N'})$ be two strictly positive probability distributions. If $(\boldsymbol{w}, \boldsymbol{r}) \succ (\boldsymbol{w}', \boldsymbol{r}')$ then*

$$\forall \alpha \in \mathcal{A} : D_\alpha(\boldsymbol{w}||\boldsymbol{r}) \geq D_\alpha(\boldsymbol{w}'||\boldsymbol{r}'), \tag{A.80}$$

for all $\alpha \in \mathcal{A} \coloneqq \left\{ \frac{2a}{2b-1} \ : \ a, b \in \mathbb{N}^+, a \geq b \right\}$.

Putting Lemma 71 together with Theorem 72, we conclude that if $(\boldsymbol{w}, \boldsymbol{r}) \succ (\boldsymbol{w}', \boldsymbol{r}')$ such that $\mathrm{supp}(\boldsymbol{w}) \subseteqq \mathrm{supp}(\boldsymbol{r})$ and therefore $\mathrm{supp}(\boldsymbol{w}') \subseteqq \mathrm{supp}(\boldsymbol{r}')$, then

$$D_\alpha(\tilde{\boldsymbol{w}}||\tilde{\boldsymbol{r}}) \geq D_\alpha(\tilde{\boldsymbol{w}}'||\tilde{\boldsymbol{r}}'), \tag{A.81}$$

for all $\alpha \in \mathcal{A}$. Since $\boldsymbol{w}$ ($\boldsymbol{w}'$) is 0 whenever $\boldsymbol{r}$ ($\boldsymbol{r}'$) is 0, we find that

$$D_\alpha(\tilde{\boldsymbol{w}}||\tilde{\boldsymbol{r}}) \coloneqq \frac{1}{\alpha - 1} \log \sum_{j \in \mathrm{supp}(\boldsymbol{r})} r_j \left(\frac{w_j}{r_j}\right)^\alpha = \frac{1}{\alpha - 1} \log \sum_j r_j \left(\frac{w_j}{r_j}\right)^\alpha \coloneqq D_\alpha(\boldsymbol{w}||\boldsymbol{r}). \tag{A.82}$$

and similarly $D_\alpha(\tilde{\boldsymbol{w}}'||\tilde{\boldsymbol{r}}') = D_\alpha(\boldsymbol{w}||\boldsymbol{r})$. We thus arrive at the extension of Theorem 72 presented in Theorem 25.

A.6.1 Entropic constraints on completely CSS-preserving protocols (proofs of Lemma 27 and Theorem 30)

Lemma 27. *Let W be a generalised Gross's Wigner representation for d-dimensional qudits with respect to the QCSI scheme $(\mathcal{R}, \mathcal{F}, \mathcal{M})$. Consider states $\rho \in \mathcal{F} \cup \mathcal{M}$, $\tau \in \mathcal{F}$ such that $\mathrm{supp}(W_\rho) \subseteqq \mathrm{supp}(W_\tau)$, and a stochastically-represented channel $\Phi \in \mathcal{R}$. Then the α-Rényi divergence $D_\alpha(\cdot||\cdot)$ is well-defined and satisfies the following properties for $\alpha \in \mathcal{A}$:*

(D1) *Non-negativity: $D_\alpha(W_\rho||W_\tau) \geq 0$.*

(D2) *Additivity: $D_\alpha(W_{\rho^{\otimes n}}||W_{\tau^{\otimes n}}) = nD_\alpha(W_\rho||W_\tau)$ for all $n \in \mathbb{N}^+$.*

(D3) *Data-Processing Inequality: $D_\alpha(W_\rho||W_\tau) \geq D_\alpha(W_{\Phi(\rho)}||W_{\Phi(\tau)})$, where by Lemma 70 we also have that $\mathrm{supp}(W_{\Phi(\rho)}) \subseteqq \mathrm{supp}(W_{\Phi(\tau)})$.*

Proof. By definition, W_ρ and W_τ are respectively a quasiprobability distribution and a probability distribution such that $\mathrm{supp}(W_\rho) \in \mathrm{supp}(W_\tau)$, and we have already seen that $D_\alpha(W_\rho||W_\tau)$ will remain well-defined from $\alpha \in \mathcal{A}$. We proceed to prove each of the properties listed in turn.

(D1) Let $\mathcal{P}$ be the phase space on which ρ and τ are represented. Consider the stochastic map $\Lambda : \mathcal{P} \to \mathcal{P}$ defined by $\Lambda(\boldsymbol{v}|\boldsymbol{u}) = \frac{1}{|\mathcal{P}|}$ for all $\boldsymbol{v}, \boldsymbol{u} \in \mathcal{P}$. Since W_ρ and W_τ are normalized, we see that $\Lambda(W_\rho) = \Lambda(W_\tau) = \frac{\mathbf{1}}{|\mathcal{P}|}$, where $\mathbf{1}$ is the unit vector. Therefore,

$(W_\rho, W_\tau) \succ \left(\frac{\mathbf{1}}{|\mathcal{P}|}, \frac{\mathbf{1}}{|\mathcal{P}|}\right)$, so by Theorem 25,

$$D_\alpha(W_\rho||W_\tau) \geq D_\alpha\left(\frac{\mathbf{1}}{|\mathcal{P}|}\middle|\middle|\frac{\mathbf{1}}{|\mathcal{P}|}\right) = \frac{1}{\alpha-1}\log\sum_{\boldsymbol{u}\in\mathcal{P}}\frac{1}{|\mathcal{P}|} = 0. \tag{A.83}$$

(D2) Let $\mathcal{P}_n$ be the phase-space of an n-qudit system. Since W_ρ respects the parallel composition of processes, we have that $W_{\rho^{\otimes n}} = (W_\rho)^{\otimes n}$ and $W_{\tau^{\otimes n}} = (W_\tau)^{\otimes n}$, which implies

$$\begin{aligned}
D_\alpha(W_{\rho^{\otimes n}}||W_{\tau^{\otimes n}}) &= \frac{1}{\alpha-1}\log\sum_{\boldsymbol{u}\in\mathcal{P}_n} W_{\rho^{\otimes n}}(\boldsymbol{u})^\alpha W_{\tau^{\otimes n}}(\boldsymbol{u})^{1-\alpha} \\
&= \frac{1}{\alpha-1}\log\sum_{\boldsymbol{u}_1\in\mathcal{P}_1}\cdots\sum_{\boldsymbol{u}_n\in\mathcal{P}_1}\left[\prod_{i=1}^n W_\rho(\boldsymbol{u}_i)^\alpha W_\tau(\boldsymbol{u}_i)^{1-\alpha}\right] \\
&= \frac{1}{\alpha-1}\log\prod_{i=1}^n\left[\sum_{\boldsymbol{u}_i\in\mathcal{P}_1} W_\rho(\boldsymbol{u}_i)^\alpha W_\tau(\boldsymbol{u}_i)^{1-\alpha}\right] \\
&= n\left(\frac{1}{\alpha-1}\log\sum_{\boldsymbol{u}_i\in\mathcal{P}_1} W_\rho(\boldsymbol{u}_i)^\alpha W_\tau(\boldsymbol{u}_i)^{1-\alpha}\right) \\
&= nD_\alpha(W_\rho||W_\tau)
\end{aligned} \tag{A.84}$$

(D3) Since W respects the sequential composition of processes, we have that $W_{\Phi(\rho)} = W_\Phi W_\rho$ and $W_{\Phi(\tau)} = W_\Phi W_\tau$. As W_Φ is stochastic, this implies $(W_\rho, W_\tau) \succ (W_{\Phi(\rho)}, W_{\Phi(\tau)})$, so we can prove this property by invoking Theorem 25. □

Theorem 30. *Let ρ be a noisy multirebit magic state and τ be a CSS state where we have that* $\mathrm{supp}(W_\rho) \subseteqq \mathrm{supp}(W_\tau)$. *If there exists a completely CSS-preserving magic distillation protocol Φ such that $\Phi(\rho^{\otimes n}) = \rho'$ and $\tau' := \Phi(\tau^{\otimes n})$, then*

$$\Delta D_\alpha := nD_\alpha(W_\rho||W_\tau) - D_\alpha(W_{\rho'}||W_{\tau'}) \geq 0 \tag{3.41}$$

for the qubit Wigner representation of Eq. (3.18) and all $\alpha \in \mathcal{A} := \{\frac{2a}{2b-1}, |a, b \in \mathbb{N}^+, a \geq b\}$.

Proof. We have seen that the qubit Wigner representation defined in Eq. (3.18) provides a generalised Gross's Wigner representation for the QCSI scheme $(\mathcal{R}, \mathcal{F}, \mathcal{M})$ where $\mathcal{R}$ is the set of completely CSS-preserving operations, $\mathcal{F}$ is the set of CSS states and $\mathcal{M}$ is the set of non-CSS

multirebit states. We can therefore apply Lemma 27 to W, ρ and τ to obtain

$$D_\alpha(W_{\rho^{\otimes n}}||W_{\tau^{\otimes n}}) = nD_\alpha(W_\rho||W_\tau). \tag{A.85}$$

We also have that $\rho^{\otimes n}$ is also a multirebit state because ρ is a multirebit state, and $\tau^{\otimes n}$ is a CSS state because τ is a CSS state and introducing a CSS state is completely CSS-preserving (see Lemma 19). As W respects the parallel composition of processes, we see that $\mathrm{supp}(W_\rho) \subseteqq \mathrm{supp}(W_\tau)$ implies

$$\begin{aligned}\mathrm{supp}(W_{\rho^{\otimes n}}) = \mathrm{supp}(W_\rho^{\otimes n}) = \mathrm{supp}(W_\rho)^{\times n} \\ \subseteqq \mathrm{supp}(W_\tau)^{\times n} = \mathrm{supp}(W_\tau^{\otimes n}) = \mathrm{supp}(W_{\tau^{\otimes n}}).\end{aligned} \tag{A.86}$$

Finally, by Theorem 18, Φ is stochastically represented. We can therefore apply Lemma 27 to W, $\rho^{\otimes n}$, $\tau^{\otimes n}$ and Φ to obtain

$$D_\alpha(W_{\rho^{\otimes n}}||W_{\tau^{\otimes n}}) \geq D_\alpha(W_{\rho'}||W_{\tau'}). \tag{A.87}$$

Substituting the left-hand side using Eq. (A.42) and rearranging completes the proof. □

A.6.2 Generating magic monotones from entropic constraints on completely CSS-preserving protocols

Theorem 28. *Let W be a generalised Gross's Wigner representation for the QCSI scheme $(\mathcal{R}, \mathcal{F}, \mathcal{M})$. We then have that*

$$M_\alpha(\rho) \coloneqq \inf_{\substack{\tau \in \mathcal{F} \text{ s.t.} \\ \mathrm{supp}(W_\rho) \subseteqq \mathrm{supp}(W_\tau)}} D_\alpha(W_\rho||W_\tau), \tag{3.38}$$

is a magic monotone for any $\alpha \in \mathcal{A}$. Furthermore, M_α can be used to lower-bound the overhead n of a magic distillation process $\rho^{\otimes n} \to \sigma$ within this QCSI scheme as

$$\forall \alpha \in \mathcal{A} : n \geq \frac{M_\alpha(\sigma)}{M_\alpha(\rho)}. \tag{3.39}$$

Proof. Letting $\Phi \in \mathcal{R}$ denote a magic distillation protocol such that $\Phi(\rho) = \rho'$, we obtain

$$M_\alpha(\rho) = \inf_{\substack{\tau\in\mathcal{F}\ s.t.\\ \mathrm{supp}(W_\rho)\subseteqq\mathrm{supp}(W_\tau)}} D_\alpha(W_\rho||W_\tau) \geq \inf_{\substack{\tau\in\mathcal{F}\ s.t.\\ \mathrm{supp}(W_\rho)\subseteqq\mathrm{supp}(W_\tau)}} D_\alpha(W_{\rho'}||W_{\Phi(\tau)}) \geq M_\alpha(\rho'), \quad \text{(A.88)}$$

where the first inequality follows from applying the data-processing property **(D3)** of the α-Rényi entropies to each (ρ, τ), and the second inequality follows from the fact that $\Phi(\tau)$ must be a non-magic state such that, by Lemma 70, we have $\mathrm{supp}(W_{\rho'}) \subseteqq \mathrm{supp}(W_{\Phi(\tau)})$. We thus conclude that $\{M_\alpha\}$ are monotonically non-increasing under the magic distillation protocols defined as $\mathcal{R}$.

By the non-negativity property **(D1)** of the α-Rényi operators, we are guaranteed that $M_\alpha(\rho) \geq 0$ for all ρ. Furthermore, given any non-magic state τ, we have that $D_\alpha(W_\tau||W_\tau) = 0$, which further guarantees that $M_\alpha(\tau) = 0$ for any $\tau \in \mathcal{F}$. We therefore conclude that $\{M_\alpha\}_{\alpha\in\mathcal{A}}$ form a set of magic monotones in the QCSI scheme $(\mathcal{R}, \mathcal{F}, \mathcal{M})$.

Now $\{M_\alpha\}$ are sub-additive, i.e. $nM_\alpha(\rho) \geq M_\alpha(\rho^{\otimes n})$, because

$$nM_\alpha(\rho) = \inf_{\substack{\tau\in\mathcal{F}\ s.t.\\ \mathrm{supp}(W_\rho)\in\mathrm{supp}(W_\tau)}} nD_\alpha(W_\rho||W_\tau) = \inf_{\substack{\tau\in\mathcal{F}\ s.t.\\ \mathrm{supp}(W_\rho)\in\mathrm{supp}(W_\tau)}} D_\alpha(W_{\rho^{\otimes n}}||W_{\tau^{\otimes n}}) \geq M_\alpha(\rho^{\otimes n}) \quad \text{(A.89)}$$

where the second equality is due to Property **(D1)** of the α-Rényi entropies, and the final inequality is due to the fact that $\tau^{\otimes n}$ must be non-magic if τ is non-magic, and W's respect for the parallel composition of processes together with $\mathrm{supp}(W_\rho) \subseteqq \mathrm{supp}(W_\tau)$ leads to $\mathrm{supp}(W_{\rho^{\otimes n}}) = \mathrm{supp}(W_\rho^{\otimes n}) = \mathrm{supp}(W_\rho)^{\times n} \subseteqq \mathrm{supp}(W_\tau)^{\times n} = \mathrm{supp}(W_\tau^{\otimes n}) = \mathrm{supp}(W_{\tau^{\otimes n}})$.

Therefore, if there exists a magic distillation protocol $\Phi \in \mathcal{R}$ such that $\Phi(\rho^{\otimes n}) = \rho'$, then we can combine Eq. (A.88) and Eq. (A.89) to show that

$$nM_\alpha(\rho^{\otimes n}) \geq M_\alpha(\rho^{\otimes n}) \geq M_\alpha(\rho'). \quad \text{(A.90)}$$

Rearranging this inequality leads to the lower bound on n found in Eq. (3.39). □

Lemma 29. *Let W be a generalised Gross's Wigner representation for a QCSI scheme $(\mathcal{R}, \mathcal{F}, \mathcal{M})$.*

Then given any state $\rho \in \mathcal{F} \cup \mathcal{M}$ covered by the QCSI scheme, we have that

$$\lim_{\alpha \to 1^+} (\alpha - 1) M_\alpha(\rho) = \mathcal{M}_W(\rho) \coloneqq \log \sum_{\boldsymbol{u}} |W_\rho(\boldsymbol{u})| \tag{3.40}$$

where the limit is taken through a sequence of rational values drawn from $\mathcal{A}$.

Proof. We first note that the values of α in $\mathcal{A}$ were so chosen to ensure that $W_\rho^\alpha(\boldsymbol{u}) = |W_\rho(\boldsymbol{u})^\alpha|$.

$$\begin{aligned}
\lim_{\alpha \to 1^+} (\alpha - 1) M_\alpha(\rho) &= \lim_{\alpha \to 1^+} \inf_{\substack{\tau \in \mathcal{F} \ s.t. \\ \mathrm{supp}(W_\rho) \in \mathrm{supp}(W_\tau)}} (\alpha - 1) D_\alpha(W_\rho || W_\tau) \\
&= \lim_{\alpha \to 1^+} \inf_{\substack{\tau \in \mathcal{F} \ s.t. \\ \mathrm{supp}(W_\rho) \in \mathrm{supp}(W_\tau)}} \log \sum_{\boldsymbol{u}} \frac{|W_\rho(\boldsymbol{u})|^\alpha}{W_\tau(\boldsymbol{u})^{\alpha - 1}} \\
&= \inf_{\substack{\tau \in \mathcal{F} \ s.t. \\ \mathrm{supp}(W_\rho) \in \mathrm{supp}(W_\tau)}} \log \sum_{\boldsymbol{u}} |W_\rho(\boldsymbol{u})| = \log \sum_{\boldsymbol{u}} |W_\rho(\boldsymbol{u})| = \mathcal{M}_W(\rho).
\end{aligned} \tag{A.91}$$

□

A.6.3 Negative α-Rényi entropies for rebit magic states

Consider the QCSI scheme $(\mathcal{R}, \mathcal{F}, \mathcal{M})$ where $\mathcal{R}$ consists of completely CSS-preserving operations, $\mathcal{F}$ are CSS states and magic states are non-CSS rebit states. Because all pure rebit magic states have negativities in their Wigner representation, and completely CSS-preserving operations are stochastically represented, we see that only rebit magic states with negativities in their Wigner distributions can be distilled within this QCSI scheme.

We now prove that, given any rebit state with negativities in its Wigner distribution, at least one α-Rényi entropy with $\alpha \in \mathcal{A}$ evaluated on the state is negative.

Lemma 73 (based on Theorem 10 of Ref. [27])**.** *There exists $\alpha \in \mathcal{A}$ such that $H_\alpha(W_\rho) < 0$ if and only if ρ is a rebit state whose Wigner representation W_ρ has negative entries.*

Proof. Since H_α is non-negative on probability distributions, $H_\alpha(W_\rho) < 0$ implies that W_ρ must contain negative entries.

In the reverse direction, if W_ρ is not a probability distribution, then we must have

$$\lim_{\epsilon \to 0} \sum_{\boldsymbol{u} \in \mathcal{P}} |W_\rho(\boldsymbol{u})|^{1+\epsilon} = \sum_{\boldsymbol{u} \in \mathcal{P}_n} |W_\rho(\boldsymbol{u})| > 1 \tag{A.92}$$

As the function $f(x_1, \ldots, x_n) \coloneqq \sum_i x_i^{1+\epsilon}$ is continuous in ϵ, there must exist some finite interval $\epsilon \in (0, \epsilon_{th}]$ such that

$$\forall \epsilon \in (0, \epsilon_{th}] : \sum_{\boldsymbol{u} \in \mathcal{P}} |W_\rho(\boldsymbol{u})|^{1+\epsilon} > 1. \tag{A.93}$$

Since $\mathcal{A}$ is dense in $(1, \infty)$, $\mathcal{A}' \coloneqq \{\alpha - 1 | \alpha \in \mathcal{A}\}$ must be dense in $(0, \infty)$, which implies it is always possible to find $\alpha' \in \mathcal{A}$ such that $(\alpha' - 1) \in (0, \epsilon_{\text{th}}]$. We then have that

$$\forall \alpha' \in (1, 1 + \epsilon_{\text{th}}), \alpha' \in \mathcal{A} : \sum_{\boldsymbol{u} \in \mathcal{P}_n} W_\rho(\boldsymbol{u})^{\alpha'} = \sum_{\boldsymbol{u} \in \mathcal{P}} |W_\rho(\boldsymbol{u})|^{\alpha'} > 1. \tag{A.94}$$

As $\alpha' > 1$, this implies $H_{\alpha'}(W_\rho) < 0$. □

Appendix B

Appendices to Chapter 4

B.1 Stochastic representation of qubit CSS code projection protocols

In this section, we prove that all qubit CSS code projection protocols are stochastic under the Wigner representation defined in Eq. (3.18).

To this end, it is helpful to first consider the impact of a CNOT gate on the CSS observables generating the stabilizer group defining a CSS code. We first remark that a CNOT channel acts on individual X- and Z-type CSS observables as

$$\mathrm{CNOT}(i,j) Z_m \,\mathrm{CNOT}(i,j) = \begin{cases} Z_i Z_j & \text{for } m = j, \\ Z_m & \text{otherwise.} \end{cases} \tag{B.1}$$

$$\mathrm{CNOT}(i,j) X_m \,\mathrm{CNOT}(i,j) = \begin{cases} X_i X_j & \text{for } m = i, \\ X_m & \text{otherwise.} \end{cases} \tag{B.2}$$

We now introduce an isomorphism M_Z that maps any set of Z-type CSS observables $\{Z(\boldsymbol{u}_i)\}_{i=1,\ldots,m}$, where each $\boldsymbol{u}_i$ is an n-bit vector, to a matrix whose ith column is $\boldsymbol{u}_i$, i.e.

$$M_Z\left(\{Z(\boldsymbol{u}_1), \ldots, Z(\boldsymbol{u}_m)\}\right) = \begin{bmatrix} | & \ldots & | \\ \boldsymbol{u}_1 & \ldots & \boldsymbol{u}_m \\ | & \ldots & | \end{bmatrix}. \tag{B.3}$$

Performing the channel for swapping qubits i and j, $\mathcal{U}_{\mathrm{SWAP}}(i,j) \coloneqq U_{\mathrm{SWAP}}(i,j)(\cdot)U_{\mathrm{SWAP}}(i,j)$

where $U_{\mathrm{SWAP}}(i,j) \coloneqq \mathrm{CNOT}(i,j)\,\mathrm{CNOT}(j,i)\,\mathrm{CNOT}(i,j)$ [1], on every observable in the set $\{Z(\boldsymbol{u}_1),\ldots,Z(\boldsymbol{u}_m)\}$ is represented in the M_Z mapping by *swapping rows i and j*.

Furthermore, by Eq. (B.1), performing the channel $\mathcal{U}_{\mathrm{Z-SUM}}(i,j) \coloneqq \mathrm{CNOT}(j,i)(\cdot)\,\mathrm{CNOT}(j,i)$ on $Z(\boldsymbol{u})$ for any n-bit vector $\boldsymbol{u}$ produces

$$\mathcal{U}_{\mathrm{Z-SUM}}(Z(\boldsymbol{u})) = Z(\boldsymbol{u} + u_i \boldsymbol{e}_j), \tag{B.4}$$

where arithmetic is modulo 2 and $\boldsymbol{e}_j$ is an n-bit vector that has 1 in its jth entry, and 0 everywhere else. Consequently, the simultaneous action of $\mathcal{U}_{\mathrm{Z-SUM}}(i,j)$ on every observable in the set $\{Z(\boldsymbol{u}_1),\ldots,Z(\boldsymbol{u}_m)\}$ is represented under the M_Z mapping as *adding row i to row j*.

Similarly, we can introduce an isomorphism M_X that maps any set of X-type CSS observables $\{X(\boldsymbol{v}_j)\}_{j=1,\ldots,m}$, where each $\boldsymbol{v}_j$ is an n-bit vector, to a matrix whose jth column is $\boldsymbol{v}_j$, i.e.

$$M_X\left(\{X(\boldsymbol{v}_1),\ldots,X(\boldsymbol{v}_m)\}\right) = \begin{bmatrix} | & \ldots & | \\ \boldsymbol{v}_1 & \ldots & \boldsymbol{v}_m \\ | & \ldots & | \end{bmatrix}. \tag{B.5}$$

Performing the channel $\mathcal{U}_{\mathrm{SWAP}}(i,j)$ for swapping qubits i and j on every observable in the set $\{X(\boldsymbol{u}_1),\ldots,X(\boldsymbol{u}_m)\}$ is also represented in the M_X mapping by *swapping rows i and j*. Furthermore, by Eq. (B.2), performing the channel $\mathcal{U}_{\mathrm{Z-SUM}}(i,j) \coloneqq \mathrm{CNOT}(i,j)(\cdot)\,\mathrm{CNOT}(i,j)$ on $X(\boldsymbol{v})$ for any n-bit vector $\boldsymbol{v}$ produces

$$\mathcal{U}_{\mathrm{X-SUM}}(X(\boldsymbol{v})) = Z(\boldsymbol{v} + v_i \boldsymbol{e}_j), \tag{B.6}$$

where arithmetic is modulo 2, which implies that the simultaneous action of $\mathcal{U}_{\mathrm{X-SUM}}(i,j)$ on every observable in the set $\{X(\boldsymbol{u}_1),\ldots,X(\boldsymbol{u}_m)\}$ is represented under the M_X mapping as *adding row i to row j*. We remark that M_Z and M_X are the Z and X blocks forming the parity check matrix [1] of a CSS code.

With these preliminaries in place, we can prove the following Lemma[1].

[1] A similar proof can be found at in the Appendix of Ref. [144].

Lemma 74. *Let $\mathcal{C}$ be an $[[n,k]]$ stabilizer code where $n > k$, so $\mathsf{S}(\mathcal{C})$ can be generated as*

$$\mathsf{S}(\mathcal{C}) \coloneqq \langle (-1)^{b_1} Z(\boldsymbol{u}_1), \ldots, (-1)^{b_r} Z(\boldsymbol{u}_r), (-1)^{b_{r+1}} X(\boldsymbol{v}_{r+1}), \ldots, (-1)^{b_{n-k}} X(\boldsymbol{v}_{n-k}) \rangle. \tag{B.7}$$

where $b_1 \ldots, b_{n-k}$ are bits, $\{\boldsymbol{u}_1, \ldots, \boldsymbol{u}_r\}$ form a set of non-zero and linearly independent n-bit vectors, $\{\boldsymbol{v}_{r+1}, \ldots, \boldsymbol{v}_{n-k}\}$ form another set of non-zero and linearly independent n-bit vectors, and r gives the number of Z-type generators and so falls into the range $r \in \{0, \ldots, n-k\}$.

The decoding channel $\mathcal{U}_{\mathcal{C}}{}^{\dagger}$ of $\mathcal{C}$ then has the form

$$\mathcal{U}_{\mathcal{C}}^{\dagger} = \mathcal{U}_{r+1}^{H} \circ \cdots \circ \mathcal{U}_{n-k}^{H} \circ \mathcal{U}, \tag{B.8}$$

where $\mathcal{U}_l^H \coloneqq H_l(\cdot)H_l$ is the Hadamard gate on the lth qubit, and $\mathcal{U}$ is a completely CSS-preserving unitary channel such that

$$\forall i \in \{1, \ldots, r\} : \mathcal{U}[(-1)^{b_i} Z(\boldsymbol{u}_i)] = Z_{k+i}, \tag{B.9}$$

$$\forall j \in \{r+1, \ldots, n-k\} : \mathcal{U}[(-1)^{b_j} X(\boldsymbol{v}_j)] = X_{k+j}. \tag{B.10}$$

Proof. The (unique) reduced row echelon form R_Z of $M_Z(\boldsymbol{u}_1, \ldots, \boldsymbol{u}_r)$ is

$$R_Z \coloneqq \begin{bmatrix} I_{r,r} \\ 0_{n-r,r} \end{bmatrix}, \tag{B.11}$$

where $I_{r,r}$ is an $r \times r$ identity matrix while $0_{n-r,r}$ is a $(n-r) \times r$ null matrix. We can use Gauss-Jordan elimination on $\mathbb{F}_2$ to convert M_Z to R_Z. As this procedure is a sequence of row swaps and additions, both of which can be executed via sequences of CNOT channels, we conclude that there exists a unitary channel $\mathcal{U}_Z$ formed by a sequence of CNOT channels such that

$$\forall i \in \{1, \ldots, r\} : \mathcal{U}_Z[Z(\boldsymbol{u}_i)] = Z_i. \tag{B.12}$$

Let us define $X(\boldsymbol{v}'_j) \coloneqq \mathcal{U}_Z[X(\boldsymbol{v}_j)]$ for $j \in \{r+1, \ldots, n-k\}$. Since the Z-type generators of $\mathsf{S}(\mathcal{C})$ commute with the X-type generators, each $X(\boldsymbol{v}'_j)$ must commute with $Z_1, \ldots, Z_r$ for all

$j \in \{r+1, \ldots, n-k\}$, which implies

$$\forall j \in \{r+1, \ldots, n-k\} : \mathcal{U}_Z[X(\boldsymbol{v}_j)] \coloneqq X(\boldsymbol{v}'_j) = \mathbb{1}^{\otimes r} \otimes X(\boldsymbol{v}''_j) \tag{B.13}$$

where $\boldsymbol{v}''_j$ is an $(n-r)$-bit vector. Repeating our reasoning for $M_Z(\boldsymbol{u}_1, \ldots, \boldsymbol{u}_r)$ for $M_X(\boldsymbol{v}''_{r+1}, \ldots, \boldsymbol{v}''_{n-k})$, we conclude that there exists a channel $\mathcal{U}_X$, which is formed from sequence of CNOT gates acting on qubits $r+1$ through n *alone*, such that

$$\forall j \in \{r+1, \ldots, n-k\} : \mathcal{U}_X \circ \mathcal{U}_Z[X(\boldsymbol{v}_j)] = X_j. \tag{B.14}$$

Finally, we define unitary channels $\mathcal{U}_C \coloneqq U_C(\cdot)U_C$ and $U_{\mathrm{MV}} \coloneqq U_{\mathrm{MV}}(\cdot)U_{\mathrm{MV}}$ where,

$$\begin{aligned} U_C &\coloneqq \left[\prod_{i=r+1}^{n-k} X_i^{b_i}\right]\left[\prod_{i=1}^{r} Z_i^{b_i}\right], \\ U_{\mathrm{MV}} &\coloneqq \prod_{i=0}^{n-k-1} U_{\mathrm{SWAP}}(n-i, n-k-i). \end{aligned} \tag{B.15}$$

which respectively remove signs from the generators of $\mathsf{S}(\mathcal{C})$ and moves qubits 1 through $n-k$ to $k+1$ through n. These enable us to construct the unitary channel $\mathcal{U} \coloneqq \mathcal{U}_{\mathrm{MV}} \circ \mathcal{U}_C \circ \mathcal{U}_X \circ \mathcal{U}_Z$ satisfying Eq. (B.9) and Eq. (B.9) from the Lemma statement. Since $\mathcal{U}$ is formed from CNOT-, single-qubit X- and Z-gates, it is completely CSS-preserving unitary channel. □

With this Lemma in hand, we can show that every trace-preserving CSS code projection can be executed as a CSS circuit, which leads to the following Lemma from the main text.

Lemma 32. *Every trace-preserving qubit CSS code projection can be executed as a CSS circuit, and is therefore stochastically represented by the Wigner representation defined in Eq. (3.18).*

Proof. When $k = n$, the trace-preserving qubit CSS code projection protocol becomes introduction of a CSS state $|0\rangle\langle 0|$ on the ancillary qubit, which is completely CSS-preserving. Otherwise, the stabilizer group $\mathsf{S}(\mathcal{C})$ of an $[[n, k]]$ CSS code can be generated by a set of non-trivial $(n-k)$ CSS observables $\{S_1, \ldots, S_{n-k}\}$, and its trace-preserving code projection protocol $\Lambda_{\mathcal{C}}$.

1. Projectively measure $S_1, \ldots, S_{n-k}$.
2. If the no-error syndrome is obtained, decode onto the first k qubits and discard the final $(n-k)$ qubits. By Lemma 74, the decoding channel is a completely CSS-preserving

unitary channel $\mathcal{U}$ followed by Hadamard gates on some of the final $(n-k)$ qubits. Because the trace is invariant under circular shifts, this step is equivalent to just performing the completely CSS-preserving unitary channel $\mathcal{U}$ and then discarding the final $(n-k)$ qubits. Furthermore, prepare an ancillary qubit in the state $|0\rangle$.

3. Otherwise, discard all qubits, prepare a k-qubit CSS state σ, and prepare the ancillary qubit in the state $|1\rangle$.

We see from Lemma 19 that $\Lambda_{\mathcal{C}}$ is completely CSS-preserving, and is therefore stochastically represented by Corollary 2. □

B.2 Entropic constraints on qubit CSS code projection protocols

B.2.1 Structure of entropic constraints

In this section, we prove several lemmas that will help us simplify the expression of entropic constraint functions ΔD_α defined for qubit CSS code projection protocols in Eq. (4.17). To this end, we find it useful to introduce the general mean $Q_\alpha(\boldsymbol{w}||\boldsymbol{r})$ for any $\alpha \in A$ on a quasiprobability distribution $\boldsymbol{w} \coloneqq (w_1, \ldots w_N)$ and probability distribution $\boldsymbol{r} \coloneqq (r_1, \ldots r_N)$, as

$$Q_\alpha(\boldsymbol{w}||\boldsymbol{r}) \coloneqq 2^{(\alpha-1)D_\alpha(\boldsymbol{w}||\boldsymbol{r})} = \sum_{i=1}^{N} w_i^\alpha r_i^{1-\alpha}. \tag{B.16}$$

Lemma 75. *Consider k-rebit states ρ_0 and ρ_1, k-qubit CSS states τ_0 and τ_1, distinct m-qubit computational basis states ψ_0 and ψ_1, and probability distributions $\{p_0, p_1\}$ and $\{q_0, q_1\}$. Given any $\alpha \in \mathcal{A}$, we then have*

$$Q_\alpha(W_{\sum_i p_i \rho_i \otimes \psi_i}||W_{\sum_i q_i \tau_i \otimes \psi_i}) = \sum_{i \in \{0,1\}} p_i^\alpha q_i^{1-\alpha} Q_\alpha(W_{\rho_i}||W_{\tau_i}), \tag{B.17}$$

which in turn implies the inequality

$$Q_\alpha(W_{\sum_i p_i \rho_i \otimes \psi_i}||W_{\sum_i q_i \tau_i \otimes \psi_i}) \geq p_i^\alpha q_i^{1-\alpha} Q_\alpha(W_{\rho_i}||W_{\tau_i}), \tag{B.18}$$

for each $i \in \{0, 1\}$.

Proof. Recalling that computational basis states are CSS and therefore have real Wigner rep-

resentations, we have by Property **(A2)** of the phase-point operators in our chosen Wigner representation that

$$\langle \psi_0, \psi_1 \rangle = 2^m \sum_{\boldsymbol{u} \in \mathcal{P}_m} W_{\psi_0}(\boldsymbol{u}) W_{\psi_1}(\boldsymbol{u}) = 0 \tag{B.19}$$

because ψ_0 and ψ_1 are orthogonal.

Since computational basis states are CSS, we must further have $W_{\psi_i}(\boldsymbol{u}) \geq 0$. We thus conclude from Eq. (B.19) that

$$V_0 := \mathrm{supp}(W_{\psi_0}) \subseteq \ker(W_{\psi_1}), \tag{B.20}$$

$$V_1 := \mathrm{supp}(W_{\psi_1}) \subseteq \ker(W_{\psi_0}). \tag{B.21}$$

With this in hand, we can explicitly evaluate

$$\begin{aligned}
&Q_\alpha(W_{\sum_i p_i \rho_i \otimes \psi_i} || W_{\sum_i q_i \tau_i \otimes \psi_i}) \\
&= \sum_{i,j \in \{0,1\}} \sum_{\boldsymbol{u} \in \mathcal{P}_k} \sum_{\boldsymbol{v} \in \mathcal{P}_m} \left(p_i W_{\rho_j}(\boldsymbol{u}) W_{\psi_j}(\boldsymbol{v})\right)^\alpha \left(q_i W_{\tau_i}(\boldsymbol{u}) W_{\psi_i}(\boldsymbol{v})\right)^{1-\alpha} \\
&= \sum_{i \in \{0,1\}} p_i^\alpha q_i^{1-\alpha} \sum_{\boldsymbol{u} \in \mathcal{P}_k} W_{\rho_i}(\boldsymbol{u})^\alpha W_{\tau_i}(\boldsymbol{u})^{1-\alpha} \\
&= \sum_{i \in \{0,1\}} p_i^\alpha q_i^{1-\alpha} Q_\alpha(W_{\rho_i} || W_{\tau_i}),
\end{aligned} \tag{B.22}$$

where in the second equality we used the normalisation of our chosen representation. The inequality in the Lemma statement then follows from the fact that both terms summed in the right hand side of Eq. (B.17) must be non-negative for all $\alpha \in \mathcal{A}$. □

With this property in hand, we obtain the following Lemma, which makes the non-trivial n-dependence in ΔD_α more explicit.

Lemma 76. *Let μ denote the maximally mixed state on k qubits. Given any $\alpha \in \mathcal{A}$, the function ΔD_α can then be expressed as*

$$\Delta D_\alpha = n\left(1 - H_\alpha[W_{\rho_{\mathrm{in}}}]\right) + k - \frac{1}{\alpha - 1} \log \left[p^\alpha Q_\alpha\left(W_{\rho_{\mathrm{out}}} || W_\mu\right) + (1-p)^\alpha \left(\frac{1}{2^{n-k} - 1}\right)^{\alpha - 1} \right] \tag{B.23}$$

Proof. By Lemma 75, we have the expansion

$$\begin{aligned} Q_\alpha\left(W_{\rho_p} \,||\, W_{\tau_{n,k}}\right) &= p^\alpha \left(2^{k-n}\right)^{1-\alpha} Q_\alpha\left(W_{\rho_{\text{out}}} \,||\, W_\mu\right) + (1-p)^\alpha \left(1-2^{k-n}\right)^{1-\alpha} Q_\alpha\left(W_\sigma \,||\, W_\sigma\right) \\ &= \left(2^{n-k}\right)^{\alpha-1} \left[p^\alpha Q_\alpha\left(W_{\rho_{\text{out}}} || W_\mu\right) + (1-p)^\alpha \left(\frac{1}{2^{n-k}-1}\right)^{\alpha-1}\right], \end{aligned} \tag{B.24}$$

where in the last equality we have used $Q_\alpha(\boldsymbol{p}||\boldsymbol{p}) = 1$ for all probability distributions $\boldsymbol{p}$. We therefore obtain

$$D_\alpha\left(W_{\rho_p} \,||\, W_{\tau_{n,k}}\right) = n - k + \frac{1}{\alpha-1} \log\left[p^\alpha Q_\alpha\left(W_{\rho_{\text{out}}} || W_\mu\right) + (1-p)^\alpha \left(\frac{1}{2^{n-k}-1}\right)^{\alpha-1}\right]. \tag{B.25}$$

Substituting Eq. (B.25) and $D_\alpha\left(W_\rho \middle|\middle| W_{\frac{\mathbb{1}}{2}}\right) = 2 - H_\alpha[W_\rho]$ into Eq. (4.17) gives the Lemma. □

By inspection, the form for ΔD_α given in Lemma 76 has no σ-dependence, which serves as another proof of Lemma 34.

B.2.2 Properties of qubit CSS entropic constraints (Proof of Lemma 35)

We now prove the following properties of the constraint function ΔD_α from the main text:

Lemma 35. *Given any noisy input rebit magic state ρ_{in}, output k-rebit magic state ρ_{out}, acceptance probability p and $\alpha \in \mathcal{A}$, we have that*

(i) *ΔD_α is concave over the domain $n \in [k, \infty]$.*

(ii) *ΔD_α becomes negative as $n \to k$, i.e. $\lim_{n\to k} \Delta D_\alpha < 0$.*

(iii) *If $H_\alpha[W_{\rho_{\text{in}}}] > 1$, then we have that ΔD_α also becomes negative in the asymptotic limit, i.e. $\lim_{n\to\infty} \Delta D_\alpha < 0$.*

Proof. Let us denote the maximally mixed state on k qubits by μ. We further simplify notation by defining the constants $c_1 \coloneqq p^\alpha Q_\alpha(W_{\rho_{\text{out}}}||W_\mu)$ and $c_2 \coloneqq (1-p)^\alpha$.

***Proof of* (i):** Let us define the following function

$$g(n) \coloneqq \left[c_1 + c_2 \left(\frac{1}{2^{n-k}-1}\right)^{\alpha-1}\right]. \tag{B.26}$$

This means that from Lemma 76 we can write

$$\Delta D_\alpha = n\left(1 - H_\alpha[W_{\rho_{\text{in}}}]\right) + k - \frac{1}{\alpha - 1}\log g(n), \tag{B.27}$$

and since the first term is linear in n we need only check the second derivative of the second term to establish that ΔD_α is concave. We have

$$\begin{aligned}\partial_n^2 \Delta D_\alpha &= -\frac{1}{\alpha - 1}\partial_n^2 \log g(n) \\ &= -\left[\frac{\ln 2\, c_2 2^{k+n}\left(c_1\left(2^k + (\alpha-1)2^n\right)\left(2^{n-k}-1\right)^\alpha + c_2\left(2^n - 2^k\right)\right)}{\left(2^n - 2^k\right)\left(c_1 2^k\left(2^{n-k}-1\right)^\alpha + c_2\left(2^n - 2^k\right)\right)^2}\right].\end{aligned}$$

Since $c_1, c_2 \geq 0$ for all ρ_{out} and p, the term in square brackets must be non-negative for all $n \geq k, \alpha > 1, \rho_{\text{out}}$ and p (strictly positive for $p < 1$). This implies $\partial_n^2 \Delta D_\alpha$ is non-positive everywhere on $n \in [k, \infty]$ for any given ρ_{out} and p. Therefore ΔD_α is concave, as claimed.

Proof of (ii)***:*** Recalling that $\alpha > 1$, we have from Eq. (B.27) that

$$\lim_{n\to k} \Delta D_\alpha = k(2 - H_\alpha[W_{\rho_{\text{in}}}]) - \frac{1}{\alpha - 1}\lim_{n\to k}\left\{\log\left[c_1 + c_2\left(\frac{1}{2^{n-k}-1}\right)^{\alpha-1}\right]\right\} = -\infty \tag{B.28}$$

so long as $c_2 > 1$, which is true if and only if $p < 1$.

Proof of (iii)***:*** We have from Eq. (B.27) that

$$\lim_{n\to\infty} \Delta D_\alpha = k - \frac{1}{\alpha - 1}\log[c_1] + \lim_{n\to\infty}\left\{n\left(1 - H_\alpha[W_{\rho_{\text{in}}}]\right)\right\} = -\infty \tag{B.29}$$

if $H_\alpha[W_{\rho_{\text{in}}}] > 1$, as claimed. □

B.3 Paired CSS code projection protocols for Hadamard distillation

Lemma 77. *Consider an $[[n, 1]]$ qubit CSS code $\mathcal{C}$ whose projection protocol sends n copies of $\rho(\epsilon_{\text{in}})$ to the rebit state ρ_{out} with acceptance probability p. Then there exists an n-to-1 qubit CSS code $\tilde{\mathcal{C}}$ whose projection protocol sends n copies of $\rho(\epsilon_{\text{in}})$ to $H(\rho_{\text{out}})H$ with the same acceptance probability.*

Proof. We can represent the action of the code projection protocol constructed around $\mathcal{C}$, $\Phi_{\mathcal{C}}$,

on n copies of $\rho(\epsilon_{\text{in}})$ in terms of the encoding unitary $U_{\mathcal{C}}$ and codespace projector $P_{\mathcal{C}}$ as

$$\Phi_{\mathcal{C}}(\rho(\epsilon)^{\otimes n}) = \text{Tr}_{2,\ldots,n}[U_{\mathcal{C}}^{\dagger} P_{\mathcal{C}} (\rho(\epsilon)^{\otimes n}) P_{\mathcal{C}} U_{\mathcal{C}}] = p\rho_{\text{out}} \tag{B.30}$$

The stabilizer group defining the codespace of $\mathcal{C}$ can be represented in the form

$$\mathsf{S}(\mathcal{C}) \coloneqq \langle (-1)^{b_1} Z(\boldsymbol{u}_1), \ldots, (-1)^{b_r} Z(\boldsymbol{u}_r), (-1)^{b_{r+1}} X(\boldsymbol{v}_{r+1}), \ldots, (-1)^{b_{n-1}} X(\boldsymbol{v}_{n-1})) \rangle. \tag{B.31}$$

for bits $b_1 \ldots, b_{n-1}$, n-bit vectors $\boldsymbol{u}_1, \ldots, \boldsymbol{u}_r$ and $\boldsymbol{v}_{r+1}, \ldots, \boldsymbol{v}_{n-1}$, and some $r \in \{0, \ldots, n-1\}$. By Lemma 74, U may be represented as $U = [\mathbb{1}^{\otimes (r+1)} \otimes H^{\otimes (n-r)}] V_{\mathcal{C}}$ for some completely CSS-preserving unitary $V_{\mathcal{C}}$ such that

$$\forall i \in \{1, \ldots, r\} : V_{\mathcal{C}} [(-1)^{b_i} Z(\boldsymbol{u}_i)] V_{\mathcal{C}}^{\dagger} = Z_{i+1}, \tag{B.32}$$

$$\forall j \in \{r+1, \ldots, n-1\} : V_{\mathcal{C}} [(-1)^{b_j} X(\boldsymbol{v}_j)] V_{\mathcal{C}}^{\dagger} = X_{j+1}. \tag{B.33}$$

We now observe that

$$pH(\rho_{\text{out}})H = H \left(\text{Tr}_{2,\ldots,n} [U_{\mathcal{C}} P_{\mathcal{C}} (\rho(\epsilon)^{\otimes n}) P_{\mathcal{C}} U_{\mathcal{C}}^{\dagger}] \right) H = \text{Tr}_{2,\ldots,n} [H_1 U_{\mathcal{C}} P_{\mathcal{C}} (\rho(\epsilon)^{\otimes n}) P_{\mathcal{C}} U_{\mathcal{C}}^{\dagger} H_1], \tag{B.34}$$

where H_1 is the Hadamard gate on the first qubit. The form of $U_{\mathcal{C}}$ and the cyclic property of the trace then means we can execute Hadamard gates on the final $(n-1)$ qubits before discarding them to obtain

$$pH(\rho_{\text{out}})H = \text{Tr}_{2,\ldots,n} [H_{\text{all}} V_{\mathcal{C}} P_{\mathcal{C}} (\rho(\epsilon_{\text{in}})^{\otimes n}) P_{\mathcal{C}} V_{\mathcal{C}}^{\dagger} H_{\text{all}}] = \text{Tr}_{2,\ldots,n} [\tilde{V}_{\mathcal{C}} (\rho(\epsilon_{\text{in}})^{\otimes n}) \tilde{P}_{\mathcal{C}} \tilde{V}_{\mathcal{C}}^{\dagger}], \tag{B.35}$$

where we have defined $\tilde{V}_{\mathcal{C}} \coloneqq H_{\text{all}}(V_{\mathcal{C}}) H_{\text{all}}$ and $\tilde{P}_{\mathcal{C}} \coloneqq H_{\text{all}}(P_{\mathcal{C}}) H_{\text{all}}$ for the simultaneous Hadamard gate on all n qubits, H_{all}. We note that the final equality relies on the invariance of $\rho(\epsilon)$ under the Hadamard gate.

From the definition of $P_{\mathcal{C}}$ in Eq. (4.5), we have that

$$P_{\mathcal{C}} = \prod_{i=1}^{n-1} \left[\frac{1}{2} (I + (-1)^{b_i} S_i) \right] \tag{B.36}$$

Therefore, $\tilde{P}_{\mathcal{C}}$ is the codespace projector onto another $[[n, 1]]$ CSS code $\tilde{\mathcal{C}}$ whose stabilizer group

is generated by

$$\mathsf{S}(\tilde{\mathcal{C}}) \coloneqq \langle (-1)^{b_1} X(\boldsymbol{u}_1), \ldots, (-1)^{b_r} X(\boldsymbol{u}_r), (-1)^{b_{r+1}} Z(\boldsymbol{v}_{r+1}), \ldots, (-1)^{b_{n-1}} Z(\boldsymbol{v}_{n-1})) \rangle. \tag{B.37}$$

Since $V_{\mathcal{C}}$ is a completely CSS-preserving unitary, it must be a product of $CNOT_{j,k}, Z_j, X_j$ for $j \neq k \in \{1, \ldots, n\}$. Therefore, $\tilde{V}_{\mathcal{C}}$ is a product of $H_{\text{all}} CNOT_{j,k} H_{\text{all}} = CNOT_{k,j}$, $H_{\text{all}} Z_j H_{\text{all}} = X_j$ and $H_{\text{all}} X_j H_{\text{all}} = Z_j$, and thus also a completely CSS-preserving unitary. Furthermore, by Eq. (B.32) and Eq. (B.33), we have that

$$\forall i \in \{1, \ldots, r\} : \tilde{V}_{\mathcal{C}}[(-1)^{b_i} X(\boldsymbol{u}_i)] \tilde{V}_{\mathcal{C}}^{\dagger} = H_{\text{all}} V_{\mathcal{C}}[(-1)^{b_i} Z(\boldsymbol{u}_i)] V_{\mathcal{C}}^{\dagger} H_{\text{all}} = X_{i+1}, \tag{B.38}$$

$$\forall j \in \{r+1, \ldots, n-1\} : \tilde{V}_{\mathcal{C}}[(-1)^{b_j} X(\boldsymbol{v}_j)] \tilde{V}_{\mathcal{C}}^{\dagger} = H_{\text{all}} V_{\mathcal{C}}[(-1)^{b_i} X(\boldsymbol{v}_j)] V_{\mathcal{C}}^{\dagger} H_{\text{all}} = Z_{j+1}. \tag{B.39}$$

Therefore, $\tilde{V}_{\mathcal{C}}$ followed by some swaps amongst the final $(n-1)$ qubits, which can be carried out by CNOT gates on those qubits, and Hadamard gates on some of the final $(n-1)$ qubits, produces the encoding unitary $U_{\tilde{\mathcal{C}}}$ of the code $\tilde{\mathcal{C}}$. We thus conclude

$$pH(\rho_{\text{out}})H = \Phi_{\tilde{\mathcal{C}}}[\rho(\epsilon)^{\otimes n}], \tag{B.40}$$

i.e. the code projection protocol of the CSS code $\tilde{\mathcal{C}}$ produces $H(\rho_{\text{out}})H$ from n copies of $\rho(\epsilon_{\text{in}})$ with acceptance probability p. □

The even mixture of ρ_{out} and $H(\rho_{\text{out}})H$, $\mathcal{G}_H(\rho_{\text{out}}) \coloneqq \frac{1}{2}((\rho_{\text{out}}) + H(\rho_{\text{out}})H)$, is the projection of ρ_{out} onto the Bloch sphere axis from $|H\rangle$ to $|\overline{H}\rangle$. Therefore, we have that $\mathcal{G}_H(\rho_{\text{out}}) = \rho(\epsilon_{\text{out}})$, an ϵ_{out}-depolarised $|H\rangle$, or $\mathcal{G}_H(\rho_{\text{out}}) = \overline{\rho}(\epsilon_{\text{out}})$, an ϵ_{out}-depolarised $|\overline{H}\rangle$, where $\epsilon_{\text{out}} \in [0, 1]$.

Let us now define the channel $\Lambda'_{\mathcal{C}}$ for an $[[n, 1]]$ CSS code $\mathcal{C}$, which performs the trace-preserving code projection protocol for $\mathcal{C}$ with the addition that we perform the gate XZ on the accepted output, i.e.

$$\Lambda'_{\mathcal{C}}(\cdot) \coloneqq [XZ(\text{Tr}_{2,\ldots,n}[\mathcal{U}_{\mathcal{C}}^{\dagger} \circ \Pi_{\mathcal{C}}(\cdot)])ZX] \otimes |0\rangle\langle 0| + \sigma \,\text{Tr} \otimes |1\rangle\langle 1| \tag{B.41}$$

Since the gate ZX is completely CSS-preserving, $\Lambda'_{\mathcal{C}}$ also lies in the CSS subtheory and is there-

fore stochastically represented. With this in hand, let us next define the channel

$$\Lambda := \begin{cases} \frac{1}{2}(\Lambda_{\mathcal{C}} + \Lambda_{\tilde{\mathcal{C}}}) & \text{if } \mathcal{G}_H(\rho_{\text{out}}) = \rho(\epsilon_{\text{out}}) \\ \frac{1}{2}(\Lambda'_{\mathcal{C}} + \Lambda'_{\tilde{\mathcal{C}}}) & \text{if } \mathcal{G}_H(\rho_{\text{out}}) = \overline{\rho}(\epsilon_{\text{out}}) \end{cases} \text{ for } \epsilon_{\text{out}} \in [0,1]/ \tag{B.42}$$

As an equal mixture of stochastically-represented channels, Λ is itself stochastically represented, and transforms n copies of $\rho(\epsilon_{\text{in}})$ into

$$\Lambda(\rho(\epsilon_{\text{in}})^{\otimes n}) = p\rho(\epsilon_{\text{out}}) \otimes |0\rangle\langle 0| + (1-p)\sigma \otimes |1\rangle\langle 1|). \tag{B.43}$$

because $XZ(\overline{\rho}(\epsilon_{\text{out}}))ZX = \rho(\epsilon_{\text{out}})$. Furthermore, Λ maps n copies of $\frac{\mathbb{1}}{2}$ to $\tau_{n,k}$. Therefore, if there exists an n-to-1 CSS code projection that transforms n copies of $\rho(\epsilon)$ to ρ_{out} with acceptance probability p, then there exists a stochastically-represented operation that transforms n copies of $\rho(\epsilon)$ to $\rho(\epsilon_{\text{out}})$ with acceptance probability p while also mapping n copies of $\frac{\mathbb{1}}{2}$ to $\tau_{n,k}$. Consequently, the entropic constraints given by 33 remain valid after replacing ρ_{out} by $\rho(\epsilon_{\text{out}})$ in ρ_p for some $\epsilon_{\text{out}} \in [0,1]$.

B.4 Bounds on code length in stabilizer code projection protocols

Throughout this section, we will use "Wigner representation", denoted W, as a shorthand covering the representation defined in Eq. (3.18) for qubits and Gross's Wigner representation $W^{(G)}$ for qudits of odd prime dimensions. We first remind the reader of the shared properties between our chosen qubit Wigner representation and Gross's Wigner representation and that power the extension of results from the former to the latter throughout this section. Given a state ρ on n qudits of odd prime dimension d, its Gross's Wigner representation $W_\rho^{(G)}$ is defined at each point $\boldsymbol{u}$ in a phase space $\mathcal{P}_n^{(G)} := \mathbb{Z}_d^{\times n} \times \mathbb{Z}_d^{\times n}$ as [145]

$$W_\rho^{(G)}(\boldsymbol{u}) := \frac{1}{d^n}\langle A_{\boldsymbol{u}}^{(G)}, \rho\rangle, \tag{B.44}$$

where $\{A_{\boldsymbol{u}}^{(G)}\}$ are the set of d^{2n} *phase-point operators* defined on each point $\boldsymbol{u} \in \mathcal{P}_n^{(G)}$ as

$$A_{\boldsymbol{0}} := \frac{1}{d^n}\sum_{\boldsymbol{v}\in\mathcal{P}_n^{(G)}} D_{\boldsymbol{v}}, \quad A_{\boldsymbol{u}} := D_{\boldsymbol{u}} A_{\boldsymbol{0}} D_{\boldsymbol{u}}^\dagger \tag{B.45}$$

for the set of d^{2n} displacement operators $\{D_{\boldsymbol{u}}^{(G)}\}$ on n qudits of prime dimension d defined for each point $\boldsymbol{u} \coloneqq (\boldsymbol{u}_x, \boldsymbol{u}_z) \in \mathcal{P}_n^{(G)}$ as

$$D_{\boldsymbol{u}} \coloneqq \omega^{2^{-1}(\boldsymbol{u}_z \cdot \boldsymbol{u}_x)} Z_d(\boldsymbol{u}_z) X_d(\boldsymbol{u}_x), \quad \omega \coloneqq \exp\left(\frac{2\pi i}{d}\right). \tag{B.46}$$

where arithmetic is modulo d. Given a channel Φ from n to m qudits of prime dimension d, Gross's Wigner representation has been extended by Ref. [61] as

$$W_{\Phi}^{(G)}(\boldsymbol{v}|\boldsymbol{u}) \coloneqq \left\langle A_{\boldsymbol{u}}^{(G)^T} \otimes A_{\boldsymbol{v}}, \otimes, \mathcal{J}(\Phi) \right\rangle, \tag{B.47}$$

where transposition is carried out in the same basis of the input system with which the Choi representation of Φ is calculated. It has been shown [145, 61] that the phase-point operators of Gross's Wigner representation shares Properties **(A1)**–**(A3)** with those of our chosen qubit representation, and Gross's Wigner representation itself shares Properties **(W1)**–**(W6)** with our chosen qubit Wigner representation. Because of these properties, Gross's Wigner representation constitutes a generalised Gross's Wigner representation with respect to the Bravyi-Kitaev model because it represents stabilizer operations stochastically [61].

B.4.1 Generalisation of entropic constraints on qubit CSS code projections to qudit stabilizer code projections

We first explain how Theorem 33 can be extended to all stabilizer code projection protocols on qudits of odd prime dimension d. Because Gross's Wigner representation stochastically represents *every* stabilizer operation for qudits of odd prime dimension [60], we can use the proof strategy of Theorem 30 to prove its analogue for qudits of odd prime dimension.

Theorem 78. *Let ρ be a noisy magic state and τ be a stabilizer state on a single qudit such that $\mathrm{supp}(W_\rho) \subsetneqq \mathrm{supp}(W_\tau)$. If there exists a magic distillation protocol Φ such that $\Phi(\rho^{\otimes n}) = \rho'$ and $\tau' \coloneqq \Phi(\tau^{\otimes n})$, then*

$$\Delta D_\alpha \coloneqq n D_\alpha(W_\rho || W_\tau) - D_\alpha(W_{\rho'} || W_{\tau'}) \geq 0 \tag{B.48}$$

for Gross's Wigner representation W and all $\alpha \in \mathcal{A} \coloneqq \{\frac{2a}{2b-1}, |a, b \in \mathbb{N}^+, a \geq b\}$.

Much as we had already done for CSS code projection protocols, we can construct a trace-preserving equivalent to any given n-to-k stabilizer code projection protocol by preparing a

specially designed stabilizer state σ on k qudits whenever the output is rejected, while recording acceptance (rejection) in computational basis states $|0\rangle$ ($|1\rangle$) of an ancillary qudit. Since such a trace-preserving stabilizer code projection belongs to the stabilizer subtheory for qudits of odd prime dimension d, it is stochastically represented under Gross's Wigner representation. Furthermore, every trace-preserving stabilizer code projection protocol has the form $\Lambda_{\mathcal{C}}$ from Eq. (4.9) once we let $\mathcal{C}$ be any $[[n,k]]$ stabilizer code and σ be any k-qudit stabilizer state, so by the reasoning of section 4.3.2, we can take the following analogous reference process to that defined for qubit CSS code projection protocols in Eq. (4.16),

$$\Lambda_{\mathcal{C}}\left(\left[\frac{\mathbb{1}}{d}\right]^{\otimes n}\right) = d^{k-n}\left[\frac{\mathbb{1}}{d}\right]^{\otimes k} \otimes |0\rangle\langle 0| + (1-d^{k-n})\sigma \otimes |1\rangle\langle 1|\,. \tag{B.49}$$

As Gross's Wigner representation for the maximally mixed state on a qudit of prime dimension d is the uniform distribution on a d^2 sample space, the representation of any qudit state ρ_{in} will have support inside the representation of the maximally mixed state. We can therefore apply the above Theorem to trace-preserving stabilizer code projection protocols on qudits of odd prime dimension d with the above reference process to obtain Theorem 37.

B.4.2 Continuity of α-Réyni entropy in state

In this section, we establish a bound on the difference between α-Réyni entropies calculated on Wigner representations of states ρ and σ based on the distance between the states themselves. We first define the ℓ_1- and ℓ_2-norms, respectively, of a vector $\boldsymbol{w} \in \mathbb{R}^d$ as

$$\|\boldsymbol{w}\|_1 \coloneqq \sum_{i=1}^{d} |w_i|, \tag{B.50}$$

$$\|\boldsymbol{w}\|_2 \coloneqq \left[\sum_{i=1}^{d} |w_i|^2\right]^{\frac{1}{2}}. \tag{B.51}$$

We now make use of the following result from the literature on real vector spaces (e.g. see [146]), which is a consequence of the Cauchy-Schwarz inequality.

Lemma 79. *For all* $\boldsymbol{w} \in \mathbb{R}^d$

$$\|\boldsymbol{w}\|_1 \leq \sqrt{d}\|\boldsymbol{w}\|_2. \tag{B.52}$$

This result enables us to show that vanishingly small variations in quantum states correspond to vanishingly small variations in their Wigner representations.

Lemma 80. *If $\|\rho - \sigma\|_1 \leq \epsilon$ then for any Wigner representation W, we have*

$$\|W_\rho - W_\sigma\|_1 \leq \sqrt{d}\epsilon. \tag{B.53}$$

Proof. To simplify notation, we first define the state difference $\Delta := \rho - \sigma$ such that $W_\Delta = W_\rho - W_\sigma$. Since the Schatten-$p$ norms are non-increasing with respect to p [146], we obtain

$$\begin{aligned}
\|\rho - \sigma\|_1 \geq \|\rho - \sigma\|_2 &= \left\| \sum_{\boldsymbol{x}} W_\Delta(\boldsymbol{x}) A_{\boldsymbol{x}} \right\|_2 \\
&= \sqrt{\sum_{\boldsymbol{x},\boldsymbol{y}} W_\Delta^*(\boldsymbol{x}) W_\Delta(\boldsymbol{y}) \langle A_{\boldsymbol{x}}, A_{\boldsymbol{y}} \rangle} \\
&= \sqrt{\sum_{\boldsymbol{x},\boldsymbol{y}} W_\Delta^*(\boldsymbol{x}) W_\Delta(\boldsymbol{y}) d\delta_{\boldsymbol{x},\boldsymbol{y}}} \\
&= \sqrt{d}\|W_\Delta\|_2 \geq \frac{1}{\sqrt{d}}\|W_\Delta\|_1,
\end{aligned} \tag{B.54}$$

where in the second inequality we employ Lemma 79. We thus obtain $\|W_\rho - W_\sigma\|_1 \leq \sqrt{d}\|\rho - \sigma\|_1$, which completes the proof. □

Lemma 81. *Let ρ and σ be two quantum states of a d-dimensional qudit such that $\|\rho - \sigma\|_1 \leq \epsilon$. Then*

$$|H_\alpha[W_\rho] - H_\alpha[W_\sigma]| \leq \frac{\alpha}{\alpha - 1} \log \left[1 + \epsilon d^{\frac{5}{2}}\right]. \tag{B.55}$$

Proof. Theorem 7 (2) of Ref. [147] tells us that for two d^2-dimensional distributions $\boldsymbol{w}, \boldsymbol{w}'$, we have the following continuity statement on the α-Rényi entropies

$$\left|H_\alpha(\boldsymbol{w}) - H_\alpha(\boldsymbol{w}')\right| \leq \frac{\alpha}{\alpha - 1} \log \left[1 + \left\|\boldsymbol{w} - \boldsymbol{w}'\right\|_1 d^2\right]. \tag{B.56}$$

The Lemma result then follows from Lemma 80. □

B.4.3 Proof of Theorem 38

We are now in a position to derive the analytic bounds on code length for stabilizer code protocols on qudits of prime dimension d presented in Theorem 38.

Theorem 38 (Qudit code length bounds)**.** *Consider the distillation of a target pure magic state ψ*

from a supply of the noisy magic state ρ_{in}, where ψ and ρ_{in} are states of a single rebit (qudit of odd prime dimension d). Any n-to-k CSS (stabilizer) code projection protocol that distils n copies of ρ_{in} into a k-rebit (qudit) state ρ_{out} with acceptance probability p and per-rebit (qudit) output error ϵ_{out} must have a code length n such that

$$n \geq \frac{k\left[\log d - H_\alpha(W_\psi)\right] - \frac{\alpha}{1-\alpha}\log\left(\frac{p}{1+k\epsilon_{\text{out}}d^{5/2}}\right)}{\left[\log d - H_\alpha(W_{\rho_{\text{in}}})\right]}, \tag{4.27}$$

for all $\alpha \in \mathcal{A}$ for which $H_\alpha(W_{\rho_{\text{in}}}) < \log d$, and

$$n \leq \frac{k\left[H_\alpha(W_\psi) - \log d\right] + \frac{\alpha}{1-\alpha}\log\left(\frac{p}{1+k\epsilon_{\text{out}}d^{5/2}}\right)}{\left[H_\alpha(W_{\rho_{\text{in}}}) - \log d\right]}, \tag{4.28}$$

for all $\alpha \in \mathcal{A}$ for which $H_\alpha(W_{\rho_{\text{in}}}) > \log d$, where W is the qubit Wigner representation defined in Eq. (3.18) (Gross's Wigner representation).

Proof. Let μ denote the maximally mixed state of k qudits with prime dimension d. We first note that Lemma 75 extends straightforwardly to Gross's Wigner representation upon replacing rebits by qudits of odd prime dimension d and CSS states by stabilizer states. Applying this extended Lemma to the parts of ρ_p and $\tau_{n,k}$ that are tagged by $|0\rangle\langle 0|$ on the ancillary qudit, we obtain

$$Q_\alpha(W_{\rho_p}||W_{\tau_{n,k}}) \geq \frac{p^\alpha}{d^{(n-k)(1-\alpha)}} Q_\alpha(W_{\rho_{\text{out}}}||W_\mu). \tag{B.57}$$

By applying the fact that $\log(\cdot)/(\alpha-1)$ is a monotonically increasing function for $\alpha > 1$ to both sides of the above equation, we see that

$$\begin{aligned} D_\alpha(W_{\rho_p}||W_{\tau_{n,k}}) &\geq \frac{1}{\alpha-1}\log\left[\frac{p^\alpha}{d^{(n-k)(1-\alpha)}} Q_\alpha(W_{\rho_{\text{out}}}||W_\mu)\right] \\ &= D_\alpha(W_{\rho_{\text{out}}}||W_\mu) + \frac{\alpha}{\alpha-1}\log p + (n-k)\log d \\ &= k\log d - H_\alpha(W_{\rho_{\text{out}}}) + \frac{\alpha}{\alpha-1}\log p + n\log d, \end{aligned} \tag{B.58}$$

where in the final equality we have used the identity $D_\alpha\left(W_\rho\middle\|W_\mu\right) = 2k\log d - H_\alpha(W_\rho)$ due to Eq. (3.33).

We can now make use of the continuity of the Rényi entropy as stated in Lemma 81 to obtain

$$\left|H_\alpha[W_{\psi^{\otimes k}}] - H_\alpha[W_{\rho_{\text{out}}}]\right| \leq \frac{\alpha}{\alpha-1}\log[1+k\epsilon_{\text{out}}d^{\frac{5}{2}}], \tag{B.59}$$

where we have applied the definition of the output error rate $\epsilon_{\text{out}} \coloneqq k^{-1}\left\|\rho_{\text{out}} - \psi^{\otimes k}\right\|_1$. By applying the above inequality to obtain

$$\begin{aligned} k\log d - H_\alpha(W_{\rho_{\text{out}}}) &= k\log d - H_\alpha(W_{\psi^{\otimes k}}) + (H_\alpha(W_{\psi^{\otimes k}}) - H_\alpha(W_{\rho_{\text{out}}}) \\ &\geq k\log d - H_\alpha(W_{\psi^{\otimes k}}) - \left|(H_\alpha(W_{\psi^{\otimes k}}) - H_\alpha(W_{\rho_{\text{out}}})\right| \\ &\geq k\log d - H_\alpha(W_{\psi^{\otimes k}}) - \frac{\alpha}{\alpha-1}\log[1+k\epsilon_{\text{out}}d^{\frac{5}{2}}], \end{aligned} \tag{B.60}$$

we can lower-bound $D_\alpha(W_{\rho_p}||W_{\tau_{n,k}})$ further as

$$\begin{aligned} D_\alpha(W_{\rho_p}||W_{\tau_{n,k}}) &\geq k\log d - H_\alpha(W_{\psi^{\otimes k}}) - \frac{\alpha}{\alpha-1}\log[1+k\epsilon_{\text{out}}d^{\frac{5}{2}}] + \frac{\alpha}{\alpha-1}\log p + n\log d \\ &= k\left[\log d - H_\alpha(W_\psi)\right] - \frac{\alpha}{1-\alpha}\log\frac{p}{1+k\epsilon_{\text{out}}d^{\frac{5}{2}}} + n\log d, \end{aligned} \tag{B.61}$$

where in the final equality we have used

$$\begin{aligned} H_\alpha(W_{\psi^{\otimes k}}) &= \frac{1}{1-\alpha}\log\sum_{\boldsymbol{u}_1\in\mathcal{P}}\cdots\sum_{\boldsymbol{u}_k\in\mathcal{P}} W_\psi(\boldsymbol{u}_1)^\alpha\ldots W_\psi(\boldsymbol{u}_k)^\alpha \\ &= \frac{1}{1-\alpha}\log\prod_{i=1}^{k}\left[\sum_{\boldsymbol{u}_i\in\mathcal{P}_d} W_\psi(\boldsymbol{u}_i)^\alpha\right] \\ &= k\left(\frac{1}{1-\alpha}\log\sum_{\boldsymbol{u}\in\mathcal{P}} W_\psi(\boldsymbol{u})^\alpha\right) = kH_\alpha(W_\psi) \end{aligned} \tag{B.62}$$

for the phase space $\mathcal{P}_d \coloneqq \mathbb{Z}_d \times \mathbb{Z}_d$ used for the Wigner representation of a single qudit with prime dimension d. Noting that

$$D_\alpha\left(W_{\rho_{\text{in}}^{\otimes n}}||W_{\left(\frac{\mathbb{1}}{d}\right)^{\otimes n}}\right) = nD_\alpha\left(W_{\rho_{\text{in}}}||W_{\left(\frac{\mathbb{1}}{d}\right)}\right) = n[2\log d - H_\alpha(W_{\rho_{\text{in}}})] \tag{B.63}$$

where we applied Property **(D2)** in the first equality and Eq. (3.33) in the second, we obtain the following upper bound on the relative entropy difference ΔD_α

$$0 \leq \Delta D_\alpha \leq n[\log d - H_\alpha(W_{\rho_{\text{in}}})] + k\left[H_\alpha(W_\psi) - \log d\right] + \frac{\alpha}{1-\alpha}\log\frac{p}{1+k\epsilon_{\text{out}}d^{\frac{5}{2}}}. \tag{B.64}$$

This gives a weaker but still necessary constraint on stochastic transformations accomplishing $\rho^{\otimes n} \mapsto \rho_p$ and $\left(\frac{\mathbb{1}}{d}\right)^{\otimes n} \mapsto \tau_{n,k}$. We can rearrange this as

$$n[H_\alpha(W_{\rho_{\text{in}}}) - \log d] \leq k\,[H_\alpha(W_\psi) - \log d] + \frac{\alpha}{1-\alpha} \log \frac{p}{1 + k\epsilon_{\text{out}} d^{\frac{5}{2}}}. \tag{B.65}$$

For $H_\alpha(W_{\rho_{\text{in}}}) < \log d$, we can rearrange Eq. (B.65) to obtain the lower bound in Eq. (4.27). For $H_\alpha(W_{\rho_{\text{in}}}) > \log d$, we obtain the upper bound in Eq. (4.28). □

B.5 Decomposition of CSS magic distillation into code projections

We first define magic distillation protocols in the CSS setting.

Definition 82. *An n-to-k CSS magic distillation protocol is any CSS circuit (as defined in Lemma 19) that takes in $n \geq 2$ qubits and outputs onto the first $1 \leq k < n$ qubits.*

With this definition in hand, we move to the main goal of this section, which is to prove the following theorem.

Theorem 83. *Every n-to-k CSS magic distillation protocol $\mathcal{E}$ can be decomposed as a sum of CSS code projection protocols followed by preparing CSS states and completely CSS-preserving post-processing. Thus one can write*

$$\Phi(\rho) = \sum_i p_i\, \mathcal{U}_i \circ (\Phi_{\mathcal{C}_i}(\rho) \otimes |\varphi_i\rangle\langle\varphi_i|)\,, \tag{B.66}$$

where p_i is a probability, $\mathcal{U}_i$ is a completely CSS-preserving unitary channel on k qubits, $\Phi_{\mathcal{C}_j}$ is the codespace projection protocol for an $[[n, k_i]]$ CSS code $\mathcal{C}_i$ for an integer k_i in the range $0 \leq k_i \leq k$, and $|\varphi_i\rangle$ is a CSS state on $(k - k_i)$ qubits.

Proof. To bring $\mathcal{E}$ into the desired form, we proceed in four steps (all auxiliary lemmas used will be presented after this proof):

1. By Lemma 85, any n-to-k CSS magic distillation protocol Φ can be decomposed as a sum of n-qubit operations

$$\Phi(\rho) = \sum_i q_i \Phi_i \text{ where } \Phi_i \coloneqq \text{Tr}_{k+1,\ldots,n+m}\left[K_i(\rho \otimes |\psi_i\rangle\langle\psi_i|)K_i^\dagger\right] \tag{B.67}$$

wherein q_i is a probability, $|\psi_i\rangle$ is a CSS state on m ancillary qubits and K_i is a Kraus

operator of the form

$$K_i = U_i \left(\prod_{j=1}^{N} P(S_{i,j}) \right), \tag{B.68}$$

for a completely CSS-preserving unitary U_i and projector $P(S_{i,j})$ onto the $+1$ eigenspace of a CSS observable $S_{i,j}$.

2. By Lemma 86, each operation Φ_i can be further decomposed into a sum

$$\Phi_i(\rho) = \sum_{\boldsymbol{s} \in \{0,1\}^{n+m-k}} \Phi_{i,\boldsymbol{s}}(\rho), \tag{B.69}$$

where each operation $\Phi_{i,\boldsymbol{s}}$ first introduces m ancillary qubits in the CSS state $|\psi_i\rangle$, then post-selects the $+1$ outcome in a sequence of projective measurements of CSS observables $S_{i,N}, \ldots, S_{i,1}$, and then performs a CSS code projection protocol on the input and ancillary qubits. Thus one can write

$$\Phi_{i,\boldsymbol{s}}(\rho) = \Phi_{\mathcal{C}_{i,\boldsymbol{s}}} \circ \Pi(S_{i,1}) \circ \cdots \circ \Pi(S_{i,N})[\rho \otimes |\psi_i\rangle\langle\psi_i|], \tag{B.70}$$

where $\Pi(S_{i,j})$ post-selects the $+1$ outcome in a projective measurement of the CSS observable $S_{i,j}$, and $\Phi_{\mathcal{C}_{i,\boldsymbol{s}}}$ is the code projection of an $[[n+m,k]]$ CSS code $\mathcal{C}_{i,\boldsymbol{s}}$.

3. By repeated applications of Lemma 87, we find that $\Phi_{i,\boldsymbol{s}}$ performs a CSS code projection on the input and ancillary qubits, followed by preparing a CSS state and completely CSS-preserving post-processing. Thus one can write

$$\Phi_{i,\boldsymbol{s}}(\rho) = q'_{i,\boldsymbol{s}}\, \mathcal{U}'_{i,\boldsymbol{s}} \circ (\Phi_{\mathcal{C}'_{i,\boldsymbol{s}}}(\rho \otimes |\psi_i\rangle\langle\psi_i|) \otimes |\varphi'\rangle\langle\varphi'|_{i,\boldsymbol{s}}), \tag{B.71}$$

where $q'_{i,\boldsymbol{s}}$ is a probability, $\mathcal{U}'_{i,\boldsymbol{s}}$ is a completely CSS-preserving unitary channel on k qubits, $|\varphi'\rangle_{i,\boldsymbol{s}}$ is a CSS state on $k - k'_{i,\boldsymbol{s}}$ qubits for some integer $k'_{i,\boldsymbol{s}}$ in the range $0 \leq k'_{i,\boldsymbol{s}} \leq k$, and $\Phi_{\mathcal{C}'_{i,\boldsymbol{s}}}$ is a code projection for an $[[n+m,k'_{i,\boldsymbol{s}}]]$ CSS code $\mathcal{C}'_{i,\boldsymbol{s}}$.

4. By Lemma 89, each CSS code projection $\Phi_{\mathcal{C}'_{i,\boldsymbol{s}}}$ on the input and ancillary qubits from Eq. (B.71) can be reduced to a CSS code projection on the input qubits *alone*, followed by preparing a CSS state and completely CSS-preserving post-processing. Thus one can

write

$$\Phi_{\mathcal{C}'_{i,\boldsymbol{s}}}(\rho \otimes |\psi_i\rangle\langle\psi_i|) = q''_{i,\boldsymbol{s}}\, \mathcal{U}''_{i,\boldsymbol{s}} \circ (\Phi_{\mathcal{C}''_{i,\boldsymbol{s}}}(\rho) \otimes |\varphi''\rangle\langle\varphi''|_{i,\boldsymbol{s}}), \tag{B.72}$$

where $q''_{i,\boldsymbol{s}}$ is a probability, $\mathcal{U}''_{i,\boldsymbol{s}}$ is a completely CSS-preserving unitary channel on $k'_{i,\boldsymbol{s}}$ qubits, $\Phi_{\mathcal{C}''_{i,\boldsymbol{s}}}$ is the code projection for an $[[n, k''_{i,\boldsymbol{s}}]]$ CSS code $\mathcal{C}''_{i,\boldsymbol{s}}$ for some integer $k''_{i,\boldsymbol{s}}$ in the range $0 \leq k''_{i,\boldsymbol{s}} \leq k'_{i,\boldsymbol{s}}$, and $|\varphi''\rangle_{i,\boldsymbol{s}}$ is a CSS state on $k'_{i,\boldsymbol{s}} - k''_{i,\boldsymbol{s}}$ qubits. Substituting back then immediately yields the result. □

Auxiliary lemmas

Before turning to the proofs of Lemmas used in each step of the main proof, we present a result on the conjugation relations between completely CSS-preserving unitaries and CSS observables as well as three representations of CSS code projection protocols, all of which will be useful throughout.

Lemma 84. *Given any completely CSS-preserving unitary U and CSS observable S on n qubits, $S' \coloneqq U^\dagger S U$ is another CSS observable of the same type as S. This further implies $P(\pm S)U = UP(\pm S')$.*

Proof. For convenience, let us label the n qubits as $1, \ldots, n$. Let $\boldsymbol{a}$ be an arbitrary n-bit string, and $\boldsymbol{e}_j$ be the n-bit string with 1 in its jth entry and 0 everywhere else. We then have the following conjugation relations

$$Z_j[X(\boldsymbol{a})]Z_j = (-1)^{a_j} X(\boldsymbol{a}) \tag{B.73}$$

$$Z_j[Z(\boldsymbol{a})]Z_j = Z(\boldsymbol{a}) \tag{B.74}$$

$$X_j[X(\boldsymbol{a})]X_j = X(\boldsymbol{a}) \tag{B.75}$$

$$X_j[Z(\boldsymbol{a})]X_j = X_j(-1)^{a_j} Z(\boldsymbol{a}) \tag{B.76}$$

$$\mathrm{CNOT}(i,j)[X(\boldsymbol{a})]\,\mathrm{CNOT}(i,j) = X(\boldsymbol{a} + a_i \boldsymbol{e}_j) \tag{B.77}$$

$$\mathrm{CNOT}(i,j)[Z(\boldsymbol{a})]\,\mathrm{CNOT}(i,j) = Z(\boldsymbol{a} + a_j \boldsymbol{e}_i), \tag{B.78}$$

for any i, j from the range $1, \ldots, n$, where arithmetic is modulo 2. Since S is of the form $\pm X(\boldsymbol{a})$ or $\pm Z(\boldsymbol{a})$, and U is a product of $\mathrm{CNOT}(i,j)$, Z_j and X_j for i, j in the range $1, \ldots, n$, we immediately arrive at the Lemma result. □

We recall from Chapter 4 and Lemma 74 that the code projection protocol for an $[[n, k]]$ CSS

code $\mathcal{C}$ with $n > k$ and minimal set of stabilizer group generators $\{S_1, \ldots, S_{n-k}\}$ can be represented in terms of a completely CSS-preserving unitary $U_{\mathcal{C}}$ and codespace projector $P_{\mathcal{C}}$ as

$$\Phi_{\mathcal{C}}(\rho) \coloneqq \mathrm{Tr}_{k+1,\ldots,n}\left[U_{\mathcal{C}}^{\dagger} P_{\mathcal{C}}(\rho) P_{\mathcal{C}} U_{\mathcal{C}}\right] \text{ where } U_{\mathcal{C}}^{\dagger}(S_i)U_{\mathcal{C}} = \begin{cases} Z_{k+i} \text{ if } S_i \text{ is } Z\text{-type} \\ X_{k+i} \text{ if } S_i \text{ is } X\text{-type} \end{cases}$$
$$\text{and } P_{\mathcal{C}} \coloneqq \prod_{i=1}^{n-k} P(S_i) \tag{B.79}$$

With slight abuse of terminology and notation, we shall also refer to $U_{\mathcal{C}}$ as the encoding unitary of $\mathcal{C}$ throughout the rest of this section.

By conjugating $U_{\mathcal{C}}$ past each $P(S_i)$ in $P_{\mathcal{C}}$, we obtain the following alternative expression of $\Phi_{\mathcal{C}}$ in terms of the decoding unitary and the zero-syndrome projector $P_{\mathbf{0}}$ (up to Hadamard gates on syndrome qubits to transform $|+\rangle$ into $|0\rangle$ as needed).

$$\Phi_{\mathcal{C}}(\rho) \coloneqq \mathrm{Tr}_{k+1,\ldots,n}\left[P_{\mathbf{0}}\, \mathcal{U}_{\mathcal{C}}^{\dagger}(\rho) U_{\mathcal{C}} P_{\mathbf{0}}\right] \text{ where } P_{\mathbf{0}} \coloneqq \prod_{i=k+1}^{n} P(C_i)$$
$$\text{for } C_i \coloneqq \begin{cases} X_{k+i} & \text{if } S_i \text{ is } Z\text{-type} \\ Z_{k+i} & \text{if } S_i \text{ is } X\text{-type,} \end{cases} \tag{B.80}$$

which is equivalent to

$$\Phi_{\mathcal{C}}(\rho) \coloneqq \mathrm{Tr}_{k+1,\ldots,n}\left[\left(I \bigotimes_{i=k+1}^{n} |\phi_i\rangle\langle\phi_i|\right) \mathcal{U}_{\mathcal{C}}^{\dagger}(\rho) U_{\mathcal{C}} \left(I \bigotimes_{i=k+1}^{n} |\phi_i\rangle\langle\phi_i|\right)\right]$$
$$\text{where } |\phi_i\rangle \coloneqq \begin{cases} |+\rangle & \text{if } S_i \text{ is } X\text{-type} \\ |0\rangle & \text{if } S_i \text{ is } Z\text{-type,} \end{cases} \tag{B.81}$$

Now $U_{\mathcal{C}}$ generates the logical basis for $\mathcal{C}$ as

$$\forall \boldsymbol{s} \in \{0,1\}^k : |\boldsymbol{s}_{\mathcal{C}}\rangle \coloneqq U_{\mathcal{C}}\left(|\boldsymbol{s}\rangle \bigotimes_{i=k+1}^{n} |\phi_i\rangle\right) \tag{B.82}$$

We can therefore express $\Phi_{\mathcal{C}}$ as the Kraus decomposition

$$\Phi_{\mathcal{C}}(\rho) = K_{\mathcal{C}} \rho K_{\mathcal{C}}^{\dagger} \text{ for } K_{\mathcal{C}}^{\dagger} \coloneqq \sum_{\boldsymbol{s}\in\{0,1\}^k} |\boldsymbol{s}_{\mathcal{C}}\rangle\langle\boldsymbol{s}|\,. \tag{B.83}$$

When $k = 0$, $\Phi_{\mathcal{C}}$ simply projects onto a pure n-qubit CSS state and then discards it.

Step 1: Standard form for CSS magic distillation protocols

Lemma 85. *Any n-to-k CSS magic distillation protocol Φ can be decomposed as a sum of n-qubit operations*

$$\Phi(\rho) = \sum_i p_i \operatorname{Tr}_{k+1,\ldots,n+m} \left[K_i(\rho \otimes |\psi_i\rangle\langle\psi_i|)K_i^\dagger \right], \tag{B.84}$$

where p_i is a probability, $|\psi_i\rangle$ is a CSS state on m ancillary qubits and K_i is Kraus operator of the form

$$K_i = U_i \left(\prod_{l=1}^{N} P(S_{i,j}) \right), \tag{B.85}$$

where U_i is a completely CSS-preserving unitary and $P(S_{i,j})$ projects onto the $+1$ eigenspace of a CSS observable $S_{i,j}$.

Proof. By Lemma 69, we can decompose Φ as follows

$$\Phi(\rho) = \sum_i p_i' \operatorname{Tr}_{k+1,\ldots,n+m} \left[K_i(\rho \otimes \sigma_i)K_i^\dagger \right] \tag{B.86}$$

where p_i' is a probability, σ_i is a CSS state on m ancillary qubits and K_i is a Kraus operator of the form

$$K_i = \prod_{j=1}^{N} P(S_{i,j}')U_{i,j}', \tag{B.87}$$

where $U_{i,j}'$ is a completely CSS-preserving unitary and $P(S_{i,j}')$ projects onto the $+1$ eigenspace of the CSS observable $S_{i,j}'$. Given any completely CSS-preserving unitary U and CSS observable S', we have that $S'U = US$, where S is another CSS observable. Therefore, we can conjugate every completely CSS-preserving unitary $U_{i,j}'$ to the beginning of its Kraus operator K_i', where we compose them into a single completely CSS-preserving unitary U_i, leaving behind a string of projections onto the $+1$ eigenspaces of another set N CSS observables $\{S_{i,j}\}$ in their wake. Decomposing each CSS state σ_i on ancillary qubits into a statistical mixture of pure CSS states then yields the Lemma result. □

Step 2: Exposing the decoding

Lemma 86. *Consider a channel Φ on n qubits that performs a completely CSS-preserving unitary U and then discards the final $n-k$ qubits,*

$$\Phi(\rho) = \mathrm{Tr}_{k+1,\dots,n}\left[U(\rho)U^\dagger\right]. \tag{B.88}$$

where $1 \leq k < n$. Then Φ can be decomposed into a sum over CSS code projections $\{\Phi_{\mathcal{C}_{\boldsymbol{s}}}\}$ onto $[[n,k]]$ CSS codes $\{C_{\boldsymbol{s}}\}$ indexed by the computational basis $\{\boldsymbol{s}\}$ on the discarded qubits,

$$\Phi(\rho) = \sum_{\boldsymbol{s}\in\{0,1\}^{n-k}} \Phi_{\mathcal{C}_{\boldsymbol{s}}}(\rho). \tag{B.89}$$

Proof. The channel Φ is unchanged by performing a measurement in the computational basis $\{|\boldsymbol{s}\rangle\}$ of the final $n-k$ qubits before discarding them. Thus one can write

$$\Phi(\rho) = \sum_{\boldsymbol{s}\in\{0,1\}^{n-k}} \Phi_{\boldsymbol{s}}(\rho), \tag{B.90}$$

where we have defined

$$\Phi_{\boldsymbol{s}}(\rho) \coloneqq \mathrm{Tr}_{k+1,\dots,n}\left[\mathbb{1}^{\otimes k}\otimes|\boldsymbol{s}\rangle\langle\boldsymbol{s}|\,(U^\dagger(\rho)U)\,\mathbb{1}^{\otimes k}\otimes|\boldsymbol{s}\rangle\langle\boldsymbol{s}|\right]. \tag{B.91}$$

Let $U_{\boldsymbol{s}}$ be a completely CSS-preserving unitary on the final $n-k$ qubits that transforms $|\boldsymbol{s}\rangle$ into $|\mathbf{0}\rangle$, e.g. $U_{\boldsymbol{s}}^\dagger \coloneqq \bigotimes_{i=1}^{n-k}(X)^{s_i}$. By the cyclic property of the trace, we obtain

$$\Phi_{\boldsymbol{s}}(\rho) = \mathrm{Tr}_{k+1,\dots,n}\left[K(\rho)K^\dagger\right], \tag{B.92}$$

for the Kraus operator

$$K \coloneqq \mathbb{1}^{\otimes k}\otimes(U_{\boldsymbol{s}}\,|\boldsymbol{s}\rangle\langle\boldsymbol{s}|)\,U = \mathbb{1}^{\otimes k}\otimes|\mathbf{0}\rangle\langle\mathbf{0}|\left([\mathbb{1}^{\otimes k}\otimes U_{\boldsymbol{s}}]U\right). \tag{B.93}$$

By Eq. (B.81), we see that $\Phi_{\boldsymbol{s}}$ can be interpreted as a code projection protocol with a completely CSS-preserving decoding unitary $U_{\boldsymbol{s}}U$ for an $[[n,k]]$ CSS code $\mathcal{C}_{\boldsymbol{s}}$ stabilized by $\mathsf{S}(\mathcal{C}_{\boldsymbol{s}}) \coloneqq \langle (U_{\boldsymbol{s}}U)^\dagger Z_{k+1}(U_{\boldsymbol{s}}U), \dots, (U_{\boldsymbol{s}}U)^\dagger Z_n(U_{\boldsymbol{s}}U)\rangle$. □

Step 3: Removing the projections

Lemma 87. *Let Φ be an operation on n qubits that projectively measures a CSS observable S, post-selects the $+1$ outcome, and then carries out a code projection $\Phi_{\mathcal{C}}$ of an $[[n,k]]$ CSS code $\mathcal{C}$, i.e.*

$$\Phi(\rho) \coloneqq \Phi_{\mathcal{C}} \circ \Pi(S)[\rho] \text{ where } \Pi(S)[\rho] \coloneqq P(S)\rho P(S). \tag{B.94}$$

where $n > k$. There are three possibilities for how Φ transforms ρ:

1. *that $\Phi(\rho) = 0$ for all ρ.*
2. *that Φ is a code projection $\Phi_{\mathcal{C}'}$ for another $[[n,k]]$ CSS code $\mathcal{C}'$, followed by completely CSS-preserving unitary post-processing $\mathcal{U}$. Thus one can write*

$$\Phi(\rho) = p\,\mathcal{U} \circ \Phi_{\mathcal{C}'}(\rho), \tag{B.95}$$

where $p > 0$ is a probability. One can further find logical bases for the new and old codes respectively, $\left\{|\boldsymbol{s}_{\mathcal{C}'}\rangle \,|\boldsymbol{s} \in \{0,1\}^k\right\}$ and $\left\{|\boldsymbol{s}_{\mathcal{C}}\rangle \,|\boldsymbol{s} \in \{0,1\}^k\right\}$, such that

$$P(S)\,|\boldsymbol{s}_{\mathcal{C}}\rangle \propto |\boldsymbol{s}_{\mathcal{C}'}\rangle\,. \tag{B.96}$$

3. *for $k \geq 1$, that Φ is a code projection $\mathcal{C}'$ for an $[[n,k-1]]$ CSS code $\mathcal{C}'$, followed by preparing a CSS state $|\varphi\rangle$ on a single qubit and completely CSS-preserving unitary post-processing $\mathcal{U}'$, i.e.*

$$\Phi(\rho) = \mathcal{U}' \circ [\Phi_{\mathcal{C}'}(\rho) \otimes |\varphi\rangle\langle\varphi|]. \tag{B.97}$$

Furthermore, we can find logical bases $\left\{\left|\boldsymbol{s}'_{\mathcal{C}'}\right\rangle \,|\boldsymbol{s}' \in \{0,1\}^{k-1}\right\}$ and $\left\{|\boldsymbol{s}_{\mathcal{C}}\rangle \,|\boldsymbol{s} \in \{0,1\}^k\right\}$ such that

$$P(S)\left|f(\boldsymbol{s}')_{\mathcal{C}}\right\rangle \propto \left|\boldsymbol{s}'_{\mathcal{C}'}\right\rangle, \tag{B.98}$$

for some function $f : \{0,1\}^{k-1} \to \{0,1\}^k$.

Proof. Let $\{S_{k+1}, \ldots S_n\}$ be CSS observables that constitute a minimal set of generators for the stabilizer group $\mathsf{S}(\mathcal{C})$ of $\mathcal{C}$. We now demonstrate how to manipulate Φ into one of the forms stated by the Lemma depending on the relationship between S and the CSS observables generating $\mathsf{S}(\mathcal{C})$.

(i) S **does *not* commute with at least one generator of** $\mathsf{S}(\mathcal{C})$**.** We assume without loss of generality that S does not commute with S_{k+1}. We now show how Φ can be manipulated into the *second* form stated by the Lemma.

By Eq. (B.83), we can express Φ as the following Kraus decomposition in terms of the logical basis $\{|\boldsymbol{s}_{\mathcal{C}}\rangle\}$ generated by the encoding unitary $U_{\mathcal{C}}$ of $\mathcal{C}$ as

$$\Phi(\rho) = K(\rho)K^\dagger \text{ for } K^\dagger := \sum_{\boldsymbol{s}\in\{0,1\}^k} P(S)\,|\boldsymbol{s}_{\mathcal{C}}\rangle\langle\boldsymbol{s}|\,. \tag{B.99}$$

The stabilizer group $\mathsf{S}(|\boldsymbol{s}_{\mathcal{C}}\rangle)$ for each logical basis state $|\boldsymbol{s}_{\mathcal{C}}\rangle$ can be expressed as

$$\mathsf{S}(|\boldsymbol{s}_{\mathcal{C}}\rangle) = \langle (-1)^{s_1}U_{\mathcal{C}}(Z_1)U_{\mathcal{C}}^\dagger, \ldots, (-1)^{s_k}U_{\mathcal{C}}(Z_k)U_{\mathcal{C}}^\dagger, S_{k+1}, \ldots, S_n\rangle, \tag{B.100}$$

We can multiply all generators of $\mathsf{S}(|\boldsymbol{s}_{\mathcal{C}}\rangle)$ that do *not* commute with S by S_{k+1} and, as all CSS observables that do not commute with S must be of the same type, arrive at another set of commuting, independent and non-trivial CSS observables $\{S_1, \ldots, S_n\}$ such that

$$\mathsf{S}(|\boldsymbol{s}_{\mathcal{C}}\rangle) = \langle (-1)^{s_1}S_1, \ldots, (-1)^{s_k}S_k, S_{k+1}, \ldots, S_n\rangle. \tag{B.101}$$

This implies (see proof of Lemma 68)

$$P(S)\,|\boldsymbol{s}_{\mathcal{C}}\rangle = \frac{1}{\sqrt{2}}\,|\psi_{\boldsymbol{s}}\rangle \tag{B.102}$$

for a CSS state $|\psi_{\boldsymbol{s}}\rangle$ stabilized by

$$\mathsf{S}(|\psi_{\boldsymbol{s}}\rangle) = \langle (-1)^{s_1}S_1, \ldots, (-1)^{s_k}S_k, S, S_{k+2} \ldots, S_n\rangle. \tag{B.103}$$

Applying Lemma 74 to $S_1, \ldots, S_k, S, S_{k+2}, \ldots, S_n$, we can find a completely CSS-preserving encoding unitary $U_{\mathcal{C}'}$ for the $[[n, k]]$ CSS code $\mathcal{C}'$ stabilized by $\langle S, S_{k+2}, \ldots, S_n\rangle$ that generates a logical basis $\{|\boldsymbol{s}_{\mathcal{C}'}\rangle\}$ in $\mathcal{C}'$ such that $|\boldsymbol{s}_{\mathcal{C}'}\rangle$ shares the stabilizer group of $|\psi_{\boldsymbol{s}}\rangle$. Therefore, $|\boldsymbol{s}_{\mathcal{C}'}\rangle$ and $|\psi_{\boldsymbol{s}}\rangle$ only differ up to a phase, and one can write

$$P(S)\,|\boldsymbol{s}_{\mathcal{C}}\rangle = \frac{1}{\sqrt{2}}e^{-i\theta_{\boldsymbol{s}}}\,|\boldsymbol{s}_{\mathcal{C}'}\rangle\,. \tag{B.104}$$

Substituting into Eq. (B.99), we obtain

$$\Phi(\rho) = \frac{1}{2} U \left[\Phi_{\mathcal{C}'}(\rho)\right] U^\dagger, \tag{B.105}$$

where we have defined the following unitary on k qubits to adjust for the phase differences between $|\psi_{\boldsymbol{s}}\rangle$ and $|\boldsymbol{s}_{\mathcal{C}'}\rangle$,

$$U \coloneqq \sum_{\boldsymbol{s}\in\{0,1\}^k} e^{i\theta_{\boldsymbol{s}}} |\boldsymbol{s}\rangle\langle\boldsymbol{s}| \,. \tag{B.106}$$

We see that Eq. (B.105) now matches the second statement of the Lemma provided that U is completely CSS-preserving, which we now proceed to show.

Consider an arbitrary pure CSS state $|\psi\rangle$ on k qubits. Since $U_{\mathcal{C}'}$ is completely CSS-preserving, it encodes $|\psi\rangle$ as another CSS state $|\psi_{\mathcal{C}'}\rangle$ in $\mathcal{C}'$. The initial form of Φ in Eq. (B.94) shows that it is completely CSS preserving, which implies $\Phi(|\psi_{\mathcal{C}'}\rangle\langle\psi_{\mathcal{C}'}|)$ is a (possibly subnormalised) CSS state. On the other hand, the form of Φ in Eq. (B.105) implies $\Phi(|\psi_{\mathcal{C}'}\rangle\langle\psi_{\mathcal{C}'}|) = \frac{1}{2} U |\psi\rangle\langle\psi| U^\dagger$, and therefore $U|\psi\rangle\langle\psi|$ must be CSS as well. We conclude that if $|\psi\rangle$ is CSS, then so is $U|\psi\rangle$, which implies that U is CSS-preserving.

From the proof of Lemma 67, we can express U as $U = \left[H^{\otimes k}\right]^a V$, where V is a *completely* CSS-preserving unitary and a is a binary digit. While the CNOT-gate and single-qubit X- and Z-gates map computational basis states onto computational basis states, the collective Hadamard gate does not. As U is diagonal in the computational basis by construction, we conclude that $a = 0$, so U is indeed completely CSS-preserving.

(ii) **S commutes with $S_{k+1}, \ldots, S_n$, and both S and $-S$ are independent of them.** This is only possible when $k \geq 1$. In this case, we show that Φ can be manipulated into the *third* form in the Lemma.

By Eq. (B.80), we can express Φ as the following once we conjugate $P(S)$ past U in accordance with Lemma 84:

$$\Phi(\rho) = \mathrm{Tr}_{k+1,\ldots,n}\left[P_{\boldsymbol{0}} P(S') U_{\mathcal{C}}^\dagger(\rho) U_{\mathcal{C}} P(S') P_{\boldsymbol{0}}\right] \tag{B.107}$$

where $S' \coloneqq U_{\mathcal{C}}^\dagger S U_{\mathcal{C}}$ is another CSS observable, and $P_{\boldsymbol{0}}$ is the zero-syndrome projector of

$\mathcal{C}$. Consider the term

$$P(S')P_{\mathbf{0}} = P(S')P(C_{k+1})\dots P(C_n), \text{ where } C_i \coloneqq \begin{cases} X_i & \text{if } S_i \text{ is } X\text{-type,} \\ Z_i & \text{if } S_i \text{ is } Z\text{-type.} \end{cases} \tag{B.108}$$

from Eq. (B.107). Given any two Pauli observables O_2 and O_2, we have that [82]

$$P(O_1)P(O_2) = P(O_1O_2)P(O_2). \tag{B.109}$$

Because S' must commute with $C_{n-k}, \dots C_n$, we conclude that S' can be non-trivial on a qubit i where $k+1 \leq i \leq n$ if and only if C_i is the same type of Pauli observable as S', which implies, since C_i must be non-trivial, that S' is non-trivial on a qubit i if and only if $S'C_i$ is trivial on qubit i. We can therefore apply Eq. (B.109) to S' and the C_i in Eq. (B.108) to eliminate the non-trivial part of $P(S')$ on the last $n-k$ qubits, and conclude that there exists a non-trivial CSS observable S'' on the first k qubits *alone* such that

$$P(S')P_{\mathbf{0}} = [P(S'') \otimes \mathbb{1}^{\otimes(n-k)}]P_{\mathbf{0}}. \tag{B.110}$$

Since both S' and $-S'$ must be independent of $C_k + 1, \dots, C_n$, we conclude that S'' is not proportional to the identity.

We now consider the cases where S'' is Z-type and X-type separately.

(a) S'' **is Z-type.** We therefore represent S'' as $S'' = (-1)^b Z(\boldsymbol{z})$, where b is a binary digit and $\boldsymbol{z}$ is an k-bit string of which at least one bit j is such that $z_j = 1$.

Consider the following completely CSS-preserving unitary on k qubits

$$\tilde{U}^\dagger \coloneqq \mathrm{SWAP}(j,k) \circ X_j^b \circ U_Z \text{ where } U_Z \coloneqq \prod_{i=1}^{k} [\mathrm{CNOT}(i,j)]^{z_i}, \tag{B.111}$$

where $\mathrm{SWAP}(j,k) \coloneqq \mathrm{CNOT}(i,j)\,\mathrm{CNOT}(j,i)\,\mathrm{CNOT}(i,j)$ swaps qubits i and j, and U_Z has been designed via Eq. (B.1) to "clear" the non-trivial parts of S'' on every qubit *except* the jth, which implies $U_C[Z(\boldsymbol{z})]U_C^\dagger = Z_j$. Therefore,

$$\tilde{U}^\dagger P(S'')\tilde{U} = P(Z_k), \tag{B.112}$$

Using Eq. (B.112) to substitute $P(S'')$ in Eq. (B.110), we have by Eq. (B.107) that

$$\Phi(\rho) = \tilde{U}\left(\mathrm{Tr}_{k+1,\ldots,n}\left[P'U'^{\dagger}(\rho)U'P'\right]\right)\tilde{U}^{\dagger}, \tag{B.113}$$

where we have defined

$$P' \coloneqq P(Z_k)P_{\mathbf{0}} \text{ and } U' \coloneqq U_{\mathcal{C}}\tilde{U} \otimes \mathbb{1}^{\otimes(n-k)}. \tag{B.114}$$

As kth qubit outputted after the projection P' is always $|0\rangle$, we can simply discard the kth qubit as well and re-prepare it. Thus one can write

$$\Phi(\rho) = \tilde{U}\left(\Phi_{\mathcal{C}'}(\rho) \otimes |0\rangle\langle 0|\right)\tilde{U}^{\dagger}, \tag{B.115}$$

where we have defined

$$\Phi_{\mathcal{C}'}(\rho) \coloneqq \mathrm{Tr}_{k,\ldots,n}\left[P'U'^{\dagger}(\rho)U'P'\right]. \tag{B.116}$$

By Eq. (B.81), we see that $\Phi_{\mathcal{C}'}$ is the code projection of an $[[n, k-1]]$ CSS code $\mathcal{C}'$ stabilized by $\langle U'Z_kU'^{\dagger}, U'C_{k+1}U'^{\dagger}, \ldots, U'C_nU'^{\dagger}\rangle$ with a completely CSS-preserving encoding unitary U'.

By Eq. (B.82), U' generates the following logical basis for $\mathcal{C}'$:

$$\begin{aligned}
&\forall \boldsymbol{s}' \in \{0,1\}^{k-1} : \left|\boldsymbol{s}'_{\mathcal{C}}\right\rangle \coloneqq U' \left|\boldsymbol{s}\right\rangle |0\rangle |\Phi\rangle \\
&\qquad \text{where } |\Phi\rangle \coloneqq \bigotimes_{i=1}^{n-k} |\phi_i\rangle \text{ for } |\phi_i\rangle \coloneqq \begin{cases} |+\rangle & \text{if } S_i \text{ is } X\text{-type.} \\ |0\rangle & \text{if } S_i \text{ is } Z\text{-type.} \end{cases}
\end{aligned} \tag{B.117}$$

Noting that $P(Z_k)P_{\mathbf{0}}$ is the zero-syndrome projectors of $\mathcal{C}'$, which has the zero syndrome state $|0\rangle|\Phi\rangle$ on its $(n-k-1)$ syndrome qubits, we can relate each logical

basis state $\left|\boldsymbol{s}'_{\mathcal{C}'}\right\rangle$ to a (unique) logical basis state in the old code $\mathcal{C}$ as

$$\begin{aligned}
\left|\boldsymbol{s}'_{\mathcal{C}'}\right\rangle &= \left[U_{\mathcal{C}}\tilde{U}\otimes\mathbb{1}^{\otimes(n-k)}\right]\left|\boldsymbol{s}'\right\rangle|0\rangle|\Phi\rangle \\
&= \left[U_{\mathcal{C}}\tilde{U}\otimes\mathbb{1}^{\otimes(n-k)}\right]P(Z_k)P_{\mathbf{0}}\left|\boldsymbol{s}'\right\rangle|0\rangle|\Phi\rangle \\
&= U_{\mathcal{C}}P\left(S''\right)\otimes\mathbb{1}^{\otimes(n-k)}P_{\mathbf{0}}\left[\left(\tilde{U}\left|\boldsymbol{s}'\right\rangle|0\rangle\right)|\Phi\rangle\right] \\
&= U_{\mathcal{C}}P(S')P_{\mathbf{0}}\left|f(\boldsymbol{s}')\right\rangle|\Phi\rangle \\
&= P(S)U_{\mathcal{C}}\left|f(\boldsymbol{s}')\right\rangle|\Phi\rangle \\
&= P(S)\left|f(\boldsymbol{s}')_{\mathcal{C}}\right\rangle, \qquad \text{(B.118)}
\end{aligned}$$

where we have defined $|f(\boldsymbol{s}')\rangle \coloneqq \tilde{U}\left|\boldsymbol{s}'\right\rangle|0\rangle$, used Eq. (B.112) for the second equality, used Eq. (B.110) for the third equality, and the fact that $P_{\mathbf{0}}$ is the zero-syndrome projector of $\mathcal{C}$, whose zero-syndrome state is $|\Phi\rangle$ on $(n-k)$ syndrome qubits, for the fourth equality. Explicitly, f evaluates to

$$\begin{aligned}
f(\boldsymbol{s}') =& (s'_1,\ldots,s'_{j-1},s,s'_j,\ldots,s'_{k-1}), \text{where} \\
& s \coloneqq \left(\sum_{i=1}^{j-1}s'_i z_i\right)+b+\left(\sum_{i=j+1}^{k}s'_{i-1}z_i\right), \qquad \text{(B.119)}
\end{aligned}$$

and arithmetic is modulo 2.

(b) S'' **is** X**-type.** We therefore represent S'' as $S'' = (-1)^b X(\boldsymbol{x})$, where b is a binary digit and $\boldsymbol{x}$ is an k-bit string of which at least one bit j is such that $x_j = 1$.

Consider the following completely CSS-preserving unitary on k qubits,

$$\tilde{U}^\dagger \coloneqq \mathrm{MV}_{j\to k}\circ Z_j^b\circ U_C, \qquad \text{(B.120)}$$

in which U_C is defined as

$$U_C \coloneqq \begin{cases} \prod\limits_{\substack{i=1\\ x_i=1, i\neq j}}^{k} \mathrm{CNOT}(j,i) & \text{if } \exists i\neq j : x_i = 1, \\ \mathbb{1}^{\otimes k} & \text{otherwise.} \end{cases} \qquad \text{(B.121)}$$

Because $U_C[X(\boldsymbol{x})]U_C^\dagger = X_j$, we have that

$$\tilde{U}^\dagger P(S'')\tilde{U} = P(X_k). \tag{B.122}$$

Using Eq. (B.122) to substitute $P(S'')$ in Eq. (B.110), we have by Eq. (B.107) that

$$\Phi(\rho) = \tilde{U}\left(\mathrm{Tr}_{k+1,\ldots,n}\left[P'U'^\dagger(\rho)U'P'\right]\right)\tilde{U}^\dagger, \tag{B.123}$$

where we have defined

$$P' \coloneqq P(X_k)P_{\mathbf{0}},\ U' \coloneqq U_{\mathcal{C}}\tilde{U} \otimes \mathbb{1}^{\otimes(n-k)}. \tag{B.124}$$

Since kth qubit outputted by the part inside $\tilde{U}(\cdot)\tilde{U}^\dagger$ is always $|+\rangle$, we can simply discard the kth qubit as well and re-prepare it. Thus one can write

$$\mathcal{E}(\rho) = \tilde{U}\left(\mathcal{K}'(\rho) \otimes |+\rangle\langle+|\right)\tilde{U}^\dagger, \tag{B.125}$$

where we have defined

$$\mathcal{K}'(\rho) \coloneqq \mathrm{Tr}_{k,\ldots,n}\left[P'U'^\dagger(\rho)U'P'\right]. \tag{B.126}$$

From Eq. (B.79), we see that $\mathcal{K}'$ is a code projection for an $[[n, k-1]]$ CSS code $\mathcal{C}'$ stabilized by $\langle U'X_kU'^\dagger, U'C_{k+1}U'^\dagger, \ldots, U'C_nU'^\dagger\rangle$ with a completely CSS-preserving encoding unitary U'. By Eq. (B.82), U' generates the following logical basis for $\mathcal{C}'$:

$$\forall \boldsymbol{s}' \in \{0,1\}^{k-1} : \left|\boldsymbol{s}'_{\mathcal{C}}\right\rangle \coloneqq U'\left|\boldsymbol{s}\right\rangle|+\rangle|\Phi\rangle \tag{B.127}$$

$$\text{where } |\Phi\rangle \coloneqq \bigotimes_{i=1}^{n-k}|\phi_i\rangle. \tag{B.128}$$

We can then relate the logical basis state $\left|\boldsymbol{s}'_{\mathcal{C}'}\right\rangle$ in the new code $\mathcal{C}'$ to a logical basis

state in the old code $\mathcal{C}$ as

$$\begin{aligned}
\left|\boldsymbol{s}'_{\mathcal{C}'}\right\rangle &= \left[U_{\mathcal{C}}\tilde{U}\otimes\mathbb{1}^{(n-k)}\right]P(X_k)P_{\mathbf{0}}\left|\boldsymbol{s}'\right\rangle|+\rangle|\Phi\rangle \\
&= \sqrt{2}\left[U_{\mathcal{C}}\tilde{U}\otimes\mathbb{1}^{\otimes(n-k)}\right]P(X_k)P_{\mathbf{0}}\left|\boldsymbol{s}'\right\rangle|0\rangle|\Phi\rangle \\
&= \sqrt{2}\,U_{\mathcal{C}}P(S'')\otimes\mathbb{1}^{\otimes(n-k)}P_{\mathbf{0}}\left(\tilde{U}\left|\boldsymbol{s}'\right\rangle|0\rangle\right)|\Phi\rangle \\
&= \sqrt{2}\,U_{\mathcal{C}}P(S')P_{\mathbf{0}}\left|f(\boldsymbol{s}')\right\rangle|\Phi\rangle \\
&= \sqrt{2}\,P(S)U_{\mathcal{C}}\left|f(\boldsymbol{s}')\right\rangle|\Phi\rangle \\
&= \sqrt{2}\,P(S)\left|f(\boldsymbol{s}')_{\mathcal{C}}\right\rangle, \qquad \text{(B.129)}
\end{aligned}$$

where we defined $|f(\boldsymbol{s}')\rangle \coloneqq \tilde{U}\,|\boldsymbol{s}'\rangle\,|0\rangle$, used Eq. (B.122) to obtain the second equality and Eq. (B.110) to obtain the third. Explicitly, $f(\boldsymbol{s}') = (s'_1,\ldots,s'_{j-1},0,s'_j,\ldots,s'_{k-1})$.

(iii) **Either S or $-S$ is not independent of $S_{k+1},\ldots,S_n$.** This implies either $-S$ or S stabilizes $\mathcal{C}$. In the former case, we see from Eq. (B.99) that $\Phi(\rho) = 0$ for all ρ. In the latter case, we have $P(S)\,|\boldsymbol{s}_{\mathcal{C}}\rangle = |\boldsymbol{s}_{\mathcal{C}}\rangle$, which implies $\Phi = \Phi_{\mathcal{C}}$ by Eq. (B.99). Together these equations match the Lemma's second form. □

Step 4. Removing ancillary qubits

Lemma 88. *Let $\mathcal{C}$ be an $[[n+m,k]]$ CSS codes where $n \geq 1, n > k$ and $m > 0$, and let $\{|\boldsymbol{s}_{\mathcal{C}}\rangle\}$ be a logical basis for $\mathcal{C}$ generated by a completely CSS-preserving encoding unitary U such that each logical basis state $|\boldsymbol{s}_{\mathcal{C}}\rangle$ factorises over n and m qubits, i.e.,*

$$|\boldsymbol{s}_{\mathcal{C}}\rangle = |\psi_{\boldsymbol{s}}\rangle\otimes|\psi\rangle, \qquad \text{(B.130)}$$

where $|\psi\rangle$ is a CSS state on m qubits. We then have that

$$|\psi_{\boldsymbol{s}}\rangle = e^{-i\theta_{\boldsymbol{s}}}\,|\boldsymbol{s}_{\mathcal{C}'}\rangle, \qquad \text{(B.131)}$$

where $\{|\boldsymbol{s}_{\mathcal{C}'}\rangle\}$ is a logical basis for an $[[n,k]]$ CSS code $\mathcal{C}'$ generated by a completely CSS-preserving unitary.

Proof. The stabilizer group for the logical basis state $|\boldsymbol{s}_{\mathcal{C}}\rangle$ can be related to the stabilizer group

$\mathsf{S}(\mathcal{C})$ of $\mathcal{C}$ as

$$\mathsf{S}(|\boldsymbol{s}_{\mathcal{C}}\rangle) = \langle (-1)^{s_1} S_1, \ldots, (-1)^{s_k} S_k \rangle \times \mathsf{S}(\mathcal{C}), \tag{B.132}$$

where we have defined $S_i \coloneqq U Z_i U^\dagger$ for $i = 1, \ldots, k$.

We observe that $\mathsf{S}(\mathcal{C})$ is a direct sum of $\mathbb{F}_2$-subspaces $\mathsf{S}_X(\mathcal{C})$ and $\mathsf{S}_Z(\mathcal{C})$ corresponding to X- and Z-type stabilizers for $\mathcal{C}$. Because $|\psi\rangle$ is a CSS state, its stabilizer group $\mathsf{S}(|\psi\rangle)$ is similarly a direct sum of $\mathbb{F}_2$-subspaces $\mathsf{S}_Z(|\psi\rangle)$ and $\mathsf{S}_X(|\psi\rangle)$ corresponding to X- and Z-type stabilizers for $|\psi\rangle$.

Since $\{|\boldsymbol{s}_{\mathcal{C}}\rangle\}$ span $\mathcal{C}$, Eq. (B.130) implies that, if S stabilizes $|\psi\rangle$, then $\mathbb{1}^{\otimes n} \otimes S$ stabilizes $\mathcal{C}$. Therefore, $\mathbb{1}^{\otimes n} \otimes \mathsf{S}_Z(|\psi\rangle)$ and $\mathbb{1}^{\otimes n} \otimes \mathsf{S}_X(|\psi\rangle)$ are $\mathbb{F}_2$-subspaces of $\mathsf{S}_Z(\mathcal{C})$ and $\mathsf{S}_X(\mathcal{C})$ respectively. By applying the basis extension theorem separately to the Z and X cases, we can represent $\mathsf{S}(\mathcal{C})$ as

$$\mathsf{S}(\mathcal{C}) = \langle S_{k+1}, \ldots, S_n, \mathbb{1}^{\otimes n} \otimes T_1, \ldots, \mathbb{1}^{\otimes n} \otimes T_m \rangle, \tag{B.133}$$

where $S_{k+1}, \ldots, S_n$ and $T_1, \ldots, T_m$ are all CSS observables such that $\mathsf{S}(|\psi\rangle) = \langle T_1, \ldots, T_m \rangle$.

The stabilizer group of $|\boldsymbol{s}_{\mathcal{C}}\rangle$ can then be represented as

$$\mathsf{S}(|\boldsymbol{s}_{\mathcal{C}}\rangle) = \langle (-1)^{s_1} S_1, \ldots, (-1)^{s_k} S_k, S_{k+1}, \ldots, S_n, \mathbb{1}^{\otimes n} \otimes T_1, \ldots, \mathbb{1}^{\otimes n} \otimes T_m \rangle, \tag{B.134}$$

from which we note that $S_1, \ldots, S_n$ are all members of $\mathsf{S}(|\mathbf{0}_{\mathcal{C}}\rangle)$, the stabilizer group for the zero logical basis state in $\mathcal{C}$. Since $|\mathbf{0}_{\mathcal{C}}\rangle$ and $|\psi\rangle$ are CSS, $|\psi_{\mathbf{0}}\rangle$ must be CSS as well (as discarding the last m qubits is CSS-preserving). Therefore, $\mathsf{S}(|\mathbf{0}_{\mathcal{C}}\rangle)$ is a direct sum of $\mathbb{F}_2$-subspaces $\mathbb{1}^{\otimes n} \otimes \mathsf{S}(|\psi\rangle)$ and $\mathsf{S}(|\psi_{\mathbf{0}}\rangle) \otimes \mathbb{1}^{\otimes m}$, where $\mathsf{S}(|\psi_{\mathbf{0}}\rangle)$ is the stabilizer group of $|\psi_{\mathbf{0}}\rangle$. Consequently, we can express S_i as $S_i = S_i' \otimes T_i'$, where S_i' and T_i' are CSS observables of the same type as S_i that respectively stabilize $|\psi_{\mathbf{0}}\rangle$ and $|\psi\rangle$. Therefore, by multiplying each S_i in Eq. (B.134) by appropriate generators of the same type in $\{\mathbb{1}^{\otimes} \otimes T_1, \ldots \mathbb{1}^{\otimes n} \otimes T_m\}$, we can represent the stabilizer group of $|\boldsymbol{s}_{\mathcal{C}}\rangle$ as

$$\mathsf{S}(|\boldsymbol{s}_{\mathcal{C}}\rangle) = \mathsf{S}(|\psi_{\boldsymbol{s}}\rangle) \otimes \mathsf{S}(|\psi\rangle), \tag{B.135}$$

where the stabilizer group $\mathsf{S}(|\psi_{\boldsymbol{s}}\rangle)$ is generated by

$$\mathsf{S}(|\psi_{\boldsymbol{s}}\rangle) = \langle (-1)^{s_1} S'_1, \ldots, (-1)^{s_k} S'_k, S'_{k+1}, \ldots, S'_n \rangle, \tag{B.136}$$

in which $S'_1, \ldots, S'_k$ are Z-type because $S_1, \ldots, S_k$ are Z-type. By applying Lemma 74 to $S'_1, \ldots, S'_n$, we can find a logical basis $\{|\boldsymbol{s}_{\mathcal{C}'}\rangle\}$ for an $[[n, k]]$ CSS code $\mathcal{C}'$ stabilized by $\langle S'_{k+1}, \ldots, S'_n \rangle$, generated by a completely CSS-preserving encoding unitary, such that $|\boldsymbol{s}_{\mathcal{C}'}\rangle$ shares the stabilizer group of $|\psi_{\boldsymbol{s}}\rangle$. Therefore, $|\boldsymbol{s}_{\mathcal{C}'}\rangle$ and $|\psi_{\boldsymbol{s}}\rangle$ only differ up to a phase, which implies the Lemma. □

Lemma 89. *Let $\Phi_{\mathcal{C}}$ be the code projection for an $[[n+m, k]]$ CSS code $\mathcal{C}$ where $n \geq 1, n > k$ and $m > 0$. Then given any m-qubit CSS state $|\psi\rangle$, we have that $\Phi_{\mathcal{C}}([\cdot] \otimes |\psi\rangle\langle\psi|)$ is equivalent to a CSS code projection on n qubits alone, followed by preparing a CSS state and completely CSS-preserving post-processing, i.e.*

$$\Phi_{\mathcal{C}}(\rho \otimes |\psi\rangle\langle\psi|) = p\, \tilde{\mathcal{U}} \circ \left(\tilde{\Phi}_{\mathcal{C}}(\rho) \otimes |\varphi\rangle\langle\varphi| \right), \tag{B.137}$$

where p is a probability, $\tilde{\mathcal{U}}$ is a completely CSS-preserving unitary channel on k qubits, $\tilde{\Phi}_{\mathcal{C}}$ is a code projection for an $[[n, k']]$ CSS code where $0 \leq k' \leq k$, and $|\varphi\rangle$ is a CSS state on $k - k'$ qubits.

Proof. Let $\{S_{n+1}, \ldots, S_{n+m}\}$ be a set of CSS observables that generate the stabilizer group defining $|\psi\rangle$. Then $\Phi_{\mathcal{C}}(\rho \otimes |\psi\rangle\langle\psi|)$ is equivalent to

$$\Phi_{\mathcal{C}}(\rho \otimes |\psi\rangle\langle\psi|) = \Phi_{\mathcal{C}}(\mathbf{P}[\rho \otimes |\psi\rangle\langle\psi|]\mathbf{P}), \tag{B.138}$$

where $\mathbf{P}$ projects the last m qubits onto $|\psi\rangle$, i.e.

$$\mathbf{P} \coloneqq \mathbb{1}^{\otimes n} \otimes |\psi\rangle\langle\psi| = \prod_{i=n+1}^{n+m} P(\mathbb{1}^{\otimes n} \otimes S_i). \tag{B.139}$$

By applying Lemma 87 to each projection carried out by $\mathbf{P}$, we obtain

$$\Phi_{\mathcal{C}}(\rho \otimes |\psi\rangle\langle\psi|) = p\, \mathcal{U} \circ \Phi'_{\mathcal{C}}(\rho \otimes |\psi\rangle\langle\psi|) \otimes |\varphi\rangle\langle\varphi|, \tag{B.140}$$

where p is a probability, $\mathcal{U}$ is a completely CSS-preserving unitary channel on k qubits, $\Phi_{\mathcal{C}'}$ is a

code projection for an $[[n+m, k']]$ CSS code $\mathcal{C}'$ where $0 \leq k' \leq k$, and $|\varphi\rangle$ is a CSS state on $k-k'$ qubits. Lemma 87 further implies there exists a logical basis $\{|\boldsymbol{s}_{\mathcal{C}'}\rangle \,|\boldsymbol{s} \in \{0,1\}^{k'}\}$ for the new code $\mathcal{C}'$, generated by a completely CSS-preserving encoding unitary, which can be related to an orthonormal set of states $\{|\Psi_{\boldsymbol{s}}\rangle \,|\boldsymbol{s} \in \{0,1\}^{k'}\}$ on $n+m$ qubits as

$$\mathbf{P}\,|\Psi_{\boldsymbol{s}}\rangle = \left(\mathbb{1}^{\otimes n} \otimes |\psi\rangle\langle\psi|\right)|\Psi_{\boldsymbol{s}}\rangle \propto |\boldsymbol{s}_{\mathcal{C}'}\rangle\,. \tag{B.141}$$

Eq. (B.141) immediately implies

$$|\boldsymbol{s}_{\mathcal{C}'}\rangle = |\psi_{\boldsymbol{s}}\rangle \otimes |\psi\rangle\,, \tag{B.142}$$

where $|\psi_{\boldsymbol{s}}\rangle$ is a state on the first n qubits *alone*. By Lemma 88, each $|\psi_{\boldsymbol{s}}\rangle$ is, *up to a phase that may vary with* $\boldsymbol{s}$, a logical basis state $|\boldsymbol{s}_{\mathcal{C}''}\rangle$ for an $[[n,k']]$ CSS code $\mathcal{C}''$ with only n physical qubits. Thus one can write

$$|\boldsymbol{s}_{\mathcal{C}'}\rangle = e^{-i\theta_{\boldsymbol{s}}}\,|\boldsymbol{s}_{\mathcal{C}''}\rangle \otimes |\psi\rangle\,. \tag{B.143}$$

By Eq. (B.83), we can express the code projection $\Phi_{\mathcal{C}'}$ for $\mathcal{C}'$ in terms of the logical basis $\{|\boldsymbol{s}_{\mathcal{C}'}\rangle\}$ as

$$\Phi_{\mathcal{C}'}(\cdot) = K(\cdot)K^\dagger,\; K^\dagger \coloneqq \sum_{\boldsymbol{s}\in\{0,1\}^{k'}} |\boldsymbol{s}_{\mathcal{C}'}\rangle\,\langle\boldsymbol{s}|\,. \tag{B.144}$$

We can then use Eq. (B.143) to show that

$$\Phi_{\mathcal{C}'}(\rho \otimes |\psi\rangle\langle\psi|) = U'\left[\tilde{\Phi}_{\mathcal{C}}(\rho)\right]U'^\dagger, \tag{B.145}$$

where we have defined the following unitary on k' qubits to adjust for the phase differences between $|\boldsymbol{s}_{\mathcal{C}'}\rangle$ and $|\psi_{\boldsymbol{s}}\rangle$,

$$U' \coloneqq \sum_{\boldsymbol{s}\in\{0,1\}^{k'}} e^{i\theta_{\boldsymbol{s}}}\,|\boldsymbol{s}\rangle\langle\boldsymbol{s}|\,, \tag{B.146}$$

as well as

$$\tilde{\Phi}_{\mathcal{C}}(\cdot) \coloneqq \tilde{K}(\cdot)\tilde{K}^{\dagger},\ \tilde{K}^{\dagger} \coloneqq \sum_{\boldsymbol{s}\in\{0,1\}^{k'}} |\boldsymbol{s}_{\mathcal{C}''}\rangle \langle \boldsymbol{s}|\,, \tag{B.147}$$

By comparison with Eq. (B.83), we see that $\tilde{\Phi}_{\mathcal{C}}$ is the code projection for the $[[n, k']]$ CSS code $\mathcal{C}''$. Following a similar argument to that at the end of case (iii) in the proof of Lemma 87, we can demonstrate that U' is completely CSS-preserving. Substituting back immediately yields the Lemma result. □

Appendix C

Appendices to Chapter 5

C.1 Properties of the conditional min-entropy

We begin by reviewing some useful properties for the *exponentiated single-shot conditional min-entropy* function $F : \mathrm{Herm}(RA) \to \mathbb{R}$, which is simply defined as

$$F(\Omega_{RA}) \coloneqq 2^{-H_{\min}(R|A)_\Omega} = \inf_{X_A \geq 0} \{\mathrm{Tr}[X_A] : I_R \otimes X_A \geq \Omega_{RA}\}. \tag{C.1}$$

on every Hermitian operator Ω_{RA} of the joint system RA.

Given any Hermitian operators Ω_{RA} and $\Omega'_{RA} \in \mathrm{Herm}(RA)$, Ref. [148] showed that

(P1) *(Scalar multiplication).* $F(\lambda\Omega_{RA}) = \lambda F(\Omega_{RA})$ for any $\lambda > 0$.

(P2) *(Convexity).* $F(p\Omega_{RA} + (1-p)\Omega'_{RA}) \leq pF(\Omega_{RA}) + (1-p)F(\Omega'_{RA})$ for any $p \in [0,1]$.

(P3) *(Invariance under local isometries).* Let $\mathcal{U}_R$ and $\mathcal{V}_A$ be local unitary channels on subsystems R and A respectively. Then $F(\mathcal{U}_R \circ \mathcal{V}_A(\Omega_{RA})) = F(\Omega_{RA})$.

(P4) *(Local data processing inequality).* Let $\Phi_R : \mathcal{B}(R) \to \mathcal{B}(R')$ be a unital CPTP map and $\Lambda_A : \mathcal{B}(A) \to \mathcal{B}(B)$ be a CPTP map. Then $F\left((\Phi_R \otimes \mathsf{id}_A)(\Omega_{RA})\right) \leq F(\Omega_{RA})$ and $F((\mathsf{id}_R \otimes \Lambda_A)(\Omega_{RA})) \leq F(\Omega_{RA})$.

We also recall the following notation for evaluating F on bipartite $\mathcal{G}$-twirled states,

$$F_\eta(\tau) \coloneqq F(\mathcal{G}(\eta_R \otimes \tau_A)), \tag{C.2}$$

and that $\Delta F_\eta \coloneqq F_\eta(\rho) - F_\eta(\sigma) \geq 0$ is equivalent to $\Delta H_\eta \coloneqq H_\eta(\sigma) - H_\eta(\rho) \geq 0$. This is because $F(X) > 0$ for any X that is positive-semidefinite with at least one non-zero eigenvalue [148], which is true of $\mathcal{G}(\eta_R \otimes \tau_A)$ formed between any reference frame η_R and input state τ_A because it is a quantum state, and $-\log(x)$ is monotonically decreasing in x for $x > 0$.

C.1.1 Invariance under local unitary channels that commute with $\mathcal{G}$

We now prove the following Lemma from the main text, which specializes property **(P3)** to the particular conditional min-entropies appearing in Theorem 42.

Lemma 46. *Let $\mathcal{U}^R$ and $\mathcal{V}^S$ be unitary channels on the reference and input systems respectively such that $\mathcal{U}^R \otimes \mathcal{V}^S$ is G-covariant (on the joint system RS). Then given any reference frame η of system R and input state τ of system S, we have that*

$$H_{\mathcal{U}^R(\eta)}\left[\mathcal{V}^S(\tau)\right] = H_\eta(\tau). \tag{5.25}$$

Proof. By a straightforward appeal to Property **(P3)**, we have

$$F\left(\mathcal{G}\left[\mathcal{U}^R(\eta_R) \otimes \mathcal{V}^S(\tau_S)\right]\right) = F\left(\mathcal{U}^R \otimes \mathcal{V}^S \circ \mathcal{G}\left[\eta_R \otimes \tau_S\right]\right) = F(\mathcal{G}[\eta_R \otimes \tau_A]). \tag{C.3}$$

Since $H_\eta(\rho) \coloneqq -\log F(\mathcal{G}[\eta_R \otimes \tau_S])$, this implies $H_{\mathcal{U}(\eta)}(\mathcal{V}(\tau)) = H_\eta(\tau)$. □

C.1.2 Symmetric input states

We now prove that, when η is symmetric, H_η has the same value on all input states, which implies that if η is symemtric, $\Delta H_\eta = 0$ and imposes no constraints on state transitions.

Lemma 44. *Given any input state ρ_S of system S and reference frame η_R of system R,*

$$H_{\mathcal{G}(\eta)}(\rho) = H_\eta(\mathcal{G}(\rho)) = H_{\mathcal{G}(\eta)}(\mathcal{G}(\rho)) = -\log \|\mathcal{G}(\eta)\|_\infty, \tag{5.24}$$

where $\|\cdot\|_\infty$ is the Schatten-∞ norm defined in Eq. (2.8). Therefore, if η is symmetric so we have $\eta = \mathcal{G}(\eta)$, then $\Delta H_\eta = 0$.

Proof. The first two equalities straightforwardly follow from

$$\mathcal{G}(\mathcal{G}(\eta) \otimes \rho) = \mathcal{G}(\eta) \otimes \mathcal{G}(\rho) = \mathcal{G}(\eta \otimes \mathcal{G}(\rho)) = \mathcal{G}(\mathcal{G}(\eta) \otimes \mathcal{G}(\rho)) \tag{C.4}$$

To show the final equality, we first examine

$$F(\mathcal{G}(\eta)\otimes\mathcal{G}(\rho)) = \inf_{X\geq 0}\{\mathrm{Tr}[X] \;:\; \mathbb{1}\otimes X - \mathcal{G}(\eta)\otimes\mathcal{G}(\rho)\geq 0\} \tag{C.5}$$

Since $\mathcal{G}(\eta)$ is Hermitian, it can be diagonalised as $\mathcal{G}(\eta) \coloneqq \sum_i \mu_i |i\rangle\langle i|$ for some orthonormal basis $\{|i\rangle\}$ of the reference system. Working in this basis, we obtain

$$\mathbb{1}\otimes X - \mathcal{G}(\eta)\otimes\mathcal{G}(\rho) = \sum_i |i\rangle\langle i|\otimes X - \sum_i \mu_i |i\rangle\langle i|\otimes\mathcal{G}(\rho) = \sum_i |i\rangle\langle i|\otimes(X-\mu_i\mathcal{G}(\rho)) \tag{C.6}$$

Therefore, $\mathbb{1}\otimes X - \mathcal{G}(\eta)\otimes\mathcal{G}(\rho)\geq 0$ if and only if $X-\mu_i\mathcal{G}(\rho)\geq 0$ for all i, which in turn is true if and only if $X-\mu_{\max}\mathcal{G}(\rho)\geq 0$, where $\mu_{\max}$ is the largest eigenvalue of $\mathcal{G}(\rho)$. We can therefore lower-bound the $\mathrm{Tr}[X]$ needed to achieve this by noting that if $X-\mu_{\max}[\mathcal{G}(\rho)]\geq 0$, then given any orthonormal basis $\{|j\rangle\}$ for the input Hilbert space, we have that

$$\forall j : \langle j|X|j\rangle - \mu_{\max}\langle j|\mathcal{G}(\rho)|j\rangle \geq 0 \tag{C.7}$$

and can therefore identify the following lower bound after summing the above inequalities over j:

$$\implies \mathrm{Tr}[X]\geq\mu_{\max}\,\mathrm{Tr}[\mathcal{G}(\rho)] = \mu_{\max} \tag{C.8}$$

This lower bound can be attained by a positive semidefinite X simply by setting $X\coloneqq\mu_{\max}\mathcal{G}(\rho)$. Therefore,

$$F(\mathcal{G}(\eta)\otimes\mathcal{G}(\rho)) = \mu_{\max} = \|\mathcal{G}(\rho)\|_\infty \Rightarrow H_{\mathcal{G}(\eta)}(\mathcal{G}(\rho)) = -\log\|\mathcal{G}(\rho)\|_\infty, \tag{C.9}$$

where we have used the fact that the largest eigenvalue of a positive semidefinite operator can be identified by its Schatten-∞ norm. □

C.2 A sufficient surface of reference frames

C.2.1 Depolarizing the reference state

Let us define the partially depolarizing channel for some fixed probability p:

$$\Lambda_p[\rho] \coloneqq p\rho + (1-p)\frac{\mathbb{1}}{d}. \tag{C.10}$$

The following lemma shows that the functional $F_\eta(\rho)$ behaves linearly when we take convex combinations of the reference state with the maximally mixed state.

Lemma 90. *For any reference state η_R and input state τ_A, we have*

$$F_{\Lambda_p(\eta)}(\tau) = pF_\eta(\tau) + (1-p)F_{\mathbb{1}/d}(\tau) = pF_\eta(\tau) + \frac{1-p}{d}, \tag{C.11}$$

where d is the Hilbert space dimension of the reference system.

Proof. Because the maximally mixed state is symmetric under any group G, we have

$$\mathcal{G}(\Lambda_p(\eta_R)\otimes\tau_A) = p\mathcal{G}(\eta_R\otimes\tau_A) + \left(\frac{1-p}{d}\right)\mathbb{1}_R\otimes\mathcal{G}(\tau_A). \tag{C.12}$$

Substituting this into Eq. (C.2) and rearranging terms gives

$$\mathcal{F}_{\Lambda_p[\eta]}(\tau) = \inf_{X_A\geq 0}\left\{\mathrm{Tr}[X_A] : \mathbb{1}_R\otimes\left(X_A - \frac{1-p}{d}\mathcal{G}(\tau_A)\right) - p\mathcal{G}\left(\eta_R\otimes\tau_A\right)\geq 0\right\}. \tag{C.13}$$

Given any positive semidefinite operators T_A and Z_{RA}, the conditions (C1) $X_A \geq 0$ and (C2) $\mathbb{1}_R\otimes(X_A - T_A) - Z_{RA}\geq 0$ on an operator $X_A\in\mathcal{B}(A)$ become equivalent to the conditions (C'1) $X_A - T_A\geq 0$ and (C2) because (C2) implies (C'1), which further implies (C1). As $\frac{1-p}{d}\mathcal{G}(\tau_A)$ and $p\mathcal{G}\left(\eta_R\otimes\tau_A\right)$ are subnormalised quantum states and there positive semidefinite, we can rewrite the set over which we perform minimisation in Eq. (C.13) by

$$F_{\Lambda_p[\eta]}(\tau) = \inf_{X_A - \frac{1-p}{d}\mathcal{G}[\tau_A]\geq 0}\left\{\mathrm{Tr}[X_A] \;:\; \mathbb{1}_R\otimes\left(X_A - \frac{1-p}{d}\mathcal{G}[\tau_A]\right) - p\mathcal{G}\left[\eta_R\otimes\tau_A\right]\geq 0\right\} \tag{C.14}$$

$$= \inf_{Y_A\geq 0}\left\{\mathrm{Tr}\left[Y_A + \frac{1-p}{d}\mathcal{G}[\tau_A]\right] \;:\; \mathbb{1}_R\otimes Y_A - p\mathcal{G}\left[\eta_R\otimes\tau_A\right]\geq 0\right\}, \tag{C.15}$$

where we have defined $Y_A \coloneqq X_A - \frac{1-p}{d}\mathcal{G}(\tau_A)$. Since $\mathrm{Tr}[\mathcal{G}(\tau_A)] = 1$ as $\mathcal{G}(\tau_A)$

$$F_{\Lambda_p[\eta]}(\tau) = \inf_{Y_A \geq 0} \{\mathrm{Tr}[Y_A] \ : \ \mathbb{1}_R \otimes Y_A - p\mathcal{G}[\eta_R \otimes \tau_A] \geq 0\} + \frac{1-p}{d}. \tag{C.16}$$

Finally, we make use of property **(P1)** to arrive at

$$F_{\Lambda_p[\eta]}(\tau) = pF_\eta(\tau) + \frac{1-p}{d}. \tag{C.17}$$

which concludes the proof. □

An immediate consequence of Lemma 90 is that taking a statistical mixture of any reference state with the maximally mixed state does not change the entropic relation in Theorem 42.

Corollary 4. *Given any reference frame η_R, we have that*

$$\Delta H_\eta \geq 0 \iff \Delta H_{\Lambda_p[\eta]} \geq 0 \textit{for any } p \in (0,1]. \tag{C.18}$$

Proof. Lemma 90 implies that $\Delta F_{\Lambda_p[\eta]} = p\Delta F_\eta$ for any probability p. Therefore, $\Delta F_\eta \geq 0$ if and only if $\Delta F_{\Lambda_p[\eta]} \geq 0$, for any $p \in (0,1]$. The Lemma statement is established by recalling that $\Delta F_\eta \geq 0$ is equivalent to $\Delta H_\eta \geq 0$. □

C.2.2 Proof of Theorem 47

We now present a proof of Theorem 47, which we restate here for clarity:

Theorem 47. *(Sufficient surfaces of states). Let ρ_A and σ_B be states of an input system A and output system B respectively, and let G be a compact symmetry group. Furthermore, let $\partial\mathcal{D}$ be any closed surface in the state space $\mathcal{D}(R)$ of the reference system R for G-covariant state transitions from A to B that encloses the maximally mixed reference state $\frac{I}{d}$, where $d = \dim(\mathcal{H}_R)$. Then $\rho \xrightarrow{G} \sigma$ if and only if*

$$\forall \eta \in \partial\mathcal{D} : \Delta H_\eta \geq 0. \tag{5.26}$$

Proof. If the transformation is possible under a G–covariant channel, then $\Delta H_\eta \geq 0$ for all reference frames η_R, and therefore for all η_R restricted to $\partial\mathcal{D}$. Conversely, suppose $\Delta H_\eta \geq 0$ for all $\eta_R \in \partial\mathcal{D}$. Let η'_R be an reference frame that is *not* maximally mixed, and consider the family of reference frames $\eta'_R(p) \coloneqq \Lambda_p(\eta'_R)$ for $p \in [0,1]$ formed by depolarising η'_R. This defines a continuous line of states connecting η'_R to the maximally mixed state $\mathbb{1}/d$. Since $\partial\mathcal{D}$

encloses the maximally mixed state, the set $\{\eta'_R(p) : 0 \leq p \leq 1\}$ must either intersect $\partial\mathcal{D}$ for some value $p_\star$ with $0 < p_\star \leq 1$ or lie entirely within the interior of $\partial\mathcal{D}$. If the entire set is inside $\partial\mathcal{D}$, then we can find another reference frame $\eta''_R \in \partial\mathcal{D}$ such that $\Lambda_q[\eta''_R] = \eta'_R$ for some probability $q \in (0, 1)$. By Corollary 4, we have that η'_R and $\eta'_R(p_\star)$ (or η'_R and η''_R for the second case) give equivalent entropic constraints. Since η'_R was arbitrarily chosen, we conclude that we can restrict to reference on $\partial\mathcal{D}$. □

C.3 The conical structure of $F_\tau(\mathbf{x})$

In this section, we make use of the co-ordinate system for Hermitian operators (of the reference system) established in Eq. (5.27). Furthermore, we introduce the notation

$$\tilde{F}_\tau(\boldsymbol{x}) \coloneqq F_{\eta(\boldsymbol{x})}(\tau) - F_{\eta(\mathbf{0})}(\tau) \tag{C.19}$$

to represent the difference in $F_\eta(\tau)$ from the reference frame designated by $\boldsymbol{x}$ in this co-ordinate system and the maximally mixed state at $\boldsymbol{x} = 0$. Let $\mathcal{X}(R) \coloneqq \{\boldsymbol{x} \in \mathbb{R}^{d^2-1} | \eta(\boldsymbol{x}) \in \mathcal{D}(R)\}$ be the co-ordinates designating valid quantum states of the reference system R.

Lemma 91. *Let $\lambda \geq 0$. Then for all $\boldsymbol{x}, \lambda\boldsymbol{x} \in \mathcal{X}(R)$, we have that*

$$\tilde{F}_\tau(\lambda\boldsymbol{x}) = \lambda\tilde{F}_\tau(\boldsymbol{x}) \tag{C.20}$$

Furthermore,

$$\tilde{F}_\tau(\boldsymbol{x}) \geq 0. \tag{C.21}$$

Proof. We first prove that $\tilde{F}_\tau$ is always non-negative. Let $\mathcal{P}_R$ be the completely depolarising channel on the reference system. Then by Property **(P4)** of $F_\tau(\mathbf{x})$, and the fact that performing a completely depolarising channel on the reference system is G-covariant, we have

$$\begin{aligned} F_{\eta(\boldsymbol{x})}(\tau) \coloneqq F[\mathcal{G}(\eta(\boldsymbol{x}) \otimes \tau)] &\geq F[\mathcal{P}_R \otimes \mathsf{id}_A \circ \mathcal{G}(\eta(\boldsymbol{x}) \otimes \tau)] \\ &= F[\mathcal{G}(\mathcal{P}_R[\eta(\boldsymbol{x})] \otimes \tau)] \\ &= F\left[\mathcal{G}\left(\eta(\mathbf{0}) \otimes \tau_A\right)\right] \equiv F_{\eta(\mathbf{0})}(\tau) \end{aligned} \tag{C.22}$$

Therefore, $\tilde{F}_\tau(\mathbf{x}) \geq 0$ as claimed.

Let p be a probability, and define $A(\boldsymbol{x}) \coloneqq \sum_{k=1}^{d^2-1} x_k X_k$. Then partially depolarising $\eta(\boldsymbol{x})$ with probability p has the effect

$$p\eta(\mathbf{x}) + (1-p)\frac{\mathbb{1}}{d} = p\frac{\mathbb{1}}{d} + pA(\mathbf{x}) + (1-p)\frac{\mathbb{1}}{d} = \frac{\mathbb{1}}{d} + A(p\mathbf{x}) \tag{C.23}$$

Using this equation, we can rewrite Lemma 90 as:

$$\forall p \in [0,1], \boldsymbol{x} \in \mathcal{X}(R) : F_{\eta(p\mathbf{x})} = pF_{\eta(\mathbf{x})}(\tau) + \frac{1-p}{d}. \tag{C.24}$$

We then immediately have

$$\forall p \in [0,1], \boldsymbol{x} \in \mathcal{X}(R) : \tilde{F}_\tau(p\mathbf{x}) = pF_{\eta(\mathbf{x})} + \frac{1-p}{d} - \frac{1}{d} = p\left(F_{\eta(\mathbf{x})}(\tau) - \frac{1}{d}\right) = p\tilde{F}_\tau(\mathbf{x}) \tag{C.25}$$

where in the final equality we applied Lemma 44 to obtain $F_{\eta(\mathbf{0})}(\tau) \equiv F_{\frac{I}{d}}(\tau) = \left\|\frac{I}{d}\right\|_\infty = \frac{1}{d}$ as the maximally mixed state is symmetric.

Making the change of variables $\mathbf{x}' \coloneqq p\mathbf{x}$, we find that Eq. (C.25) is equivalent to:

$$\forall p \in [0,1], p^{-1}\boldsymbol{x}' \in \mathcal{X}(R) : \tilde{F}_\tau(\mathbf{x}') = p\tilde{F}\left(p^{-1}\mathbf{x}'\right) \Rightarrow p^{-1}\tilde{F}_\tau(\mathbf{x}') = \tilde{F}_\tau\left(p^{-1}\mathbf{x}'\right). \tag{C.26}$$

Since $1 \leq p^{-1} \leq \infty$, we can combine this fact with Eq. (C.25) to conclude that

$$\forall \lambda > 0, \boldsymbol{x}, \lambda\boldsymbol{x} \in \mathcal{X}(R) : \tilde{F}_\tau(\lambda\mathbf{x}) = \lambda\tilde{F}_\tau(\mathbf{x}), \tag{C.27}$$

as claimed. □

Lemma 91 implies that $\tilde{F}_\tau(\boldsymbol{x})$, and consequently $F_\eta(\tau)$, is linearly non-decreasing in every direction out of the maximally mixed state. This means $F_\eta(\tau)$ will, in general, have a conical form at the maximally mixed state. As a result, unless $\Delta F_{\eta(\boldsymbol{x})}$ is completely linear in $\boldsymbol{x}$, it too will have a conical form at the maximally mixed state.

We are now in a position to prove Lemma 49.

Lemma 49. *In a sufficiently small neighbourhood $\epsilon > 0$ around the maximally mixed state, ΔH_η*

changes linearly along any direction out of the maximally mixed state. Thus one can write

$$\forall \boldsymbol{x} \in \mathcal{S} : \Delta H_{\eta(\epsilon \boldsymbol{x})} = f(\boldsymbol{x})\epsilon + O(\epsilon^2), \tag{5.28}$$

for some function $f : \mathcal{S} \to \mathbb{R}$.

Proof. Since $H_{\eta(\epsilon \boldsymbol{x})}(\tau) \coloneqq -\log F_{\eta(\epsilon \boldsymbol{x})}(\tau)$, we can Taylor-expand around $F_{\eta(\mathbf{0})} = \frac{1}{d}$ to obtain

$$\begin{aligned} H_{\eta(\epsilon \boldsymbol{x})}(\tau) = -\log\Big(F_{\eta(\mathbf{0})}(\tau) + \tilde{F}_\tau(\epsilon \boldsymbol{x})\Big) &= -\log\left(\frac{1}{d} + \epsilon \tilde{F}_\tau(\boldsymbol{x})\right) \\ &= \log d - d\epsilon \tilde{F}_\tau(\boldsymbol{x}) + O(\epsilon^2) \end{aligned} \tag{C.28}$$

in a small ϵ-neighbourhood around the maximally mixed state, which implies

$$\Delta H(\epsilon \boldsymbol{x}) = d\epsilon \Delta F_{\eta(\boldsymbol{x})} + O(\epsilon^2) \tag{C.29}$$

as claimed. □

C.4 Calculating $F_\eta(\tau)$ for time-symmetric transformations in a non-degenerate qubit

In this section, we calculate $F_\eta(\tau)$ for the case of time-symmetric transformations from a qubit S with Hamiltonian $H_S \propto Z$ to itself, which forms the symmetry group U(1) represented on $\mathcal{H}_S$ as $\{\mathcal{U}_t^S \coloneqq e^{-iZt} | t \in [0, 2\pi)\}$. In this case, the reference system R is a qubit with Hamiltonian $H_R = -Z$, and we will describe reference states by their Bloch vector (x, y, z) as

$$\eta(x, y, z) = \frac{\mathbb{1}}{2} + x\frac{X}{2} + y\frac{Y}{2} + z\frac{Z}{2}, \tag{C.30}$$

where $x^2 + y^2 + z^2 \leq 1$. The derivation will be conducted entirely in the energy eigenbasis.

We first introduce the following simplifying lemma that allows us to vastly reduce the feasible set over which we must carry out the optimisation problem in $F_\eta(\tau)$.

Lemma 92. *When calculating $F_\eta(\tau)$ defined by the minimisation problem*

$$F_\eta(\tau) \coloneqq \inf_{X_S \geq 0} \{\mathrm{Tr}[X_S] : \mathbb{1}_R \otimes X_S \geq \mathcal{G}(\eta_R \otimes \tau_S)\}, \tag{C.31}$$

it is sufficient to minimise over positive-semidefinite operators that are G-covariant, i.e.

$$F_\eta(\tau) = \inf_{X_S \geq 0, X_S = \mathcal{G}(X_S)} \{\mathrm{Tr}[X_S] : \mathbb{1}_R \otimes X_S \geq \mathcal{G}(\eta_R \otimes \tau_S)\}. \tag{C.32}$$

Proof. Suppose the compact group G has a representation $g \to U(g)$ on the Hilbert space $\mathcal{H}$, and let K be a Hermitian operator on $\mathcal{B}(H)$. We first observe that

$$K \geq 0 \implies \langle\psi|\,\mathcal{G}(K)\,|\psi\rangle = \int_G dg\, [\langle\psi|\,U(g)]K[U^\dagger(g)\,|\psi\rangle] \geq 0 \iff \mathcal{G}(K) \geq 0 \tag{C.33}$$

Furthermore,

$$\mathrm{Tr}[\mathcal{G}(K)] = \int_G dg\ \mathrm{Tr}\Big[U(g)KU^\dagger(g)\Big] = \int_G dg\ \mathrm{Tr}[K] = \mathrm{Tr}[K]. \tag{C.34}$$

The set of operators $\mathcal{S}$ over which $F_\eta(\tau)$ minimises is

$$\mathcal{S} \coloneqq \{X_S \in \mathcal{B}(\mathcal{H}_S) | X_S \geq 0, \mathbb{1}_R \otimes X_S - \mathcal{G}(\eta_R \otimes X_S) \geq 0\}. \tag{C.35}$$

By Eq. (C.33), we see that $X_S \geq 0 \implies \mathcal{G}(X_S) \geq 0$, and furthermore that

$$\begin{aligned} \mathbb{1}_R \otimes X_S - \mathcal{G}(\eta_R \otimes X_S) \geq 0 \implies & \mathcal{G}[\mathbb{1}_R \otimes X_S - \mathcal{G}(\eta_R \otimes X_S)] \\ = & \mathbb{1}_R \otimes \mathcal{G}(X_S) - \mathcal{G}(\eta_R \otimes \tau_A) \geq 0. \end{aligned} \tag{C.36}$$

We thus conclude that if $X_S \in \mathcal{S}$, then $\mathcal{G}(X_S) \in \mathcal{S}$. Furthermore, by Eq. (C.34), we see that $\mathcal{G}(X_S)$ achieves the same value on the objective function $\mathrm{Tr}[\cdot]$ to be minimised in $F_\eta(\tau)$ as X_S. Consequently, we can restrict the minimisation in $F_\eta(\tau)$ to X_S where $X_S = \mathcal{G}(X_S)$. □

This Lemma allows us to take X_S to be diagonal in the energy eigenbasis, i.e.

$$X = \begin{pmatrix} x_1 & 0 \\ 0 & x_2 \end{pmatrix} \Rightarrow \mathbb{1} \otimes X = \begin{pmatrix} x_1 & 0 & 0 & 0 \\ 0 & x_2 & 0 & 0 \\ 0 & 0 & x_1 & 0 \\ 0 & 0 & 0 & x_2 \end{pmatrix} \tag{C.37}$$

Local evolution of the reference system R under its Hamiltonian is described by the unitary

channels $\{\mathcal{U}_t^R \otimes \mathsf{id}_S := e^{iZt}(\cdot)e^{-iZt} \otimes I_S(\cdot)I_S | t \in [0, 2\pi)$. Since all these channels are time-symmetric on the joint reference and qubit system RS, we can apply Lemma 46 to conclude that $F_{\eta(x,y,z)}(\tau)$ is invariant under all rotations of the η Bloch vector around the Z-axis of the Bloch sphere, which are induced by these channels. Consequently, we can restrict our calculation to $x \geq 0, y = 0$, and then equate

$$F_{\eta(x,y,z)}(\tau) = F_{\eta\left(\sqrt{x^2+y^2},0,z\right)}(\tau) \tag{C.38}$$

Furthermore, simultaneously performing the Pauli-X unitary channel on the reference and input system, which is described by $\mathcal{X}^R \otimes \mathcal{X}^S := X(\cdot)X \otimes X(\cdot)X$, is time-symmetric because $[X \otimes X, e^{iZt} \otimes e^{-iZt}] = 0$ for all $t \in [0, 2\pi)$. We therefore conclude that

$$F_{\eta(x,0,z)}(X\tau_A X) = F_{\eta(x,0,-z)}(\tau), \tag{C.39}$$

which means we can additionally restrict our attention to $z \geq 0$. We therefore only need to consider reference states of the form:

$$\eta(x, y, z) = \frac{1}{2}\begin{pmatrix} 1+z & x \\ x & 1-z \end{pmatrix}, \; x, z \geq 0. \tag{C.40}$$

Parameterising each state of the non-degenerate qubit in the energy eigenbasis as shown in Eq. (5.22), we can write

$$\eta_R \otimes \tau_S = \frac{1}{2}\begin{pmatrix} 1+z & x \\ x & 1-z \end{pmatrix} \otimes \begin{pmatrix} p_\tau & c_\tau \\ c_\tau^* & 1-p_\tau \end{pmatrix} \tag{C.41}$$

$$= \frac{1}{2}\begin{pmatrix} p_\tau(1+z) & c_\tau(1+z) & p_\tau x & c_\tau x \\ c_\tau^*(1+z) & (1-p_\tau)(1+z) & c_\tau^* x & (1-p_\tau)x \\ p_\tau x & c_\tau x & p_\tau(1-z) & c_\tau(1-z) \\ c_\tau^* x & (1-p_\tau)x & c_\tau^*(1-z) & (1-p_\tau)(1-z) \end{pmatrix} \tag{C.42}$$

The bipartite G-twirl on the joint reference and qubit system RS projects out the time-symmetric part of $\eta \otimes \tau$ (see discussion around Eq. (6.18)), i.e. the parts of $\eta_R \otimes \tau_S$ with support on the

total energy eigenspaces $\mathrm{span}\,|00\rangle, |11\rangle$, $\mathrm{span}\,|01\rangle$ and $\mathrm{span}\,|10\rangle$ of RS. We therefore conclude

$$\mathcal{G}(\eta_R \otimes \tau_S) = \frac{1}{2}\begin{pmatrix} p_\tau(1+z) & 0 & 0 & c_\tau x \\ 0 & (1-p_\tau)(1+z) & 0 & 0 \\ 0 & 0 & p_\tau(1-z) & 0 \\ c_\tau^* x & 0 & 0 & (1-p_\tau)(1-z) \end{pmatrix}. \tag{C.43}$$

and so obtain

$$\mathbb{1}_R \otimes X_S - \mathcal{G}(\eta_R \otimes \tau_S) = \begin{pmatrix} x_1 - p_\tau\frac{1+z}{2} & 0 & 0 & -c_\tau\frac{x}{2} \\ 0 & x_2 - (1-p_\tau)\frac{1+z}{2} & 0 & 0 \\ 0 & 0 & x_1 - p_\tau\frac{1-z}{2} & 0 \\ -c_\tau^*\frac{x}{2} & 0 & 0 & x_2 - (1-p_\tau)\frac{1-z}{2} \end{pmatrix}. \tag{C.44}$$

Re-ordering our energy eigenbasis from $|00\rangle, |01\rangle, |10\rangle, |11\rangle$ to $|00\rangle, |11\rangle, |01\rangle, |10\rangle$ so that the matrix above appears block-diagonal (in total energy), we obtain

$$\mathbb{1}_R \otimes X_S - \mathcal{G}(\eta_R \otimes \tau_S) = \begin{pmatrix} x_1 - p_\tau\frac{1+z}{2} & -c_\tau\frac{x}{2} & 0 & 0 \\ -c_\tau^*\frac{x}{2} & x_2 - (1-p_\tau)\frac{1-z}{2} & 0 & 0 \\ 0 & 0 & x_2 - (1-p_\tau)\frac{1+z}{2} & 0 \\ 0 & 0 & 0 & x_1 - p_\tau\frac{1-z}{2} \end{pmatrix}. \tag{C.45}$$

The Sylvester Criterion [149] states that a matrix is positive semidefinite if and only if all its upper-left determinants are non-negative, which produces the following criteria:

1. $x_1 \geq p_\tau\frac{1+z}{2}$.
2. $(x_1 - p_\tau\frac{1+z}{2})(x_2 - (1-p_\tau)\frac{1-z}{2}) \geq \frac{|c_\tau|^2 x^2}{4}$.
3. $x_2 \geq (1-p_\tau)\frac{1+z}{2}$.
4. $x_1 \geq p_\tau\frac{1-z}{2}$.

and we conclude that

$$F_\eta(\tau) = \min_{x_1,x_2\geq 0}\{x_1 + x_2 : \text{conditions 1-4}\} \tag{C.46}$$

The restriction $z \geq 0$ implies that condition 4 is redundant given condition 1. For the same reason, any value of x_2 satisfying condition 3 also satisfies $x_2 \geq x_2 - (1-p_\tau)\frac{1-z}{2}$, so we can transform condition 2 into

$$x_1 - p_\tau\frac{1+z}{2} \geq \frac{|c_\tau|^2x^2}{4}\frac{1}{x_2-(1-p_\tau)\frac{1-z}{2}}. \tag{C.47}$$

where the right hand side of the above equation is non-negative and thus yields a value of x_1 that satisfies condition 1. Therefore, we need only consider conditions 2-3 under our restriction $z \geq 0$, yielding

$$F_\eta(\tau) = \min_{x_1,x_2\geq 0}\{x_1 + x_2 : \text{ conditions 2-3}\} \tag{C.48}$$

Given any value of x_2 satisfying condition 3, x_1 is minimised when the inequality from the equivalent expression of condition 2 in Eq. (C.47) becomes an equality, i.e. at

$$x_1 = \frac{|c_\tau|^2x^2}{4}\frac{1}{x_2-(1-p_\tau)\frac{1-z}{2}} + p_\tau\frac{1+z}{2} \tag{C.49}$$

We can therefore eliminate x_1 from $F_\eta(\tau)$ and obtain

$$F_\eta(\tau) = \min_{x_2\geq 0}\left\{x_2 + \frac{|c_\tau|^2x^2}{4}\frac{1}{x_2-(1-p_\tau)\frac{1-z}{2}} + p_\tau\frac{1+z}{2} : x_2 \geq (1-p_\tau)\frac{1+z}{2}\right\} \tag{C.50}$$

For $x_2 \geq x_2 - (1-p_\tau)\frac{1-z}{2}$, the objective function in Eq. (C.50) is a non-negative hyperbola with asymptotes $x = (1-p_\tau)\frac{1-z}{2}$ and $y = x$ attaining minimum (by standard derivative tests) at

$$x_2 = (1-p_\tau)\frac{1-z}{2} + |c_\tau|\frac{x}{2}. \tag{C.51}$$

Therefore, the value of x_2 solving the minimisation of $F_\eta(\tau)$ should be

$$x_2 = \max\left\{(1-p_\tau)\frac{1+z}{2}, (1-p_\tau)\frac{1-z}{2} + |c_\tau|\frac{x}{2}\right\} \tag{C.52}$$

i.e. the value of x_2 where the objective function attains its minimum if that value of x_2 satisfies the constraint, or the minimum value of x_2 satisfying the constraint if this is not the case. Substituting this into Eq. (C.50), we obtain

$$F_{\eta(x\geq 0,y=0,z\geq 0)}(\tau) = \begin{cases} \frac{|c_\tau|^2}{1-p_\tau}\frac{x^2}{4z} + \frac{z}{2} + \frac{1}{2} & \text{for } \frac{x}{2z} \leq \frac{1-p_\tau}{|c_\tau|} \\ \left(p_\tau - \frac{1}{2}\right) z + |c_\tau| x + \frac{1}{2} & \text{for } \frac{x}{2z} \geq \frac{1-p_\tau}{|c_\tau|} \end{cases} \tag{C.53}$$

Finally, applying Equations C.38 and C.39, and noting that $X\tau_S X$ translates to $p_\tau \to 1 - p_\tau$ and $c_\tau \to c_\tau^*$ in our parameterisation of τ, we can calculate F_τ for all reference states from the above result as

$$F_{\eta(x,y,z)}(\tau) = \begin{cases} \frac{|c_\tau|^2}{1-p_\tau}\frac{x^2+y^2}{4z} + \frac{z}{2} + \frac{1}{2} & \text{for } 0 \leq \frac{\sqrt{x^2+y^2}}{2z} \leq \frac{1-p_\tau}{|c_\tau|} \\ \left(p_\tau - \frac{1}{2}\right) z + |c_\tau|\sqrt{x^2+y^2} + \frac{1}{2} & \text{for } \frac{\sqrt{x^2+y^2}}{2z} \geq \frac{1-p_\tau}{|c_\tau|} \text{ and } \frac{\sqrt{x^2+y^2}}{2z} \leq -\frac{p_\tau}{|c_\tau|} \\ -\frac{|c_\tau|^2}{p_\tau}\frac{x^2+y^2}{4z} - \frac{z}{2} + \frac{1}{2} & \text{for } 0 \geq \frac{\sqrt{x^2+y^2}}{2z} \geq -\frac{p_\tau}{|c_\tau|}. \end{cases} \tag{C.54}$$

From the above evaluation of F_η, we can see that unless $\frac{|c_\rho|^2}{1-p_\rho} = \frac{|c_\sigma|^2}{1-p_\sigma}$ and $\frac{|c_\rho|^2}{p_\rho} = \frac{|c_\sigma|^2}{p_\sigma}$, we can always find a region around the poles of the Bloch sphere were $\Delta F_{\eta(x,y,z)}$ is not going to be linear. Since these conditions are met if and only if $p_\rho = p_\sigma$ and $|c_\rho| = |c_\sigma|$, i.e. ρ can be transformed into σ by time evolution under the Hamiltonian, ΔF_η has a conic singularity unless ρ can be transformed into σ.

C.4.1 Two entropic conditions suffice to characterise time-covariant transformations in a non-degenerate qubit

Ref. [133] identified the following two necessary and sufficient conditions for the existence of a time-covariant transformation from ρ to σ in a non-degenerate qubit with Hamiltonian proportional to Z:

$$\frac{|c_\rho|^2}{1-p_\rho} \geq \frac{|c_\sigma|^2}{1-p_\sigma} \text{ and } \frac{|c_\rho|^2}{p_\rho} \geq \frac{|c_\sigma|^2}{p_\sigma} \tag{C.55}$$

We see from Eq. (C.54) that these conditions are equivalent to

$$\Delta F_{\eta(x_1,y_1z_1)} \geq 0 \text{ where } 0 > \frac{\sqrt{x_1^2+y_1^2}}{2z_1} \leq \min\left\{\frac{1-p_\rho}{|c_\rho|}, \frac{1-p_\sigma}{|c_\sigma|}\right\} \tag{C.56}$$

$$\Delta F_{\eta(x_2,y_2,z_2)} \geq 0 \text{ where } 0 > \frac{\sqrt{x_2^2+y_2^2}}{2z_2} \geq \max\left\{-\frac{p_\rho}{|c_\rho|}, -\frac{p_\sigma}{|c_\sigma|}\right\}. \tag{C.57}$$

We can alternatively characterise each qubit reference frame as $\eta(r,\theta,\phi)$, where $(r,\theta,\phi)^T$ is the state's Bloch vector in spherical polar co-ordinates (radial, polar and azimuthal respectively). Using the standard conversion between spherical and Cartesian co-ordinates, which is given by $x = r\sin(\theta)\cos(\phi)$, $y = r\sin(\theta)\sin(\phi)$ and $z = r\cos(\theta)$, we can rewrite Equation C.54 as

$$F_{\eta(r,\theta,\phi)} = \begin{cases} r\frac{|c_\tau|^2}{1-p_\tau}\frac{\tan(\theta)\sin(\theta)}{4} + r\frac{\cos(\theta)}{2} + \frac{1}{2} & \text{for } 0 \leq \frac{\tan(\theta)}{2} \leq \frac{1-p_\tau}{|c_\tau|} \\ r\left(p_\tau - \frac{1}{2}\right)\cos(\theta) + r|c_\tau|\sin(\theta) + \frac{1}{2} & \text{for } \frac{\tan(\theta)}{2} \geq \frac{1-p_\tau}{|c_\tau|} \text{ and } \frac{\tan(\theta)}{2} \leq -\frac{p_\tau}{|c_\tau|} \\ -r\frac{|c_\tau|^2}{p_\tau}\frac{\tan(\theta)\sin(\theta)}{4} - r\frac{\cos(\theta)}{2} + \frac{1}{2} & \text{for } 0 \geq \frac{\tan(\theta)}{2} \geq -\frac{p_\tau}{|c_\tau|}. \end{cases} \tag{C.58}$$

from which we see that F_η, and therefore H_η as well as ΔH_η, is infinitely partially differentiable with respect to θ at $\theta = 0$ and $\theta = \pi$.

Lemma 43. *In spherical polar co-ordinates on the Bloch sphere, there exists a time-covariant transformation ρ to σ in a qubit with Hamiltonian $\propto Z$ if and only if*

$$\partial_\theta^2(\Delta H_\eta)|_{\theta=0} \geq 0 \textit{ and } \partial_\theta^2(\Delta H_\eta)|_{\theta=\pi} \geq 0. \tag{5.21}$$

Proof. From Eq. (C.58), we can straightforwardly evaluate:

$$\partial_\theta^2 F_\eta(\tau)|_{\theta=0} = \frac{r}{2}\left(\frac{|c_\tau|^2}{1-p_\tau} - 1\right) \text{ and } \partial_\theta^2 F_\eta(\tau)|_{\theta=\pi} = \frac{r}{2}\left(\frac{|c_\tau|^2}{p_\tau} - 1\right), \tag{C.59}$$

from which it immediately follows that

$$\partial_\theta^2(\Delta F_\eta)|_{\theta=0} \geq 0 \iff c_\sigma \leq c_\rho\sqrt{\frac{1-p_\sigma}{1-p_\rho}} \text{ and } \partial_\theta^2(\Delta F_\eta)|_{\theta=\pi} \geq 0 \iff c_\sigma \leq c_\rho\sqrt{\frac{p_\sigma}{p_\rho}}, \tag{C.60}$$

which reproduce the necessary and sufficient conditions on a time-covariant transformation from ρ to σ given in Eq. (C.55).

We can also evaluate from Eq. (C.58) that $\partial_\theta F_\eta(\tau)|_{0,\pi} = 0$. Since η is a time-symmetric state at

$\theta = 0, \pi$, this means $F_{\eta(r,0,\phi)} = \|\eta(r,0,\phi)\|_\infty = \frac{1}{2}(r+1)$ by Lemma 44. With these in hand, we straightforwardly evaluate

$$\partial_\theta^2 H_\eta(\tau)|_{\theta=0,\pi} = \partial_\theta^2(-\log F_\eta[\tau])|_{\theta=0,\pi} = -\left(\frac{2}{r+1}\right)\partial^2 F_\eta(\tau)|_{\theta=0,\pi}, \tag{C.61}$$

from which we obtain $\partial^2 \Delta H_\eta|_{\theta=0,\pi} \geq 0$ if and only if $\partial^2 \Delta F_\eta|_{\theta=0,\pi} \geq 0$, completing the proof. □

Appendix D

Appendices to Chapter 6

D.1 Coarse-graining conditions

D.1.1 Continuity of entropic relations under variations of the reference state

In this section, we will make use of the following Theorem, which was proved in Ref. [148]:

Theorem 93 ([148]). *(Continuity of the exponentiated min-entropy F). Let ρ_{RA} and σ_{RA} be two quantum states on the bipartite system RA. Then*

$$|F(\rho_{RA}) - F(\sigma_{RA})| \leq dD(\rho_{RA}, \sigma_{RA}), \tag{D.1}$$

where d is the dimension of system R.

We now use the above theorem to establish the continuity of $F_\eta(\tau)$ in the reference state η.

Lemma 94. *Let η and $\tilde{\eta}$ be two quantum states in system R where $\tilde{\eta} \in \mathcal{B}_\varepsilon(\eta)$. Then given any state τ on system A, we have that*

$$|F_\eta(\tau) - F_{\tilde{\eta}}(\tau)| \leq d\epsilon, \tag{D.2}$$

where d is the dimension of the reference system R.

Proof. By Theorem 93, we have that

$$|F_\eta(\tau) - F_{\tilde{\eta}}(\tau)| = |F(\mathcal{G}(\eta \otimes \tau)) - F(\mathcal{G}(\tilde{\eta} \otimes \tau))| \leq dD(\mathcal{G}(\eta \otimes \tau), \mathcal{G}(\tilde{\eta} \otimes \tau)). \tag{D.3}$$

Because the trace distance is contractive under quantum operations [1], we have that

$$\begin{aligned} D(\mathcal{G}(\eta_R \otimes \rho_A), \mathcal{G}(\tilde{\eta}_R \otimes \rho_A)) &\coloneqq \frac{1}{2}\|\mathcal{G}[\eta_R \otimes \rho_A] - \mathcal{G}[\tilde{\eta}_R \otimes \rho_A]\|_1 \\ &\leq \frac{1}{2}\|\eta_R \otimes \rho_A - \tilde{\eta}_R \otimes \rho_A\|_1 \\ &= \frac{1}{2}\|(\eta_R - \tilde{\eta}_R) \otimes \rho\|_1 \\ &= \frac{1}{2}\|(\eta_R - \tilde{\eta}_R)\|_1 =: D(\eta_R, \tilde{\eta}_R) \end{aligned} \tag{D.4}$$

where in the second equality we have used the identity $\|A \otimes B\|_1 = \|A\|_1\|B\|_1$ from Ref. [1]. Since $\tilde{\eta} \in \mathcal{B}_\epsilon(\eta)$ implies $D(\eta_R, \tilde{\eta}_R) \leq \epsilon$, we conclude from Eq. (D.4) that $D(\mathcal{G}(\eta_R \otimes \rho_A), \mathcal{G}(\tilde{\eta}_R \otimes \rho_A)) \leq \epsilon$. Substituting this into Eq. (D.3) leads to the lemma result. □

We now prove the continuity of $\Delta F_\eta(\rho, \sigma)$ in reference frame η as presented in the main text.

Lemma 52. *Given any input state ρ of system A, output state σ of system B, reference frame η of system R with $\dim(R) = \dim(B) \coloneqq d$ and state $\tilde{\eta} \in \mathcal{B}_\epsilon(\eta)$, we have that*

$$|\Delta F_{\tilde{\eta}}(\rho, \sigma) - \Delta F_\eta(\rho, \sigma)| \leq 2d\epsilon \coloneqq r(\epsilon). \tag{6.4}$$

Proof. We first observe that we can carry out the rearrangement

$$\begin{aligned} \Delta F_\eta(\rho, \sigma) - \Delta F_{\tilde{\eta}}(\rho, \sigma) &= [F_\eta(\rho) - F_\eta(\sigma)] - [F_{\tilde{\eta}}(\rho) - F_{\tilde{\eta}}(\sigma)] \\ &= [F_\eta(\rho) - F_{\tilde{\eta}}(\rho)] - [F_\eta(\sigma) - F_{\tilde{\eta}}(\sigma)], \end{aligned} \tag{D.5}$$

which implies by Lemma 94 that

$$|\Delta F_\eta(\rho, \sigma) - \Delta F_{\tilde{\eta}}(\rho, \sigma)| \leq |F_\eta(\rho) - F_{\tilde{\eta}}(\rho)| + |F_\eta(\sigma) - F_{\tilde{\eta}}(\sigma)| \leq 2d\epsilon, \tag{D.6}$$

completing the proof. □

The end of Appendix C.1 demonstrates that $F_\eta(\tau) > 0$ for any reference frame η and quantum state τ. As $-\log(x)$ is a continuous function on $x > 0$, we can therefore conclude that the continuity of $F_\eta(\tau)$ in reference frame shown in Lemma 94 implies the continuity of $H_\eta(\tau)$ in reference frame as well, which further implies the continuity of $\Delta H_\eta(\rho, \sigma) \coloneqq H_\eta(\sigma) - H_\eta(\rho)$ as the difference between two continuous functions in η.

D.1.2 Proof of Lemma 51

We first introduce the definition of an ϵ-net $\mathcal{N}_\epsilon$ for a subset X in a metric space $(V, \|\cdot\|)$, which is another subset in V such that every element of X is within ϵ-distance of an element in $\mathcal{N}_\epsilon$ as measured by the chosen norm.

Definition 95 (ϵ-covering). *Given a metric space $(V, \|\cdot\|)$ and $\epsilon > 0$, an ϵ-covering of the set $X \in V$ is a set $\mathcal{N}_\epsilon$ such that for any $x \in X$, there exists $y \in \mathcal{N}_\epsilon$ such that $\|x - y\| < \epsilon$.*

When V is a real vector space, we have the following upper bound on the cardinality of ϵ-coverings.

Theorem 96 ([150, 151]). *Let $\|\cdot\|$ be a norm on $\mathbb{R}^n$. Then for every $\epsilon > 0$, the unit sphere $\{x \in \mathbb{R}^n, \|x\| = 1\}$ admits an ϵ-covering $\mathcal{N}_\epsilon$ with respect to the distance measured by $\|\cdot\|$ of cardinality*

$$|\mathcal{N}_\epsilon| \leq \left(1 + \frac{2}{\epsilon}\right)^n \tag{D.7}$$

We can leverage the above theorem to show the following result from the main text:

Lemma 51. *For every $\epsilon > 0$, there exists an ϵ-covering of sufficient reference frames, i.e. a finite set of reference frames $\mathcal{N}_\epsilon$ such that every state η on a sufficient surface of reference frames $\partial\mathcal{D}$ lies within an ϵ-ball around some state in $\mathcal{N}_\epsilon$. The cardinality of $\mathcal{N}_\epsilon$ is upper-bounded as*

$$|\mathcal{N}_\epsilon| \leq \left(1 + \frac{1}{\epsilon}\right)^{d^2-1}. \tag{6.3}$$

Proof. By Theorem 47, we can choose our sufficient set of reference states to be

$$\partial\mathcal{D} = \left\{\eta(A) \coloneqq \frac{1}{d}(\mathbb{1} + A) \mid A \in \mathrm{Herm}(R) \text{ s. t. } \|A\|_\infty = 1 \text{ and } \mathrm{Tr}[A] = 0\right\}, \tag{D.8}$$

where d is the dimension of reference system R, as $\partial\mathcal{D}$ constitutes a closed surface of quantum states around the maximally mixed state.

Traceless $d \times d$ Hermitian matrices form a real linear vector space $\mathbb{R}^{d^2-1}$. We now observe that the subset of such matrices bijective to $\partial\mathcal{D}$,

$$\mathcal{S} \coloneqq \{A \in \mathrm{Herm}(R) \text{ s. t. } \|A\|_\infty = 1 \text{ and } \mathrm{Tr}[A] = 0\}, \tag{D.9}$$

constitutes a unit ball defined by $\|A\|_\infty = 1$. By Theorem 96, given any $\delta > 0$, we can find a

δ-covering $\mathcal{N}_\delta$ for $\mathcal{S}$ with cardinality bounded as

$$|\mathcal{N}_\delta| \leq \left(1 + \frac{2}{\delta}\right)^{d^2-1}. \tag{D.10}$$

Therefore, given any matrix $A \in \mathcal{S}$, there exists another matrix $A_k \in \mathcal{N}_\delta \subset \mathcal{S}$ such that $\|A - A_k\| \leq \delta$, which further implies

$$||A - A_k||_1 \leq \text{rank}(A - A_k)||A - A_k||_\infty \leq d\delta, \tag{D.11}$$

where the first inequality was proved in Ref. [149]. By the bijection $A \leftrightarrow \eta(A)$ betwee $\mathcal{S}$ and $\partial\mathcal{D}$, we conclude that given any reference state $\eta(A) \coloneqq \frac{1}{d}(\mathbb{1} + A)$ in $\partial\mathcal{D}$, there is a reference state $\eta(A_k) = \frac{1}{d}(\mathbb{1} + A_k)$ in $\partial\mathcal{D}$ where $A_k \in \mathcal{N}_\delta \subset \mathcal{S}$ such that

$$D(\eta(A), \eta(A_k)) = \frac{1}{2}||\eta(A) - \eta(A_k)||_1 = \frac{1}{2d}\|A - A_k\|_1 \leq \frac{\delta}{2}. \tag{D.12}$$

We thus conclude that if we choose $\delta = 2\epsilon$, then given any reference state $\eta(A) \in \partial\mathcal{D}$, one can find another reference state $\eta(A_k) \in \partial\mathcal{D}$ with $A_k \in \mathcal{N}_\delta$ such that $\eta(A)$ falls into the ϵ-ball around $\eta(A_k)$. Substituting $\delta = 2\epsilon$ into Eq. (D.10), we see that one can always find a finite set of reference states $\mathcal{N}_\epsilon$ with cardinality upper bounded as

$$\mathcal{N}_\epsilon \leq \left(1 + \frac{1}{\epsilon}\right)^{d^2-1} \tag{D.13}$$

such that every reference state in the surface $\partial\mathcal{D}$ of sufficient reference states lies within the ϵ-ball around a reference state from $\mathcal{N}_\epsilon$. □

D.1.3 Proof of Theorem 53

Theorem 53. *Given any $\epsilon > 0$, there is a finite set of reference frames $\mathcal{N} \coloneqq \{\eta_k\}_{k=1}^N$ with $|\mathcal{N}| \leq (1 + \frac{1}{\varepsilon})^{d^2-1}$, where d is the dimension of the reference system, such that:*

- *if $\Delta F_{\eta_k}(\rho, \sigma) < 0$ for any $\eta_k \in \mathcal{N}$ then $\rho \to \sigma$ is forbidden under G-covariant channels.*
- *if $\Delta F_{\eta_k}(\rho, \sigma) \geq r(\epsilon)$ for all $\eta_k \in \mathcal{N}$ then $\rho \to \sigma$ under a G-covariant quantum channel.*

- *For each $\eta_k \in \mathcal{N}$ where $0 \leq \Delta F_{\eta_k}(\rho, \sigma) < r(\epsilon)$ we obtain the lower bound*

$$p \geq \frac{r(\epsilon) - \Delta F_{\eta_k}(\rho, \sigma)}{\Delta F_{\eta_k}\left(\sigma, \frac{I}{d}\right)} \tag{6.5}$$

on the amount of depolarisation required to ensure $\rho \to \sigma_p$ under a G-covariant channel.

Proof. From Lemma 51, we see that it is always possible to find a finite set $\mathcal{N}$ of reference states with cardinality upper-bounded as $|\mathcal{N}| \leq (1 + \frac{1}{\epsilon})^{d^2-1}$ such that there exists a sufficient surface of states $\partial\mathcal{D}$ in which every reference state lies within an ϵ-ball of a state in $\mathcal{N}$.

Given any input state ρ and output state σ, we know that if $\Delta F_{\eta_k}(\rho, \sigma) < 0$ for any reference state $\eta_k \in \mathcal{N}$, then $\rho \to \sigma$ is forbidden under G-covariant channels. However, if we find that

$$\Delta F_{\eta_k}(\rho, \sigma) \geq r(\epsilon) \text{ for all } \eta_k \in \mathcal{N}, \tag{D.14}$$

then Lemma 52 implies that $\Delta F_\eta(\rho, \sigma) \geq 0$ for all reference frame states $\eta \in \bigcup_k \mathcal{B}_\epsilon(\eta_k)$ and therefore all reference states $\eta \in \partial\mathcal{D}$, so we conclude that the state transition $\rho \to \sigma$ must be possible under some G-covariant channel.

The final case where there exists at least one η_k for which $0 \leq \Delta F_{\eta_k}(\rho, \sigma) < r(\epsilon)$ can be handled as follows. Because

$$\mathcal{G}(\eta \otimes \sigma_p) = p\mathcal{G}\left(\eta \otimes \frac{\mathbb{1}}{d}\right) + (1-p)\mathcal{G}\left(\eta \otimes \sigma\right), \tag{D.15}$$

the convexity of F_η (Property **(P2)**) implies

$$F_\eta(\sigma_p) \leq pF_\eta\left(\frac{\mathbb{1}}{d}\right) + (1-p)F_\eta(\sigma) \tag{D.16}$$

$$\implies \Delta F_\eta(\sigma, \sigma_p) \geq p\left(F_\eta(\sigma) - F_\eta\left(\frac{\mathbb{1}}{d}\right)\right) = p\Delta F_\eta\left(\sigma, \frac{\mathbb{1}}{d}\right). \tag{D.17}$$

We then observe that if we have

$$p \geq \frac{r(\epsilon) - \Delta F_{\eta_k}(\rho, \sigma)}{\Delta F_\eta\left(\sigma, \frac{\mathbb{1}}{d}\right)}, \tag{D.18}$$

then we would obtain

$$\Delta F_{\eta_k}(\rho,\sigma_p) = \Delta F_{\eta_k}(\sigma,\sigma_p) + \Delta F_{\eta_k}(\rho,\sigma) \geq \Delta F_{\eta_k}(\rho,\sigma) + p\Delta F_{\eta_k}\left(\sigma,\frac{\mathbb{1}}{d}\right) \geq r(\epsilon). \tag{D.19}$$

We therefore conclude that if we can find a depolarisation probability p^* such that Eq. (D.18) holds for every $\eta_k \in \mathcal{N}$ for which $0 \leq \Delta F_{\eta_k}(\rho,\sigma) < r(\epsilon)$, then $\Delta F_{\eta_k}(\rho,\sigma_p) \geq r(\epsilon)$ for all $\eta_k \in \mathcal{N}$ and therefore for all $\eta \in \partial\mathcal{D}$, which ensures that the state transition $\rho \to \sigma_p$ can be carried out via a G-covariant channel. □

D.2 Properties of modes of asymmetry

D.2.1 Proof that $\mathcal{P}_k^{(\mu)}$ projects out the (μ,k) mode of asymmetry

Lemma 97. *Given any bounded operator $O \in \mathcal{B}(S)$ mode of asymmetry $(\mu,k) \in \mathrm{modes}(S)$ for a quantum system S, we have that*

$$\mathcal{P}_k^{(\mu)}(O) = O_k^{(\mu)}. \tag{D.20}$$

Proof. We show by direct calculation that

$$\begin{aligned}
\mathcal{P}_k^{(\mu)}(O) &= \int dg d_\mu u_{kk}^{(\mu^*)}(g)\, \mathcal{U}_g(O) \\
&= \int dg d_\mu u_{kk}^{(\mu^*)}(g)\, \mathcal{U}_g\left(\sum_{\lambda,j}\left(\sum_\alpha \langle X_j^{(\lambda,\alpha)}, O\rangle X_j^{(\lambda,\alpha)}\right)\right) \\
&= \sum_{\lambda,j}\left(\int dg d_\mu u_{kk}^{(\mu^*)}(g) \sum_\alpha \langle X_j^{(\lambda,\alpha)}, O\rangle \left(\sum_{j'} u_{j'j}^{(\lambda)}(g) X_{j'}^{(\lambda,\alpha)}\right)\right) \\
&= \sum_{\lambda,j}\left(\sum_{\alpha,j'}\left[\int dg d_\mu u_{kk}^{(\mu^*)}(g) u_{j'j}^{(\lambda)}(g)\right] \langle X_j^{(\lambda,\alpha)}, O\rangle X_{j'}^{(\lambda,\alpha)}\right) \\
&= \sum_{\lambda,j,j'} \delta_{\lambda,\mu}\delta_{j,k}\delta_{j',k}\left(\sum_\alpha \langle X_j^{(\lambda,\alpha)}, O\rangle X_{j'}^{(\lambda,\alpha)}\right) \\
&= \sum_\alpha \langle X_k^{(\mu,\alpha)}, O\rangle X_k^{(\mu,\alpha)} = O_k^{(\mu)},
\end{aligned} \tag{D.21}$$

where we applied the Schur orthogonality relations in Eq. (D.34) to obtain the fifth equality. □

D.2.2 Proof that distinct modes of asymmetry are orthogonal

Lemma 98. *Given any two operators A and B in $\mathcal{B}(S)$ for a quantum system S, we have that*

$$\langle A_j^{(\lambda)}, B_k^{(\mu)} \rangle = \delta_{\lambda,\mu}\delta_{j,k}\langle A_j^{(\lambda)}, B_k^{(\mu)} \rangle, \tag{D.22}$$

Proof. Because the ITO basis is orthonormal by construction, we have that

$$\langle A_j^{(\lambda)}, B_k^{(\mu)} \rangle = \left\langle \sum_\alpha \langle X_j^{(\lambda,\alpha)}, A\rangle X_j^{(\lambda,\alpha)}, \sum_\beta \langle X_k^{(\mu,\beta)}, B\rangle X_k^{(\mu,\beta)} \right\rangle \tag{D.23}$$

$$= \sum_{\alpha,\beta} \langle X_j^{(\lambda,\alpha)}, A\rangle \langle X_k^{(\mu,\beta)}, B\rangle \langle X_j^{(\lambda,\alpha)}, X_k^{(\mu,\beta)}\rangle \tag{D.24}$$

$$= \sum_{\alpha,\beta} \langle X_j^{(\lambda,\alpha)}, A\rangle \langle X_k^{(\mu,\beta)}, B\rangle \delta_{\lambda,\mu}\delta_{j,k} \langle X_j^{(\lambda,\alpha)}, X_k^{(\mu,\beta)}\rangle \tag{D.25}$$

$$= \delta_{\lambda,\mu}\delta_{j,k} \left\langle \sum_\alpha \langle X_j^{(\lambda,\alpha)}, A\rangle X_j^{(\lambda,\alpha)}, \sum_\beta \langle X_k^{(\mu,\beta)}, B\rangle X_k^{(\mu,\beta)} \right\rangle = \delta_{\lambda,\mu}\delta_{j,k}\langle A_j^{(\lambda)}, B_k^{(\mu)} \rangle, \tag{D.26}$$

which completes the proof. □

D.3 Modal sufficient conditions

D.3.1 Complex conjugate irreps

Lemma 99. *Given a complex irrep $U^{(\mu)}(g)$ of a compact group G, its complex conjugate $U^{(\mu)}(g)^*$ is also a complex irrep of G. Furthermore, two complex irreps $U^{(\mu_1)}(g)$ and $U^{(\mu_2)}(g)$ of G are equivalent if and only if $U^{(\mu_1)}(g)^*$ and $U^{(\mu_2)}(g)^*$ of G are also equivalent.*

Proof. We first prove that $U^{(\mu)}(g)^*$ is an irrep of G. Given any two elements g_1, g_2 of G, we have $U^{(\mu)}(g_1)U^{(\mu)}(g_2) = U^{(\mu)}(g_1 g_2)$. By taking the complex conjugate on both sides, we obtain $U^{(\mu)}(g_1)^* U^{(\mu)}(g_2)^* = U^{(\mu)}(g_1 g_2)^*$, so $U^{(\mu)}(g)^*$ is also a valid representation of G. If $U^{(\mu)}(g)^*$ were reducible, then there must exist a unitary operator V such that $VU^{(\mu)}(g)^*V^\dagger$ is non-trivially block-diagonal. However, the complex conjugate of this construction, $V^*U^{(\mu)}(g)V^T = V^*U^{(\mu)}(g)(V^*)^\dagger$, must therefore be non-trivially block diagonal too, so there exists a unitary operator V^* that places $U^{(\mu)}(g)$ into a non-trivial block-diagonal form. As this would generate the contradiction that $U^{(\mu)}(g)$ is also reducible, we conclude that $U^{(\mu)}(g)^*$ is irreducible.

We now assume that two complex irreps $U^{(\mu_1)}(g)$ and $U^{(\mu_2)}(g)$ of G are equivalent, i.e. there exists a similarity transformation generated by a unitary V such that $U^{(\mu_1)}(g) = VU^{(\mu_2)}(g)V^\dagger$. By taking the complex conjugate of both sides, we see that there exists a similarity transformation generate by the unitary V^* such that $U^{(\mu_1)}(g)^* = V^*U^{(\mu_2)}(g)^*V^T = V^*U^{(\mu_2)}(g)^*(V^*)^\dagger$, so $U^{(\mu_1)}(g)^*$ and $U^{(\mu_2)}(g)^*$ are equivalent. By the same reasoning, we can show that if $U^{(\mu_1)}(g)^*$ and $U^{(\mu_2)}(g)^*$ are equivalent, then $U^{(\mu_1)}(g)$ and $U^{(\mu_2)}(g)$ are also equivalent. We therefore conclude that $U^{(\mu_1)}(g)$ and $U^{(\mu_2)}(g)$ of G are equivalent if and only if $U^{(\mu_1)}(g)^*$ and $U^{(\mu_2)}(g)^*$ of G are also equivalent. □

D.3.2 Proof of Lemma 57

In this section, we prove that it is possible to choose an ITO basis that satisfies the conditions laid out in Lemma 57, which we reproduce below for convenience.

Lemma 57. *Let ρ be the input state for a G-covariant state transition from a quantum system S, and consider the representation $\mathcal{U}_g$ of a compact group G on $\mathcal{B}(S)$ whose multidimensional irreps have no multiplicity. Then there exists an ITO basis $\{X_j^{(\lambda,\alpha)}\}$ for $\mathcal{U}_g$ such that*

1. *corresponding basis elements of complex conjugate irreps are Hermitian conjugates of each other. Thus one can write*

$$X_j^{(\lambda^*,\alpha)} = X_j^{(\lambda,\alpha)\dagger}. \tag{6.19}$$

2. *in every irrep λ, the input state ρ has at most a single non-zero mode of asymmetry, which we label $(\lambda, 0)$. Thus one can write*

$$\rho_j^{(\lambda)} = \delta_{j,0}\rho_0^{(\lambda)}. \tag{6.20}$$

Proof. We first show how, given an arbitrary ITO basis for $\mathcal{U}_g$ on $\mathcal{B}(S)$, we can construct new ITO basis elements for each irrep μ appearing in $\mathcal{U}_g$ and its complex conjugate μ^* that satisfy the first condition in the Lemma statement. The method of construction depends as follows on the dimension of μ and whether μ is its own complex conjugate irrep.

Case I: μ and μ^* are distinct irreps. In this case, we are free to simply choose $X_j^{(\mu^*,\alpha)} := X_j^{(\mu,\alpha)\dagger}$.

Case II: μ and μ^* are the same one-dimensional irrep. In this case, we need to prove that,

given an arbitrary ITO basis $\{X^{(\mu,\alpha)}\}$ for the μ-irrep, we can construct a *Hermitian* ITO basis $\mathcal{S} \coloneqq \{S^{(\mu,\alpha)}\}$, which we proceed to do as follows.

1. Construct a Hermitian but potentially overcomplete basis for μ-irrep from the initial ITO basis as $\mathcal{B} \coloneqq \{(X^{(\mu,\alpha)}) + X^{(\mu,\alpha)\dagger}), i(X^{(\mu,\alpha)} - X^{(\mu,\alpha)\dagger})\}$.

2. Pick out α_μ linearly-independent non-zero elements from $\mathcal{B}$, where α_μ denotes the multiplicity of the μ-irrep, to construct a Hermitian basis $\mathcal{B}'$ for the μ-irrep that is no longer overcomplete. Let us denote the elements of $\mathcal{B}'$ by $\{X^{(\mu,\alpha)'}\}$.

3. Apply the Gram-Schmidt procedure to $\mathcal{B}'$ and generate the basis $\mathcal{B}''$ for the μ-irrep, whose elements we denote by $\{X^{(\mu,\alpha)''}\}$. By construction, $\mathcal{B}''$ constitutes an orthogonal basis for the μ-irrep, so we now prove that its elements are Hermitian by induction. Suppose that elements $1, \ldots, \beta$ of $\mathcal{B}''$ are Hermitian. Then the $\beta + 1$ element of $\mathcal{B}''$ is defined as

$$X^{(\mu,\beta+1)''} \coloneqq X^{(\mu,\beta+1)'} - \sum_{\alpha=1}^{\beta} \frac{\langle X^{(\mu,\alpha)'}, X^{(\mu,\alpha)''}\rangle}{\langle X^{(\mu,\alpha)''}, X^{(\mu,\alpha)''}\rangle} X^{(\mu,\alpha)''}. \tag{D.27}$$

   We now observe that the Hilbert-Schmidt inner product of two Hermitian operators A and B, which is simply $\langle A, B\rangle \coloneqq \mathrm{Tr}\left[A^\dagger B\right] = \mathrm{Tr}[AB]$, is always real, because we have $\mathrm{Tr}[AB] = \mathrm{Tr}\left[A^\dagger B^\dagger\right] = \mathrm{Tr}\left[B^\dagger A^\dagger\right] = \mathrm{Tr}\left[(AB)^\dagger\right] = \mathrm{Tr}[AB]^*$. As every $X^{(\mu,\alpha)'}$ is Hermitian by construction as a member of $\mathcal{B}'$, the assumption that $X^{(\mu,1)''}, \ldots, X^{(\mu,\beta)''}$ are Hermitian implies that the right hand side of Eq. (D.27) is a real linear combination of Hermitian operators, and therefore $X^{(\mu,\beta+1)''}$ must also be Hermitian. Noting that $X^{(\mu,1)''} = X^{(\mu,1)'}$ and is therefore Hermitian, we see by induction that all elements of $\mathcal{B}''$ are Hermitian.

4. Normalise each element of $\mathcal{B}''$ to obtain an orthonormal Hermitian basis for the μ-irrep, $\mathcal{S} \coloneqq \left\{\frac{X^{(\mu,\alpha)''}}{\langle X^{(\mu,\alpha)''}, X^{(\mu,\alpha)''}\rangle}\right\}$.

   We now show that $\mathcal{S}$ is a valid ITO basis for the μ irrep. To do so, we first show by induction that each element of $\mathcal{B}''$ is a (complex) linear combination of the initial ITO basis elements $\{X^{(\mu,\alpha)}\}$. Suppose that elements $1, \ldots, \beta$ of $\mathcal{B}''$ are all linear combinations of $\{X^{(\mu,\alpha)}\}$. As $X^{(\mu,\beta+1)'}$ is a linear combination of $\{X^{(\mu,\alpha)}\}$ by construction, this assumption implies that the right hand side of Eq. (D.27) is also a linear combination of $\{X^{(\mu,\alpha)}\}$, and therefore the $\beta + 1$ element of $\mathcal{B}''$ is also a linear combination of $\{X^{(\mu,\alpha)}\}$. Noting that $X^{(\mu,1)''} = X^{(\mu,1)'}$, wich is a linear combination of $\{X^{(\mu,\alpha)}\}$, we conclude by induction that every element of $\mathcal{B}''$, and therefore every element of $\mathcal{S}$, which merely scales the cor-

responding element in $\mathcal{B}''$ its norm, is a linear combination of $\{X^{(\mu,\alpha)}\}$. Because $\{X^{(\mu,\alpha)}\}$ is an ITO basis for the μ-irrep, we know that $\mathcal{U}_g(X^{(\mu,\alpha)}) = U^{(\mu)}(g)X^{(\mu,\alpha)}$, where $U^{(\mu)}(g)$ is a real number because we have $\mu^* = \mu$. As each element of $\mathcal{S}'$ is a linear combination of $\{X^{(\mu,\alpha)}\}$, we conclude that $\mathcal{U}_g(S^{(\mu,\alpha)}) = U^{(\mu)}(g)S^{(\mu,\alpha)}$, which proves that $\mathcal{S}$ is a valid ITO basis for the μ-irrep.

Case III: μ and μ^* are the same multidimensional irrep. In this case, we also need to prove that, given an arbitrary ITO basis for the μ-irrep, we can construct a *Hermitian* ITO basis. This can be done using the same method as Case II. When unitary representations $\mathcal{U}_g^A$ and $\mathcal{U}_g^B$ of the same compact group G on two different physical systems A and B both contain a single copy of μ, this procedure generates ITO bases for μ on $\mathcal{B}(A)$ and $\mathcal{B}(B)$ over which G acts equivalently. This is because any complex irrep appears on the complex span of a unique real irrep (the Hermitian matrices span a real irrep) [152].

Finally, we show that, given an ITO basis satisfying the first condition in the Lemma statement, it is possible to construct another ITO basis that also satisfies the second condition. Consider any irrep μ that appears in $\mathcal{U}_g$. When μ is one-dimensional, the second condition in the Lemma statement is trivially satisified, so we only need to consider the case where μ is multidmensional (and assumed to be without multiplicity).

Let $\{X_j^{(\mu)}\}$ and $\{X_j^{(\mu^*)}\}$ be elements of the given ITO basis for the irrep μ and its complex conjugate μ^* respectively. Consequently, we have $X_j^{(\mu^*)} = X_j^{(\mu)\dagger}$. We can then generate another pair of ITO bases for μ and μ^* defined as $\{X_j^{(\mu)'} \coloneqq \sum_i U_{ij}X_i^{(\mu)}\}$ and $\{X_j^{(\mu^*)'} \coloneqq \sum_i U_{ij}^* X_i^{(\mu^*)}\}$ respectively, where U is a unitary matrix for $\mu \neq \mu^*$ and an orthogonal matrix otherwise. By construction, we have that $X_j^{(\mu)'\dagger} = X_j^{(\mu^*)'}$. We choose the first column of U so it will rotate $X_0^{(\mu)}$ into a normalised vector parallel to the projection of ρ in the irrep subspace of μ, i.e.

$$U \coloneqq \begin{pmatrix} \frac{\langle X_0^{(\mu)}, \rho\rangle}{\sqrt{\sum_j \langle \rho_j^{(\mu)}, \rho_j^{(\mu)}\rangle}} & \cdots \\ \vdots & \\ \frac{\langle X_{d_\mu-1}^{(\mu)}, \rho\rangle}{\sqrt{\sum_j \langle \rho_j^{(\mu)}, \rho_j^{(\mu)}\rangle}} & \cdots \end{pmatrix}. \tag{D.28}$$

Consequently, in the new ITO basis $\{X_j^{(\mu)'}\}$ for the μ irrep defined by this choice of U, we have $\rho_j^{(\mu)'} = \delta_{j,0}\rho_j^{(\mu)'}$ as desired. All that remains is to show that the above choice for U has the same impact in the complex conjugate irrep μ^*. We first note that, by the cyclic property of the trace

and the hermiticity of ρ, we have

$$\begin{aligned}\langle X_j^{(\mu)}, \rho\rangle \coloneqq \mathrm{Tr}\left[X_j^{(\mu)\dagger}\rho\right] &= \mathrm{Tr}\left[X_j^{(\mu^*)}\rho\right] \\ &= \mathrm{Tr}\left[X_j^{(\mu^*)}\rho^\dagger\right] \\ &= \mathrm{Tr}\left[\left(\rho X_j^{(\mu^*)\dagger}\right)^\dagger\right] \\ &= \mathrm{Tr}\left[\rho X_j^{(\mu^*)\dagger}\right]^* = \mathrm{Tr}\left[X_j^{(\mu^*)\dagger}\rho\right]^* = \langle X_j^{(\mu^*)}, \rho\rangle \end{aligned} \tag{D.29}$$

and consequently

$$\langle \rho_j^{(\mu)}, \rho_j^{(\mu)}\rangle = \langle X_j^{(\mu)}, \rho\rangle^* \langle X_j^{(\mu)}, \rho\rangle = \langle X_j^{(\mu^*)}, \rho\rangle \langle X_j^{(\mu^*)}, \rho\rangle^* = \langle \rho_j^{(\mu^*)}, \rho_j^{(\mu^*)}\rangle. \tag{D.30}$$

We can therefore write the first column of U^* as

$$U^* \coloneqq \begin{pmatrix} \frac{\langle X_0^{(\mu)},\, \rho\rangle^*}{\sqrt{\sum_j \langle \rho_j^{(\mu)},\, \rho_j^{(\mu)}\rangle}} & \cdots \\ \vdots & \\ \frac{\langle X_{d_\mu - 1}^{(\mu)},\, \rho\rangle^*}{\sqrt{\sum_j \langle \rho_j^{(\mu)},\, \rho_j^{(\mu)}\rangle}} & \cdots \end{pmatrix} = \begin{pmatrix} \frac{\langle X_0^{(\mu^*)},\, \rho\rangle}{\sqrt{\sum_j \langle \rho_j^{(\mu^*)},\, \rho_j^{(\mu^*)}\rangle}} & \cdots \\ \vdots & \\ \frac{\langle X_{d_\mu - 1}^{(\mu^*)},\, \rho\rangle}{\sqrt{\sum_j \langle \rho_j^{(\mu^*)},\, \rho_j^{(\mu^*)}\rangle}} & \cdots \end{pmatrix} \tag{D.31}$$

from which we see that U^* rotates $X_0^{(\mu^*)}$ to a normalised vector parallel to the projection of ρ in the irrep subspace of μ^*. We therefore also have $\rho_j^{(\mu^*)\prime} = \delta_{j,0}\rho_j^{(\mu^*)\prime}$ in the new basis $\{X_j^{(\mu^*)\prime}\}$ for the irrep μ^* as desired. □

In an ITO basis $\{X_j^{(\lambda,\alpha)}\}$ satisfying the conditions laid out in Lemma 57, we have that

$$\begin{aligned}\rho_j^{(\lambda^*)} &= \sum_{\alpha=1}^{\alpha_{\lambda^*}} \langle X_j^{(\lambda^*,\alpha)}, \rho\rangle X_j^{(\lambda^*,\alpha)} \\ &= \sum_{\alpha=1}^{\alpha_\lambda} \langle X_j^{(\lambda,\alpha)\dagger}, \rho\rangle X_j^{(\lambda,\alpha)\dagger} \\ &= \sum_{\alpha=1}^{\alpha_\lambda} \langle X_j^{(\lambda,\alpha)\dagger}, \rho^\dagger\rangle X_j^{(\lambda,\alpha)\dagger} \\ &= \sum_{\alpha=1}^{\alpha_\lambda} \langle X_j^{(\lambda,\alpha)}, \rho\rangle^* X_j^{(\lambda,\alpha)\dagger} = \rho_j^{(\lambda)\dagger} \end{aligned} \tag{D.32}$$

where in going from the first to the second equality we have used the fact that an irrep and its complex conjugate have the same multiplicity, and in going from the second to the third

equality we have used the fact that ρ is Hermitian.

D.3.3 Proof of Lemma 59

We first note the following relations between the (complex) matrix components $U_{ij}^{(\lambda)}(g)$ representing the λ-irrep of G in some ITO basis. These components must obey the Schur orthogonality relations

$$\int_G dg\, U_{ip}^{(\lambda)}(g)^* U_{jq}^{(\mu)}(g) = \frac{1}{d_\lambda}\delta_{\lambda,\mu}\delta_{i,j}\delta_{p,q}. \tag{D.33}$$

The choices of ITO basis made in Lemma 57 implies that $U_{ip}^{(\lambda)}(g)^* = U_{ip}^{(\lambda^*)}(g)$. Consequently, under this choice of basis, the Schur orthogonality relations become

$$\int_G dg\, U_{ip}^{(\lambda)}(g) U_{jq}^{(\mu)}(g) = \frac{1}{d_\lambda}\delta_{\lambda^*,\mu}\delta_{i,j}\delta_{p,q}. \tag{D.34}$$

Having specialised the Schur orthogonality relations to our choice of ITO basis, we can now prove Lemma 59, which we reproduce below for convenience.

Lemma 59. *Let ρ, σ and η be quantum states of systems A, B and R respectively, where we have $\dim(\mathcal{H}_R) = \dim(\mathcal{H}_B)$. Given a symmetry group G whose multidimensional irreps have no multiplicity, we have $\rho \xrightarrow{G} \sigma$ if for any reference frame η, we can find a state $\tau(\eta)$ such that*

$$\sum_{\lambda,j} \left\langle \eta_j^{(\lambda)}, f_j^{(\lambda)}(\rho)\tau(\eta)_j^{(\lambda)} - \sigma_j^{(\lambda)} \right\rangle \geq 0, \tag{6.26}$$

where we have introduced $f_j^{(\lambda)}(\rho) \coloneqq \frac{1}{d_\lambda}\left\langle \rho_j^{(\lambda^*)}, \overline{\rho}_j^{(\lambda^*)} \right\rangle$.

Proof. By Eq. (D.32), we can express the impact of the PGM prepare-and-measure channel on the (λ, j) mode of asymemtry for ρ as

$$\begin{aligned}\Phi_{\text{pgm}}^{\eta}\left(\rho_j^{(\lambda)}\right) = \int_G dg\ \text{Tr}\left[M_{\text{pgm}}(g)\rho_j^{(\lambda)}\right]\mathcal{U}_g(\tau(\eta)) &= \int_G dg\ \left\langle \rho_j^{(\lambda)\dagger}, M_{\text{pgm}}(g)\right\rangle \mathcal{U}_g(\tau(\eta)) \\ &= \int_G dg\ \left\langle \rho_j^{(\lambda^*)}, M_{\text{pgm}}(g)\right\rangle \mathcal{U}_g(\tau(\eta)).\end{aligned} \tag{D.35}$$

Because $\mathcal{G}(\rho)$ is a symmetric state, we have that $M_{\mathrm{pgm}}(g) = \mathcal{U}_g(\overline{\rho})$, which implies

$$\langle \eta_j^{(\lambda)}, \Phi^{\eta}_{\mathrm{pgm}}(\rho_j^{(\lambda)})\rangle = \int_G dg\, \langle \eta_j^{(\lambda)}, \mathcal{U}_g(\tau(\eta))\rangle \langle \rho_j^{\lambda^*}, \mathcal{U}_g(\overline{\rho})\rangle = \left\langle \eta_j^{(\lambda)} \otimes \rho_j^{(\lambda^*)}, \int_G dg\, \mathcal{U}_g(\tau(\eta)) \otimes \mathcal{U}_g(\overline{\rho}) \right\rangle$$
$$= \left\langle \eta_j^{(\lambda)} \otimes \rho_j^{(\lambda^*)}, \mathcal{G}(\tau(\eta) \otimes \overline{\rho}) \right\rangle. \tag{D.36}$$

We now calculate the bipartite G-twirl of $\tau(\eta)$ with $\overline{\rho}$.

Momentarily adopting the simplifying notation $t_i^{(\lambda,\alpha)} := \langle X_i^{(\lambda,\alpha)}, \tau(\eta)\rangle$ and $r_j^{(\mu,\beta)} := \langle Y_j^{(\mu,\beta)}, \overline{\rho}\rangle$ for the components of $\tau(\eta)$ and $\overline{\rho}$ in the ITO bases $\{X_i^{(\lambda,\alpha)}\}$ and $\{Y_j^{(\mu,\beta)}\}$ of output system B and input system A respectively, we first show that

$$\begin{aligned}
\mathcal{G}(\tau(\eta)_i^{(\lambda)} \otimes \overline{\rho}_j^{(\mu)}) &= \int_G dg\, \mathcal{U}_g(\tau(\eta)_i^{(\lambda)}) \otimes \mathcal{U}_g(\overline{\rho}_j^{(\mu)}) \\
&= \int_G dg\, \left(\sum_\alpha t_i^{(\lambda,\alpha)} \mathcal{U}_g\left(X_i^{(\lambda,\alpha)}\right)\right) \otimes \left(\sum_\beta r_j^{(\mu,\beta)} \mathcal{U}_g\left(Y_j^{(\mu,\beta)}\right)\right) \\
&= \int_G dg\, \left(\sum_\alpha t_i^{(\lambda,\alpha)} \sum_{i'} U^{(\mu)}_{i'i}(g) X_{i'}^{(\lambda,\alpha)}\right) \otimes \left(\sum_\beta r_j^{(\mu,\beta)} \sum_{j'} U^{(\mu)}_{j'j}(g) Y_{j'}^{(\mu,\beta)}\right) \\
&= \sum_{i'j'} \left[\int_G U^{(\lambda)}_{i'i}(g) U^{(\mu)}_{j'j}(g) dg\right] \left(\sum_\alpha t_i^{(\lambda,\alpha)} X_{i'}^{(\lambda,\alpha)}\right) \otimes \left(\sum_\beta r_j^{(\mu,\beta)} Y_{j'}^{(\mu,\beta)}\right) \\
&= \sum_{i'j'} \frac{\delta_{i',j'}\delta_{i,j}\delta_{\mu^*,\lambda}}{d_\lambda} \left(\sum_\alpha t_i^{(\lambda,\alpha)} X_{i'}^{(\lambda,\alpha)}\right) \otimes \left(\sum_\beta r_j^{(\mu,\beta)} Y_{j'}^{(\mu,\beta)}\right) \\
&= \delta_{i,j}\delta_{\mu^*,\lambda} \sum_{i'} \frac{1}{d_\lambda} \left(\sum_\alpha t_i^{(\lambda,\alpha)} X_{i'}^{(\lambda,\alpha)}\right) \otimes \left(\sum_\beta r_i^{(\lambda^*,\beta)} Y_{i'}^{(\lambda^*,\beta)}\right)
\end{aligned} \tag{D.37}$$

We now observe that, by construction, the modes of asymmetry in $\overline{\rho}$ are related to those of ρ in a particularly straightforward way. Using the projector $\mathcal{P}_j^{(\lambda)}$ defined in Eq. (6.12) to project out the (λ, j) mode of asymmetry for $\overline{\rho}$, we obtain

$$\begin{aligned}
\overline{\rho}_j^{(\lambda)} = \mathcal{P}_j^{(\lambda)}(\overline{\rho}) &= \int dg d_\lambda u_{jj}^{(\lambda^*)}(g) \mathcal{U}_g(\overline{\rho}) \\
&= \int dg d_\lambda u_{jj}^{(\lambda^*)}(g) \mathcal{U}_g(\mathcal{G}(\rho)^{-\frac{1}{2}} \rho \mathcal{G}(\rho)^{-\frac{1}{2}}) \\
&= \mathcal{G}(\rho)^{-\frac{1}{2}} \left[\int dg d_\lambda u_{jj}^{(\lambda^*)}(g) \mathcal{U}_g(\rho)\right] \mathcal{G}(\rho)^{-\frac{1}{2}} \\
&= \mathcal{G}(\rho)[\mathcal{P}_j^{(\lambda)}(\rho)]\mathcal{G}(\rho)^{-\frac{1}{2}} = \mathcal{G}(\rho)^{-\frac{1}{2}}(\rho_j^{(\lambda)})\mathcal{G}(\rho)^{-\frac{1}{2}},
\end{aligned} \tag{D.38}$$

which implies that, under the choice of ITO basis made in Lemma 57, we have

$$\overline{\rho}_j^{(\lambda)} = \delta_{j,0}\overline{\rho}_0^{(\lambda)}. \tag{D.39}$$

Combining this fact with Eq. (D.37) implies that

$$\mathcal{G}(\tau(\eta)_i^{(\lambda)} \otimes \overline{\rho}_j^{(\mu)}) = \delta_{i,0}\delta_{i,j}\delta_{\mu^*,\lambda} \sum_{i'} \frac{1}{d_\lambda} \left(\sum_\alpha t_0^{(\lambda,\alpha)} X_{i'}^{(\lambda,\alpha)} \right) \otimes \left(\sum_\beta r_0^{(\lambda^*,\beta)} Y_{i'}^{(\lambda^*,\beta)} \right), \tag{D.40}$$

under the choice of ITO basis made in Lemma 57.

The bipartite G-twirl of $\tau(\eta)$ and $\overline{\rho}$ is therefore

$$\begin{aligned} \mathcal{G}(\tau(\eta) \otimes \overline{\rho}) &= \sum_{\lambda,\mu,i,j} \mathcal{G}(\tau(\eta)_i^{(\mu)} \otimes \overline{\rho}_j^{(\lambda)}) \\ &= \sum_{\lambda,i'} \frac{1}{d_\lambda} \left(\sum_\alpha t_0^{(\lambda,\alpha)} X_{i'}^{(\lambda,\alpha)} \right) \otimes \left(\sum_\beta r_0^{(\lambda^*,\beta)} Y_{i'}^{(\lambda^*,\beta)} \right). \end{aligned} \tag{D.41}$$

Plugging this back into Eq. (D.36), we obtain

$$\begin{aligned} \langle \eta_j^{(\lambda)}, \Phi_{\text{pgm}}^\eta(\rho_j^{(\lambda)}) \rangle &= \left\langle \eta_j^{(\lambda)} \otimes \rho_j^{(\lambda^*)}, \sum_{\mu,i'} \frac{1}{d_\mu} \left(\sum_\alpha t_0^{(\mu,\alpha)} X_{i'}^{(\mu,\alpha)} \right) \otimes \left(\sum_\beta r_0^{(\mu^*,\beta)} Y_{i'}^{(\mu^*,\beta)} \right) \right\rangle \\ &= \delta_{j,0} \left\langle \eta_0^{(\lambda)} \otimes \rho_0^{(\lambda^*)}, \sum_{\mu,i'} \frac{1}{d_\mu} \left(\sum_\alpha t_0^{(\mu,\alpha)} X_{i'}^{(\mu,\alpha)} \right) \otimes \left(\sum_\beta r_0^{(\mu^*,\beta)} Y_{i'}^{(\mu^*,\beta)} \right) \right\rangle \\ &= \delta_{j,0} \sum_{\mu,i'} \frac{1}{d_\mu} \left\langle \eta_0^{(\lambda)}, \left(\sum_\alpha t_0^{(\mu,\alpha)} X_{i'}^{(\mu,\alpha)} \right) \right\rangle \left\langle \rho_0^{(\lambda^*)}, \left(\sum_\beta r_0^{(\mu^*,\beta)} Y_{i'}^{(\mu^*,\beta)} \right) \right\rangle \\ &= \frac{\delta_{j,0}}{d_\lambda} \left\langle \eta_0^{(\lambda)}, \left(\sum_\alpha t_0^{(\lambda,\alpha)} X_0^{(\lambda,\alpha)} \right) \right\rangle \left\langle \rho_0^{(\lambda^*)}, \left(\sum_\beta r_0^{(\lambda^*,\beta)} Y_0^{(\lambda^*,\beta)} \right) \right\rangle \\ &= \delta_{j,0} \left(\frac{1}{d_\lambda} \langle \eta_0^{(\lambda)}, \tau(\eta)_0^{(\lambda)} \rangle \langle \rho_0^{(\lambda^*)}, \overline{\rho}_0^{(\lambda^*)} \rangle \right) \\ &= \frac{1}{d_\lambda} \langle \eta_j^{(\lambda)}, \tau(\eta)_j^{(\lambda)} \rangle \langle \rho_j^{(\lambda^*)}, \overline{\rho}_j^{(\lambda^*)} \rangle \end{aligned} \tag{D.42}$$

where we have used the fact that $\rho_j^{(\lambda)} = \delta_{j,0}\rho_0^{(\lambda)}$ under the choice of ITO basis made in Lemma 57 to arrive at the second and last equalities, and invoked the orthogonality of modes of asymmetry (Eq. (6.14)) to arrive at the third equality. Applying Corollary 3 for the choice $\Phi^\eta = \Phi_{\text{pgm}}^\eta$ using the above equation leads to the Lemma statement. □

D.3.3.1 Sufficiency of reference frames whose modes of asymmetery are a subset of the input state for state transitions covariant with Abelian groups

In the special case of Abelian groups, whose complex irreps are always one-dimensional, we see from Eq. (D.37) that, given any reference frame η and state τ, we have

$$\mathcal{G}(\eta \otimes \tau) = \sum_{\lambda} \eta^{\lambda^*} \otimes \tau^{\lambda}, \tag{D.43}$$

from which we conclude that if η contains a mode of asymmetry that does *not* appear in τ,then η^{μ^*} does not contribute to the bipartite G-twirl of η with τ.

As we saw in Section 6.3.3, a necessary condition for a G-covariant state transition is that the modes of asymmetry in the output state σ must form a subset of those in the input state ρ. Consequently, if η contains a mode of asymmetry μ that does *not* appear in ρ, then for a G-covariant transition to be possible, this mode of asymmetry also must not appear in σ, and so also cannot contribute to the bipartite G-twirl of η and σ. From Eq. (D.32), we see that, given any mode of asymmetry λ, we have that $\rho^{\lambda\dagger} = \rho^{\lambda^*}$ and $\sigma^{\lambda\dagger} = \sigma^{\lambda^*}$, so if the mode μ does not exist in ρ and σ, neither does its complex conjugate mode μ^*, so η^{μ} as well as η^{μ^*} would not contribute to the bipartite G-twirls of η with ρ and σ.

In the remainder of this section, we concentrate on the non-trivial scenario where the modes of asymmetry in the output state σ is a subset of ρ. From Theorem 47, we see that if we range over all η in a region $\mathcal{R}$ bound by a small surface enclosing the maximally mixed reference frame $\frac{\mathbb{1}}{d}$, then we obtain a complete set of entropic conditions $\{\Delta H_\eta \geq 0 | \eta \in \mathcal{R}\}$. Let μ be a mode of asymmetry in η that is not present in ρ and σ, so η^{μ} and η^{μ^*} do contribute to the G-twirl of η with the desired input and output states. If the region is chosen sufficiently small, we claim that $\eta' = \eta - c(\eta^{\mu} + \eta^{\mu^*})$ is still a valid state for $\mu \neq 0$, where we have $c = 1$ if $\mu \neq \mu^*$ and $c = \frac{1}{2}$ otherwise. To see this, firstly note that, by orthonormality, the term in the brackets is traceless and so the net result still has trace 1. Secondly, by choosing the surface appropriately, the eigenvalues of η can be chosen arbitrarily close to the uniform distribution, and those of the term in bracket made arbitrarily small. Therefore the eigenvalues of the resultant operator η' are all non-negative. We thus obtain a reference frame η' by removing the μ (and μ^* mode) entirely from η, which nevertheless gives the same bipartite G-twirls $\mathcal{G}(\eta \otimes \rho)$ and $\mathcal{G}(\eta \otimes \sigma)$ as did η. This implies it suffices to range over reference frame states η with modes the same as ρ.

D.3.4 Properties of $f_j^{(\lambda)}(\rho)$

Lemma 100. *Given any state ρ and mode of asymmetry (λ, j), the function*

$$f_j^{(\lambda)}(\rho) \coloneqq \frac{1}{d_\lambda}\langle \rho_j^{(\lambda^*)}, \overline{\rho}_j^{(\lambda^*)}\rangle = \frac{1}{d_\lambda}\operatorname{Tr}\left[\rho_j^{(\lambda)}\mathcal{G}(\rho)^{-\frac{1}{2}}\rho_j^{(\lambda^*)}\mathcal{G}(\rho)^{-\frac{1}{2}}\right] \tag{D.44}$$

where we have used Eq. (D.32) to obtain the first equality, has the following properties:

(F1) *(Invariant under complex conjugation).* $f_j^{(\lambda)}(\rho) = f_j^{(\lambda^*)}(\rho)$.

(F2) *(Non-negative).* $f_j^{(\lambda)}(\rho) \geq 0$.

(F3) *(Trivial mode).* $f^0(\rho) = 1$.

Proof. **Proof of (F1)** By the cyclic property of the trace, and that an irrep λ has the same dimension as its complex conjugate λ^*, we obtain

$$f_j^{(\lambda)}(\rho) = \frac{1}{d_\lambda}\operatorname{Tr}\left[\rho_j^{(\lambda)}\mathcal{G}(\rho)^{-\frac{1}{2}}\rho_j^{(\lambda^*)}\mathcal{G}(\rho)^{-\frac{1}{2}}\right] = \frac{1}{d_{\lambda^*}}\operatorname{Tr}\left[\rho_j^{(\lambda^*)}\mathcal{G}(\rho)^{-\frac{1}{2}}\rho_j^{(\lambda)}\mathcal{G}(\rho)^{-\frac{1}{2}}\right] = f_j^{(\lambda^*)}(\rho). \tag{D.45}$$

Proof of (F2) By the cyclic property of the trace and Eq. (D.32), we have

$$\begin{aligned}\operatorname{Tr}\left[\rho_j^{(\lambda)}\mathcal{G}(\rho)^{-\frac{1}{2}}\rho_j^{(\lambda^*)}\mathcal{G}(\rho)^{-\frac{1}{2}}\right] &= \operatorname{Tr}\left[\mathcal{G}(\rho)^{-\frac{1}{4}}\rho_j^{(\lambda)}\mathcal{G}(\rho)^{-\frac{1}{4}}\mathcal{G}(\rho)^{-\frac{1}{4}}\rho_j^{(\lambda^*)}\mathcal{G}(\rho)^{-\frac{1}{4}}\right] \\ &= \operatorname{Tr}\left[\mathcal{G}(\rho)^{-\frac{1}{4}}\rho_j^{(\lambda)}\mathcal{G}(\rho)^{-\frac{1}{4}}\mathcal{G}(\rho)^{-\frac{1}{4}}\rho_j^{(\lambda)\dagger}\mathcal{G}(\rho)^{-\frac{1}{4}}\right] =: \operatorname{Tr}\left[AA^\dagger\right]\end{aligned} \tag{D.46}$$

where we have defined $A \coloneqq \mathcal{G}(\rho)^{-\frac{1}{4}}\rho_j^{(\lambda)}\mathcal{G}(\rho)^{-\frac{1}{4}}$. Regardless of what A is, any operator of the form $AA^\dagger$ is semidefinite positive, which implies

$$\operatorname{Tr}\left[\rho_j^{(\lambda)}\mathcal{G}(\rho)^{-\frac{1}{2}}\rho_j^{(\lambda^*)}\mathcal{G}(\rho)^{-\frac{1}{2}}\right] = d_\lambda f_j^{(\lambda)}(\rho) \geq 0, \tag{D.47}$$

and immediately leads to the desired property.

Proof of (F3) Recalling from Eq. (6.18) that $\rho^{(0)} = \mathcal{G}(\rho)$, and noting that the trivial irrep is one-dimensional, we immediately obtain

$$f_0^{(0)}(\rho) = \operatorname{Tr}\left[\mathcal{G}(\rho)\mathcal{G}(\rho)^{-\frac{1}{2}}\mathcal{G}(\rho)\mathcal{G}(\rho)^{-\frac{1}{2}}\right] = \operatorname{Tr}[\mathcal{G}(\rho)] = 1. \tag{D.48}$$

which completes the proof. □

D.3.5 Proof of Lemma 60

We first note that the choice of ITO basis made in Lemma 57, and the cyclic property of the trace, imply that, given any Hermitian operator O, the ITO components on corresponding basis elements of complex conjugate irreps are likewise complex conjugates of each other because

$$\begin{aligned}\langle X_j^{(\lambda,\alpha)}, O\rangle \coloneqq \mathrm{Tr}\left[X_j^{(\lambda,\alpha)\dagger} O\right] &= \mathrm{Tr}\left[X_j^{(\lambda^*,\alpha)} O\right] \\ &= \mathrm{Tr}\left[X_j^{(\lambda^*,\alpha)} O^\dagger\right] \\ &= \mathrm{Tr}\left[\left(O X_j^{(\lambda^*,\alpha)\dagger}\right)^\dagger\right] \\ &= \mathrm{Tr}\left[O X_j^{(\lambda^*,\alpha)\dagger}\right]^* = \mathrm{Tr}\left[X_j^{(\lambda^*,\alpha)\dagger} O\right]^* = \langle X_j^{(\lambda^*,\alpha)}, O\rangle^*. \end{aligned} \tag{D.49}$$

Lemma 60. *Let ρ, σ and η be quantum states of systems A, B and R respectively, where we have $\dim(\mathcal{H}_R) = \dim(\mathcal{H}_B)$. Furthermore, let us define components for states η, $\tau(\eta)$ and σ in the ITO basis $\{X_j^{(\lambda,\alpha)}\}$ of system B as $n_j^{(\lambda,\alpha)} \coloneqq \langle X_j^{(\lambda,\alpha)}, \eta\rangle$, $t_j^{(\lambda,\alpha)} \coloneqq \langle X_j^{(\lambda,\alpha)}, \tau(\eta)\rangle$ and $s_j^{(\lambda,\alpha)} \coloneqq \langle X_j^{(\lambda,\alpha)}, \sigma\rangle$ respectively. Given a symmetry group G whose multidimensional irreps have no multiplicity, we have that $\rho \xrightarrow{G} \sigma$ if for any reference frame η, we can find a state $\tau(\eta)$ of system B such that*

$$\forall(\lambda,\alpha,j): \begin{cases} f_j^{(\lambda)}(\rho) t_j^{(\lambda,\alpha)} = s_j^{(\lambda,\alpha)} \text{ or} \\ f_j^{(\lambda)}(\rho)\left|t_j^{(\lambda,\alpha)}\right| \geq \left|s_j^{(\lambda,\alpha)}\right| \text{ and } \arg(t_j^{(\lambda,\alpha)}) = \arg(n_j^{(\lambda,\alpha)}). \end{cases} \tag{6.30}$$

Proof. Together with the fact that $\lambda \to \lambda^*$ is an isomorphism on the set of distinct irreps λ appearing the representation $G \to \mathcal{U}_g^B$ in system B, Lemma 59 showed that $\rho \xrightarrow{G} \sigma$ if, for every reference state η, we can find a quantum state $\tau(\eta)$ of system B such that

$$\frac{1}{2}\sum_{\lambda,j}\left\{\left\langle \eta_j^{(\lambda)}, f_j^{(\lambda)}(\rho)\tau(\eta)_j^{(\lambda)} - \sigma_j^{(\lambda)}\right\rangle + \left\langle \eta_j^{(\lambda^*)}, f_j^{(\lambda)}(\rho)\tau(\eta)_j^{(\lambda^*)} - \sigma_j^{(\lambda^*)}\right\rangle\right\} \geq 0 \tag{D.50}$$

where we have used $f_j^{(\lambda)}(\rho) = f_j^{(\lambda^*)}(\rho)$ by Property **(F1)** of the function $f_j^{(\lambda)}(\rho)$. We can re-express these sufficient conditions in terms of ITO basis components as

$$\frac{1}{2}\sum_{\lambda,\alpha,j}\left[n_j^{(\lambda,\alpha)^*}\left(f_j^{(\lambda)}(\rho) t_j^{(\lambda,\alpha)} - s_j^{(\lambda,\alpha)}\right) + n_j^{(\lambda^*,\alpha)^*}\left(f_j^{(\lambda)}(\rho) t_j^{(\lambda^*,\alpha)} - s_j^{(\lambda^*,\alpha)}\right)\right] \geq 0. \tag{D.51}$$

Because all quantum states are Hermitian, we conclude by Eq. (D.49) that the choice of ITO basis we made in Lemma 57 implies $n_j^{(\lambda^*,\alpha)} = n_j^{(\lambda,\alpha)^*}$, $s_j^{(\lambda^*,\alpha)} = s_j^{(\lambda,\alpha)^*}$ and $t_j^{(\lambda^*,\alpha)} = t_j^{(\lambda,\alpha)^*}$. Since $f_j^{(\lambda)}(\rho)$ is always real (by Property **(F2)**), we see that the two terms in the summand of Eq. (D.50) are complex conjugates, so we can convert the sufficient conditions in Eq. (D.50) to

$$\begin{aligned}&\sum_{\lambda,\alpha,j} \mathrm{Re}\left\{ n_j^{(\lambda,\alpha)^*} \left(f_j^{(\lambda)}(\rho) t_j^{(\lambda,\alpha)} - s_j^{(\lambda,\alpha)} \right) \right\} \\ &= \sum_{\lambda,\alpha,j} \left| n_j^{(\lambda,\alpha)} \right| \mathrm{Re}\left\{ e^{-i \arg\left[n_j^{(\lambda,\alpha)} \right]} \left[f_j^{(\lambda)}(\rho) t_j^{(\lambda,\alpha)} - s_j^{\lambda,\alpha} \right] \right\} \geq 0. \end{aligned} \tag{D.52}$$

which is satisfied if each summand is itself non-negative, i.e.

$$\forall(\lambda,\alpha,j) : \mathrm{Re}\left\{ e^{-i \arg\left[n_j^{(\lambda,\alpha)} \right]} \left[f_j^{(\lambda)}(\rho) t_j^{(\lambda,\alpha)} - s_j^{(\lambda,\alpha)} \right] \right\} \geq 0. \tag{D.53}$$

For any given (λ,α,j), one way of satisfying Eq. (D.53) is if we can choose $\tau(\eta)$ such that

$$f_j^{(\lambda)}(\rho) t_j^{(\lambda,\alpha)} = s_j^{(\lambda,\alpha)}. \tag{D.54}$$

Another way is if we can choose a $\tau(\eta)$ such that the phase of $t_j^{(\lambda,\alpha)}$ cancels that of $n_j^{(\lambda,\alpha)^*}$, i.e.

$$\arg[t_j^{(\lambda,\alpha)}] = \arg[n_j^{(\lambda,\alpha)}]. \tag{D.55}$$

Substituting this phase choice for $\tau(\eta)$ into the sufficient condition in Eq. (D.53) at a particular (λ,α,j), we obtain

$$f_j^{(\lambda)}(\rho) \left| t_j^{(\lambda,\alpha)} \right| - \left| s_j^{(\lambda,\alpha)} \right| \cos\left(\arg\left[s_j^{(\lambda,\alpha)} \right] - \arg\left[n_j^{(\lambda,\alpha)} \right] \right) \geq 0. \tag{D.56}$$

Noting that $\max_x \cos x = 1$, we see that the above equation holds for any η and σ if we can choose $\tau(\eta)$ such that

$$f_j^{(\lambda)}(\rho) \left| t_j^{(\lambda,\alpha)} \right| \geq \left| s_j^{(\lambda,\alpha)} \right|. \tag{D.57}$$

Together with Eq. (D.55), this provides a second way of satisfying Eq. (D.53) for a given (λ,α,j). Therefore, if, given any (λ,α,j), either Eq. (D.54) is true or Eq. (D.57) and Eq. (D.55) are true, then Eq. (D.53) holds, which implies the sufficient conditions identified by Lemma 59

in Eq. (D.50) hold. □

D.3.6 Proof of Theorem 61

Theorem 61. *Let ρ and σ be two states of systems A and B respectively. Given a symmetry group G whose multidimensional irreps have no multiplicity, we have $\rho \xrightarrow{G} \sigma$ if*

$$\forall \lambda \neq 0 : \begin{cases} \mu_{\min} n^{-1} f_j^{(\lambda)}(\rho) \geq g_j^{(\lambda)}(\sigma) & \text{for } j = 0 \\ \sigma_j^{(\lambda)} = 0 & \text{for } j \neq 0. \end{cases} \tag{6.32}$$

where we have defined $g_j^{(\lambda)}(\sigma) \coloneqq \sum_\alpha \left| \langle X_j^{(\lambda,\alpha)}, \sigma \rangle \right|$ for the ITO basis $\{X_j^{(\lambda,\alpha)}\}$ of system B, $\mu_{\min}$ as the smallest eigenvalue of $\mathcal{G}(\sigma)$, and n as the number of distinct non-trivial irreps appearing in the representation of G on $\mathcal{B}(B)$.

Proof. Let us define components for states η and σ, as well a linear operator $\tau(\eta)$, in the ITO basis $\{X_j^{(\lambda,\alpha)}\}$ of system B as $n_j^{(\lambda,\alpha)} \coloneqq \langle X_j^{(\lambda,\alpha)}, \eta \rangle$, $t_j^{(\lambda,\alpha)} \coloneqq \langle X_j^{(\lambda,\alpha)}, \tau(\eta) \rangle$ and $s_j^{(\lambda,\alpha)} \coloneqq \langle X_j^{(\lambda,\alpha)}, \sigma \rangle$ respectively.

We first present our choices for the ITO components of $\tau(\eta)$. Assuming that these choices produce a valid quantum state, we next show that when these choices satisfy the sufficient conditions of Lemma 60, we obtain the Theorem statement. We then conclude the proof by verifying that the choices for ITO components we have made always makes $\tau(\eta)$ a valid quantum state.

1. **For trivial mode of asymmetry $\lambda = 0$.** In this case, we simply set the trivial mode of asymmetry for $\tau(\eta)$ to match that of the output state σ, i.e.

$$\tau(\eta)^{(0)} \coloneqq \sigma^{(0)} \iff \mathcal{G}(\tau(\eta)) = \mathcal{G}(\sigma). \tag{D.58}$$

Because $f_0^{(0)}(\rho) = 1$ by Property **(F3)** of the function $f_j^{(\lambda)}(\rho)$, this choice guarantees

$$\forall \alpha : f_0^{(0)}(\rho) t^{(0,\alpha)} = s^{(0,\alpha)}, \tag{D.59}$$

which satisfies sufficient conditions on ITO components for the trivial mode of asymmetry of $\tau(\eta)$ set out in Lemma 60.

2. **For the nontrivial mode of asymmetry** $(\lambda \neq 0, 0)$**.** In this case, we choose

$$t_0^{(\lambda,\alpha)} := e^{i \arg\left[n_0^{(\lambda,\alpha)}\right]} c^{(\lambda)} \left| s_0^{(\lambda,\alpha)} \right|, \tag{D.60}$$

where we have defined

$$c^{(\lambda)} := \frac{\mu_{\min}}{n \left(\sum_\alpha \left| s_0^{(\lambda,\alpha)} \right| \right)}. \tag{D.61}$$

This choice satisfies the sufficient condition imposed on the $(\lambda, \alpha, 0)$ ITO component of $\tau(\eta)$ from Lemma 60 if

$$f_0^{(\lambda)}(\rho) c^{(\lambda)} \geq 1 \iff \mu_{\min} n^{-1} f_0^{(\lambda)}(\rho) \geq \sum_\alpha \left| s_0^{(\lambda,\alpha)} \right| = \sum_\alpha \left| \langle X_0^{(\lambda,\alpha)}, \sigma \rangle \right| =: g_0^{(\lambda)}(\sigma), \tag{D.62}$$

which leads to the sufficient conditions stated in the theorem statement for any mode of asymmetry $(\lambda \neq 0, 0)$.

Given a mode of asymmetry $(\lambda \neq 0, 0)$, the choice we made for $t_0^{(\lambda,\alpha)}$ in Eq. (D.60) becomes undefined if and only if $s_0^{(\lambda,\alpha)} = 0$ for all α. In this case, *any* choice of $t_0^{(\lambda,\alpha)}$ satisfies the sufficient conditions in Lemma 60; equivalently, the sufficient conditions in the theorem statement are satisfied if $s_0^{(\lambda,\alpha)} = 0$ for all α. We therefore clarify that $t_0^{(\lambda,\alpha)} := 0$ when $s_0^{(\lambda,\alpha)} = 0$ for all α.

3. **For the non-trivial mode of asymmetry** $(\lambda \neq 0, j \neq 0)$**.** In this case, the choice of ITO basis made in Lemma 57 imposes

$$\rho_{j\neq 0}^{(\lambda)} = 0 \iff f_{j\neq 0}^{(\lambda)}(\rho) = 0, \tag{D.63}$$

so the sufficient condition stated in Lemma 60 for the $(\lambda, \alpha, j \neq 0)$ ITO component of $\tau(\eta)$ can be satisfied if and only if

$$g_{j\neq 0}^{(\lambda)}(\sigma) = 0 \iff \forall \alpha : s_{j\neq 0}^{(\lambda,\alpha)} = 0 \iff \sigma_{j\neq 0}^{(\lambda)} = 0, \tag{D.64}$$

which leads to the sufficient condition stated in the theorem for any mode of asymmetry $(\lambda \neq 0, j \neq 0)$. Provided that the output state satisfies these conditions, we can then

choose the $(\lambda, \alpha, j \neq 0)$ component of $\tau(\eta)$ to be whatever we like, so we take

$$t_{j\neq 0}^{(\lambda,\alpha)} := 0. \tag{D.65}$$

for simplicity.

We now prove that the choices of ITO components laid out above, which produces

$$\tau(\eta) := \mathcal{G}(\sigma) + \sum_{\lambda\neq 0,\alpha} e^{i \arg\left[n_0^{(\lambda,\alpha)}\right]} c^{(\lambda)} \left|s_0^{(\lambda,\alpha)}\right| X_0^{(\lambda,\alpha)}, \tag{D.66}$$

constitutes a valid quantum state regardless of η.

We first note that, because $\mathbb{1}$ transforms trivially, it must lie in the span of ITO basis elements that transform trivially, i.e. $\mathbb{1} \in \mathrm{Span}(\{X^{0,\alpha}\})$. The orthonormality between ITO basis elements for distinct irreps then implies that ITO basis elements for *non-trivial* irreps must be *traceless*, i.e. $\mathrm{Tr}\left[X_j^{(\lambda\neq 0,\alpha)}\right] = 0$ for $\lambda \neq 0$. We thus conclude that

$$\mathrm{Tr}[\tau(\eta)] = \mathrm{Tr}[\mathcal{G}(\sigma)] = 1, \tag{D.67}$$

so it only remains to show that $\tau(\eta)$ as defined in Eq. (D.66) is positive semidefinite.

Because $\lambda \to \lambda^*$ is an isomorphism on the set of distinct irreps λ appearing the representation $G \to \mathcal{U}_g^B$ in system B, and the trivial irrep is its own complex conjugate, we can write

$$\tau(\eta) = \frac{1}{2}\left(\sum_{\lambda,\alpha,j} t_j^{(\lambda,\alpha)} X_j^{(\lambda,\alpha)} + t_j^{(\lambda^*,\alpha)} X_j^{\lambda^*,\alpha}\right) \tag{D.68}$$

$$= \mathcal{G}(\sigma) + \frac{1}{2}\left(\sum_{\lambda\neq 0,\alpha} e^{i \arg\left[n_0^{(\lambda,\alpha)}\right]} c^{(\lambda)} \left|s_0^{(\lambda,\alpha)}\right| X_0^{(\lambda,\alpha)} + e^{i \arg\left[n_0^{(\lambda^*,\alpha)}\right]} c^{(\lambda^*)} \left|s^{(\lambda^*,\alpha)}\right|_0 X_0^{(\lambda^*,\alpha)}\right)$$

$$= \mathcal{G}(\sigma) + \frac{1}{2}\left(\sum_{\lambda\neq 0,\alpha} e^{i \arg\left[n_0^{(\lambda,\alpha)}\right]} c^{(\lambda)} \left|s_0^{(\lambda,\alpha)}\right| X_0^{(\lambda,\alpha)} + e^{-i \arg\left[n_0^{(\lambda,\alpha)}\right]} c^{(\lambda)} \left|s_0^{(\lambda,\alpha)}\right| X_0^{(\lambda,\alpha)\dagger}\right) \tag{D.69}$$

We obtained the final equality by invoking the choice of ITO basis made in Lemma 57, which imposed $X_j^{(\lambda^*,\alpha)} = X_j^{(\lambda,\alpha)\dagger}$. Because σ is a quantum state, this choice implies by Eq. (D.49) that $s_0^{(\lambda^*,\alpha)} = s_0^{(\lambda,\alpha)^*}$, and consequently $\left|s_0^{(\lambda^*,\alpha)}\right| = \left|s_0^{(\lambda,\alpha)}\right|$ as well as $c^{(\lambda)} = c^{(\lambda^*)}$. Similarly, because η is a quantum state, this choice implies by Eq. (D.49) that $n_j^{(\lambda^*,\alpha)} = n_j^{(\lambda,\alpha)^*}$ by Eq. (D.49) and consequently $\arg[n_j^{(\lambda^*,\alpha)}] = -\arg[n_j^{(\lambda,\alpha)}]$.

Let $|\psi\rangle$ be any pure state of the output system B. Starting from Eq. (D.69), we obtain

$$\begin{aligned}
\langle\psi|\,\tau(\eta)\,|\psi\rangle &\geq \mu_{\min} + \frac{1}{2}\left(\sum_{\lambda\neq 0,\alpha} e^{i\arg\left[n_0^{(\lambda,\alpha)}\right]} c^{(\lambda)} \left|s_0^{(\lambda,\alpha)}\right| \langle\psi|\,X_0^{(\lambda,\alpha)}\,|\psi\rangle + e^{-i\arg\left[n_0^{(\lambda,\alpha)}\right]} c^{(\lambda)} \left|s_0^{(\lambda,\alpha)}\right| \langle\psi|\,X_0^{(\lambda,\alpha)\dagger}\,|\psi\rangle\right) \\
&= \mu_{\min} + \sum_{\lambda\neq 0,\alpha} \mathrm{Re}\left(e^{i\arg\left[n_0^{(\lambda,\alpha)}\right]} c^{(\lambda)} \left|s_0^{(\lambda,\alpha)}\right| \langle\psi|\,X_0^{(\lambda,\alpha)}\,|\psi\rangle\right) \\
&= \mu_{\min} + \sum_{\lambda\neq 0,\alpha} c^{(\lambda)} \left|s_0^{(\lambda,\alpha)}\right| \left|\langle\psi|\,X_0^{(\lambda,\alpha)}\,|\psi\rangle\right| \cos\left(\arg\left[n_0^{(\lambda,\alpha)}\right] + \arg\left[\langle\psi|\,X_0^{(\lambda,\alpha)}\,|\psi\rangle\right]\right) \\
&\geq \mu_{\min} - \sum_{\lambda\neq 0,\alpha} c^{(\lambda)} \left|s_0^{(\lambda,\alpha)}\right| \left|\langle\psi|\,X_0^{(\lambda,\alpha)}\,|\psi\rangle\right| \qquad \text{(D.70)}
\end{aligned}$$

as $\min_x \cos(x) = -1$. Because $\langle X_j^{(\lambda,\alpha)}, X_j^{(\lambda,\alpha)}\rangle = 1$ by construction, and because, given any bounded operator A, we know that $A^\dagger A$ is always positive semidefinite, we conclude that $X_j^{(\lambda,\alpha)\dagger} X_j^{(\lambda,\alpha)}$ is diagonalisable with eigenvalues between 0 and 1. Therefore

$$\left|X_j^{(\lambda,\alpha)}\,|\psi\rangle\right|^2 = \langle\psi|\,X_j^{(\lambda,\alpha)\dagger} X_j^{(\lambda,\alpha)}\,|\psi\rangle \leq \left\|X_j^{(\lambda,\alpha)\dagger} X_j^{(\lambda,\alpha)}\right\|_\infty \leq 1. \qquad \text{(D.71)}$$

By the Cauchy-Schwarz inequality, we then have

$$\left|\langle\psi|\,X_j^{(\lambda,\alpha)}\,|\psi\rangle\right| \leq ||\psi\rangle| \times \left|X_j^{(\lambda,\alpha)}\,|\psi\rangle\right| \leq 1, \qquad \text{(D.72)}$$

which can then be used to lower-bound Eq. (D.70) as

$$\begin{aligned}
\langle\psi|\,\tau(\eta)\,|\psi\rangle &\geq \mu_{\min} - \sum_{\lambda\neq 0,\alpha} c^{(\lambda)} \left|s_0^{(\lambda,\alpha)}\right| \\
&= \mu_{\min} - \sum_{\lambda\neq 0,\alpha} \frac{\mu_{\min}}{n\left(\sum_{\alpha'} \left|s_0^{(\lambda,\alpha')}\right|\right)} \left|s_0^{(\lambda,\alpha)}\right| \\
&= \mu_{\min} - \sum_{\lambda\neq 0} \frac{\mu_{\min}}{n} \\
&= \mu_{\min} - \mu_{\min} = 0, \qquad \text{(D.73)}
\end{aligned}$$

which confirms that $\tau(\eta)$ is positive semidefinite regardless of η. □

D.3.7 Conditions for identical input and output systems

We begin with the following analogue of Lemma 58, which was established at the start of Section 6.4.4.

Lemma 101. *Let ρ and σ be states of a system S, and η be a state of reference system R where we have $\dim(\mathcal{H}_R) = \dim(\mathcal{H}_S)$. Then $\rho \xrightarrow{G} \sigma$ if there exists a family of G-covariant channels $\{\Phi^\eta | \eta \in \mathcal{D}(R)\}$ from S to itself and a probability $q \in [0,1]$ such that*

$$\forall \eta : \langle \eta, q\Phi^\eta(\rho) \rangle \geq \langle \eta, \sigma(q) \rangle. \tag{D.74}$$

Recalling that distinct modes of asymmetry cannot be mixed and are orthogonal, we see that Lemma 101 implies the following analogue of Corollary 3.

Corollary 5. *Let ρ and σ be states of system S, and η be a state of reference system R where we have $\dim(\mathcal{H}_R) = \dim(\mathcal{H}_S)$. Then $\rho \xrightarrow{G} \sigma$ if for any reference frame η, we can find a G-covariant channel Φ^η from S to itself and a probability $q \in [0,1]$ such that*

$$\sum_{\lambda,j} \left\langle \eta_j^{(\lambda)}, q\Phi^\eta \left(\rho_j^{(\lambda)} \right) - \sigma(q)_j^{(\lambda)} \right\rangle \geq 0. \tag{D.75}$$

Corollary 5 is identical to Corollary 3 upon substituting $\Phi^\eta \to q\Phi^\eta$ and $\sigma \to \sigma(q)$ at any probability $q \in [0,1]$. Running through the proof of Lemma 59 from Corollary 5 instead of Corollary 3 using these substitutions, we arrive at the following analogue of Lemma 59.

Lemma 102. *Let ρ and σ be states of system S, and let η be a state of reference system R where we have $\dim(\mathcal{H}_R) = \dim(\mathcal{H}_B)$. Given a symmetry group G whose multidimensional irreps have no multiplicity, we have $\rho \xrightarrow{G} \sigma$ if for any reference frame η, we can find a state $\tau(\eta)$ and a probability $q \in [0,1]$ such that*

$$\sum_{\lambda,j} \left\langle \eta_j^{(\lambda)}, f_j^{(\lambda)}(\rho) \left(q\tau(\eta)_j^{(\lambda)} \right) - \sigma(q)_j^{(\lambda)} \right\rangle \geq 0. \tag{D.76}$$

Lemma 102 is identical to Lemma 59 upon substituting $\tau(\eta) \to q\tau(\eta)$ and $\sigma \to \sigma(q)$ at any probability $q \in [0,1]$. Running through the proof of Lemma 60 from Lemma 102 using these substitutions, for which we note that $\sigma(q)$ is still Hermitian even though it is not necessarily a quantum state, we obtain the following analogue of Lemma 60.

Lemma 103. *Let ρ and σ be states of a system S, and let η be a state of reference system R where $\dim(\mathcal{H}_R) = \dim(\mathcal{H}_S)$. Furthermore, let us define components for states η, $\tau(\eta)$ and σ in the ITO basis $\{X_j^{(\lambda,\alpha)}\}$ of system S as $n_j^{(\lambda,\alpha)} := \langle X_j^{(\lambda,\alpha)}, \eta \rangle$, $t_j^{(\lambda,\alpha)} := \langle X_j^{(\lambda,\alpha)}, \tau(\eta) \rangle$ and $s(q)_j^{(\lambda,\alpha)} := \langle X_j^{(\lambda,\alpha)}, \sigma(q) \rangle$ respectively. Given a symmetry group G whose multidimensional irreps have no multiplicity, we have*

that $\rho \xrightarrow{G} \sigma$ *if for any reference frame* η*, we can find a state* $\tau(\eta)$ *of system* S *and a probability* $q \in [0,1]$ *such that*

$$\forall(\lambda,\alpha,j): \begin{cases} f_j^{(\lambda)}(\rho)(qt_j^{(\lambda,\alpha)}) = s(q)_j^{(\lambda,\alpha)} \text{ or} \\ f_j^{(\lambda)}(\rho)\left|qt_j^{(\lambda,\alpha)}\right| \geq \left|s(q)_j^{(\lambda,\alpha)}\right| \text{ and } \arg(t_j^{(\lambda,\alpha)}) = \arg(n_j^{(\lambda,\alpha)}). \end{cases} \tag{D.77}$$

$t_j^{(\lambda,\alpha)} \to qt_j^{\lambda,\alpha}$ and $s_j^{(\lambda,\alpha)} \to s(q)_j^{(\lambda,\alpha)}$.

We are now in a position to derive the following analogue of Theorem 61.

Theorem 104. *Let* ρ *and* σ *be states of a system* S*. Given a symmetry group* G *whose multidimensional irreps have no multiplicity, we have* $\rho \xrightarrow{G} \sigma$ *if* $\sigma = \rho$ *or if there exists a probability* $q \in (q^*, 1]$

$$\forall\lambda \neq 0: \begin{cases} \mu_{\min}(q)n^{-1}f_0^{(\lambda)}(\rho) \geq g_0^{(\lambda)}(\sigma(q)) & \text{for } j=0 \\ \sigma_j^{(\lambda)} = 0 & \text{for } j \neq 0. \end{cases} \tag{D.78}$$

where we have defined $q^* \coloneqq \min\{q \in [0,1] | \mathcal{G}(\sigma(q)) \geq 0\}$*,* $g_j^{(\lambda)}(\sigma(q)) \coloneqq \sum_\alpha \left|\langle X_j^{(\lambda,\alpha)}, \sigma(a)\rangle\right|$ *for the ITO basis* $\{X_j^{(\lambda,\alpha)}\}$ *of system* S*,* $\mu_{\min}(q)$ *as the smallest eigenvalue of* $\mathcal{G}(\sigma(q))$*, and* n *as the number of distinct non-trivial irreps appearing in the representation of* G *on* $\mathcal{B}(S)$*.*

Proof. The only covariantly accessible state at $q = 0$ from ρ is ρ itself. At any other probability $q \in (0,1]$, the proof follows that of Theorem 61 with the substitutions $\tau(\eta) \to q\tau(\eta)$ and $\sigma \to \sigma(q)$ except for the following caveat. The part of that proof demonstrating why $\tau(\eta)$ is positive semidefinite relies on $\mathcal{G}(\tau(\eta))$ being positive semidefinite, which may not be true for all $q \in (0,1]$ since we now assign $\mathcal{G}(\tau(\eta)) \coloneqq q^{-1}\mathcal{G}(\sigma(q))$, which is semidefinite positive if and only if $\mathcal{G}(\sigma(q))$ is semidefinite positive. We therefore restrict ourselves to the interval of $q \in (0,1]$ in which $\mathcal{G}(\sigma(q)) \geq 0$ by finding the minimum probability q^* at which this is achieved and noting that if $\mathcal{G}((\sigma(q^*))) \geq 0$, then at any higher probability $q > q^*$ we also have $\mathcal{G}(\sigma(q)) = \mathcal{G}([\sigma - (1-q^*) + (q^*-q))\rho]) = \mathcal{G}([\sigma - (1-q^*)\rho + (q-q^*)\rho]) = \mathcal{G}(\sigma(q^*)) + (q-q^*)\mathcal{G}(\rho) \geq 0$ as the sum of two semidefinite positive operators. We conclude this proof by noting that q^* always exists since at $q = 1$ we have $\mathcal{G}(\sigma(1)) = \mathcal{G}(\sigma) \geq 0$. □

By applying Theorem 104 to a partially depolarised version of σ, $\sigma_p \coloneqq (1-p)\sigma + p\frac{I}{d}$, we arrive at Theorem 63.

D.4 Truncation of output system

The following two Lemmas detail when one can truncate the Hilbert space of the output system without affecting the possibility of G-covariant interconversion to a particular output state σ. This is of use in our analysis of state interconversion with partial depolarisation.

Lemma 105. *Let ρ be a state of the input system A, associated to the Hilbert space $\mathcal{H}_A$. Let σ be a state of the output system B, associated to the Hilbert space $\mathcal{H}_B$.*

Let $\mathcal{H}_S$ be any subspace of $\mathcal{H}_B$ with the following two properties:

1. *$\mathcal{H}_S$ carries its own representation of G, i.e. $\mathcal{H}_B$ can be decomposed into $\mathcal{H}_B = \mathcal{H}_S \oplus \mathcal{H}_{S^\perp}$ such that $U_B(g) = U_S(g) \oplus U_{S^\perp}(g)$.*
2. *The support of σ is contained entirely within $\mathcal{H}_S$, i.e. letting Π_S be the projector onto $\mathcal{H}_S$, $\Pi_S \sigma \Pi_S = \sigma$.*

Then there exists a G-covariant operation from A to B that takes ρ to σ if and only if there exists a G–covariant operation from A to S that takes ρ to σ.

Proof. Let us first assume that there exists a G–covariant operation $\mathcal{E}_{\text{cov}} : \mathcal{B}(A) \to \mathcal{B}(B)$ such that $\mathcal{E}_{\text{cov}}(\rho) = \sigma$. We then observe that

$$\begin{aligned} \Pi_S U_B(g) &= (\mathbb{1}_S \oplus 0_{S^\perp})\,(U_S(g) \oplus U_{S^\perp}(g)) \\ &= U_S(g) \oplus 0_{S^\perp} \\ &= (U_S(g)\mathbb{1}_S) \oplus 0_{S^\perp} = U_S(g)\Pi_S \end{aligned} \tag{D.79}$$

Therefore, $\Pi_S(\cdot)\Pi_S$ is a covariant map from $\mathcal{B}(B)$ to $\mathcal{B}(S)$. As a result, $\Pi_S[\mathcal{E}_{\text{cov}}(\cdot)]\Pi_S$ is a covariant operation from $\mathcal{B}(A)$ to $\mathcal{B}(S)$ such that

$$\Pi_S \mathcal{E}_{\text{cov}}(\rho)\Pi_S = \Pi_S \sigma \Pi_S = \sigma. \tag{D.80}$$

Conversely, let us now assume that there exists a covariant transformation $\mathcal{F}_{\text{cov}} : \mathcal{B}(A) \to \mathcal{B}(S)$ such that $\mathcal{F}_{\text{cov}}(\rho) = \sigma$. We then extend $\mathcal{H}_S$ into the bigger Hilbert space $\mathcal{H}_B = \mathcal{H}_S \oplus \mathcal{H}_{S^\perp}$ such that $\mathcal{H}_S$ still forms its own representation of G, e.g. as $U_B(g) = U_S(g) \oplus I_{S^\perp}$. Then $\mathcal{F}_{\text{cov}}$ can be reinterpreted as a covariant channel from $\mathcal{B}(A)$ to $\mathcal{B}(B)$.

We therefore conclude that G-covariant interconversion from ρ to σ is unaffected by treating σ as a state of B or as a state of S. □

Lemma 106. *Given any particular output state σ, it is always possible to truncate the Hilbert space of the output system, $\mathcal{H}_B$, to the support of $\mathcal{G}(\sigma)$ without affecting the possibility of G–covariant interconversion.*

Proof. The representation of G on $\mathcal{H}_B$ splits up in the following way [7]:

$$\mathcal{H}_B = \bigoplus_q \mathcal{H}_q. \tag{D.81}$$

The $\mathcal{H}_q$ are known as the *charge sectors* of $\mathcal{H}_B$, and they each carry an *inequivalent* representation of G. Each $\mathcal{H}_q$ can be further decomposed into a tensor product

$$\mathcal{H}_q = \mathcal{M}_q \otimes \mathcal{N}_q. \tag{D.82}$$

The $\mathcal{M}_q$ carry inequivalent *irreps* of G, while the $\mathcal{N}_q$ carry *trivial* representations of G (so $\dim(\mathcal{N}_q)$ yields the multiplicity of q in the irrep decomposition for the representation of G on $\mathcal{H}_B$). This means every element g is represented on $\mathcal{H}_q$ in the form $U_{\mathcal{M}_q}(g) \otimes \mathbb{1}_{\mathcal{N}_q}$. As a result, given any pure state $|\psi_q\rangle$ in $\mathcal{N}_q$, $\mathcal{M}_q \otimes \mathrm{Span}(|\psi_q\rangle)$ is an irrep of G. Projectors onto irreps of G thus take the form $\mathbb{1}_{\mathcal{M}_q} \otimes |\psi_q\rangle\langle\psi_q|$.

We note the following properties about the projector $\mathbb{1}_{\mathcal{M}_q} \otimes |\psi_q\rangle\langle\psi_q|$. Because $\mathcal{M}_q \otimes \mathrm{Span}\left(|\psi_q\rangle\right)$ is a subspace of $\mathcal{H}_q$,

$$\left[\Pi_q, \mathbb{1}_{\mathcal{M}_q} \otimes |\psi_q\rangle\langle\psi_q|\right] = 0 \tag{D.83}$$

For the same reason, Π_q is identity on $\mathcal{M}_q \otimes \mathrm{Span}\left(|\psi_q\rangle\right)$, which means

$$\mathbb{1}_{M_q} \otimes |\psi_q\rangle\langle\psi_q| = \Pi_q \left(\mathbb{1}_{\mathcal{M}_q} \otimes |\psi_q\rangle\langle\psi_q|\right). \tag{D.84}$$

A subspace $\mathcal{H}$ of $\mathcal{H}_B$ lies inside the kernel of σ if and only if $\mathrm{Tr}[\Pi\sigma] = 0$, where Π is the projector

onto $\mathcal{H}$. Therefore, the irrep $\mathcal{M}_q \otimes \mathrm{Span}\left(|\psi_q\rangle\right)$ lies in the kernel of σ if and only if

$$\begin{aligned}
\mathrm{Tr}\left[\left(\mathbb{1}_{\mathcal{M}_q} \otimes |\psi_q\rangle\langle\psi_q|\right)\sigma\right] &= \mathrm{Tr}\left[\Pi_q\left(\mathbb{1}_{\mathcal{M}_q} \otimes |\psi_q\rangle\langle\psi_q|\right)\sigma\right] \\
&= \mathrm{Tr}\left[\Pi_q\Pi_q\left(\mathbb{1}_{\mathcal{M}_q} \otimes |\psi_q\rangle\langle\psi_q|\right)\sigma\right] \\
&= \mathrm{Tr}\left[\Pi_q\left(\mathbb{1}_{\mathcal{M}_q} \otimes |\psi_q\rangle\langle\psi_q|\right)\sigma\Pi_q\right] \\
&= \mathrm{Tr}\left[\left(\mathbb{1}_{\mathcal{M}_q} \otimes |\psi_q\rangle\langle\psi_q|\right)\Pi_q\sigma\Pi_q\right] \\
&= \langle\psi_q|\,\mathrm{Tr}_{\mathcal{M}_q}[\Pi_q\sigma\Pi_q]\,|\psi_q\rangle = 0
\end{aligned} \tag{D.85}$$

where in the first equality we made use of Eq. (D.84), and in the fourth equality we made use of Eq. (D.83). This short calculation means $\mathcal{M}_q \otimes \mathrm{Span}(|\psi_q\rangle)$ lies inside the kernel of σ if and only if $|\psi_q\rangle$ lies inside the kernel of $\mathrm{Tr}_{\mathcal{M}_q}[\Pi_q\sigma\Pi_q]$.

Let $\{|\psi_{q,i}\rangle\}$ be an orthonormal basis for $\mathcal{N}_q$ in which $\mathrm{Tr}_{\mathcal{M}_q}[\Pi_q\sigma\Pi_q]$ is *diagonalised*. One possible irrep decomposition for $\mathcal{H}_B$ is then

$$\mathcal{H}_B = \bigoplus_{q,i} \mathcal{M}_q \otimes \mathrm{Span}\left(|\psi_{q,i}\rangle\right). \tag{D.86}$$

An irrep in this decomposition lies inside the kernel of σ if and only if $|\psi_{q,i}\rangle$ is a basis element for the kernel of $\mathrm{Tr}_{\mathcal{M}_q}[\Pi_q\sigma\Pi_q]$. This means

$$\mathcal{H}_S^{\perp} \coloneqq \bigoplus_{q,\ |\psi_{q,i}\rangle \in \ker(\mathrm{Tr}_{\mathcal{M}_q}[\Pi_q\sigma\Pi_q])} \mathcal{M}_q \otimes \mathrm{Span}\left(|\psi_{q,i}\rangle\right) = \bigoplus_q \mathcal{M}_q \otimes \ker(\mathrm{Tr}_{\mathcal{M}_q}[\Pi_q\sigma\Pi_q]) \tag{D.87}$$

must lie inside the kernel of σ on $\mathcal{H}_B$. Conversely, the support of σ must lie inside the subspace of $\mathcal{H}_B$ that is orthogonal to $\mathcal{H}_S^{\perp}$, i.e.

$$\mathcal{H}_S = \bigoplus_q \mathcal{M}_q \otimes \mathrm{supp}(\mathrm{Tr}_{\mathcal{M}_q}[\Pi_q\sigma\Pi_q]). \tag{D.88}$$

As we see from the above equation, $\mathcal{H}_S$ is also a direct sum over irreps of G and so carries its own representation of G. Thus by Lemma 105, the possibility of interconversion is unaffected if we truncate $\mathcal{H}_B$ to $\mathcal{H}_S$.

The action of the G-twirl is given by [7]:

$$\mathcal{G} = \sum_q (\mathcal{D}_{\mathcal{M}_q} \otimes \mathcal{I}_{\mathcal{N}_q}) \circ \mathcal{P}_q, \tag{D.89}$$

where $\mathcal{P}_q := \Pi_q(\cdot)\Pi_q$ is the projector onto the charge sector $\mathcal{H}_q$, $\mathcal{D}_{\mathcal{M}_q}$ is the completely depolarising channel on $\mathcal{M}_q$ and $I_{\mathcal{N}_q}$ is the identity channel on $\mathcal{N}_q$. Therefore,

$$\mathcal{G}(\sigma) = \sum_q \frac{\mathbb{1}}{d_{\mathcal{M}_q}} \otimes \mathrm{Tr}_{\mathcal{M}_q}[\Pi_q \sigma \Pi_q], \tag{D.90}$$

where $d_{\mathcal{M}_q}$ is the dimension of $\mathcal{M}_q$. Looking back at Equation (D.88), we see that $\mathcal{H}_S = \mathrm{supp}[\mathcal{G}(\sigma)]$. It is therefore always possible to truncate the output Hilbert space to the support of $\mathcal{G}(\sigma)$ without affecting possibility of interconversion. □

References

[1] M. A. Nielsen and I. L. Chuang. *Quantum Computation and Quantum Information: 10th Anniversary Edition*. Cambridge University Press, 2012.

[2] W. K. Wootters and W. H. Zurek. *A Single Quantum Cannot Be Cloned*. *Nature* 299 (1982).

[3] P. W. Shor. *Polynomial-Time Algorithms for Prime Factorization and Discrete Logarithms on a Quantum Computer*. *SIAM Journal on Computing* 26 (1997).

[4] C. H. Bennett and G. Brassard. *Quantum Cryptography: Public Key Distribution and Coin Tossing*. *Theoretical Computer Science* 560 (2014).

[5] E. Chitambar and G. Gour. *Quantum Resource Theories*. *Reviews of Modern Physics* 91 (2019).

[6] R. Horodecki, P. Horodecki, M. Horodecki, and K. Horodecki. *Quantum Entanglement*. *Reviews of Modern Physics* 81 (2009).

[7] S. D. Bartlett, T. Rudolph, and R. W. Spekkens. *Reference Frames, Superselection Rules, and Quantum Information*. *Reviews of Modern Physics* 79 (2007).

[8] A. Kitaev. *Fault-Tolerant Quantum Computation by Anyons*. *Annals of Physics* 303 (2003).

[9] E. Dennis, A. Kitaev, A. Landahl, and J. Preskill. *Topological Quantum Memory*. *Journal of Mathematical Physics* 43 (2002).

[10] M. H. Freedman. *Quantum Computation and the Localization of Modular Functors*. *Foundations of Computational Mathematics* 1 (2001).

[11] S. Aaronson and D. Gottesman. *Improved Simulation of Stabilizer Circuits*. *Physical Review A* 70 (2004).

[12] D. Gottesman. *The Heisenberg Representation of Quantum Computers*. 1998. arXiv: quant-ph/9807006.

[13] D. Gottesman. *Stabilizer Codes and Quantum Error Correction*. PhD thesis. California Institute of Technology, 1997. arXiv: quant-ph/9705052.

[14] V. Veitch, S. A. H. Mousavian, D. Gottesman, and J. Emerson. *The Resource Theory of Stabilizer Quantum Computation*. *New Journal of Physics* 16 (2014).

[15] C. E. Shannon. *A Mathematical Theory of Communication*. *Bell System Technical Journal* 27 (1948).

[16] M. Müller-Lennert, F. Dupuis, O. Szehr, S. Fehr, and M. Tomamichel. *On Quantum Rényi Entropies: A New Generalization and Some Properties*. *Journal of Mathematical Physics* 54 (2013).

[17] M. Tomamichel, M. Berta, and M. Hayashi. *Relating Different Quantum Generalizations of the Conditional Rényi Entropy*. *Journal of Mathematical Physics* 55 (2014).

[18] S. Bravyi and A. Kitaev. *Universal Quantum Computation with Ideal Clifford Gates and Noisy Ancillas*. *Physical Review A* 71 (2005).

[19] P. Sales Rodriguez, J. M. Robinson, P. N. Jepsen, Z. He, C. Duckering, C. Zhao, K.-H. Wu, J. Campo, K. Bagnall, M. Kwon, T. Karolyshyn, P. Weinberg, M. Cain, S. J. Evered, A. A. Geim, M. Kalinowski, S. H. Li, T. Manovitz, J. Amato-Grill, J. I. Basham, L. Bernstein, B. Braverman, A. Bylinskii, A. Choukri, R. J. DeAngelo, F. Fang, C. Fieweger, P. Frederick, D. Haines, M. Hamdan, J. Hammett, N. Hsu, M.-G. Hu, F. Huber, N. Jia, D. Kedar, M. Kornjača, F. Liu, J. Long, J. Lopatin, P. L. S. Lopes, X.-Z. Luo, T. Macrì, O. Marković, L. A. Martínez-Martínez, X. Meng, S. Ostermann, E. Ostroumov, D. Paquette, Z. Qiang, V. Shofman, A. Singh, M. Singh, N. Sinha, H. Thoreen, N. Wan, Y. Wang, D. Waxman-Lenz, T. Wong, J. Wurtz, A. Zhdanov, L. Zheng, M. Greiner, A. Keesling, N. Gemelke, V. Vuletić, T. Kitagawa, S.-T. Wang, D. Bluvstein, M. D. Lukin, A. Lukin, H. Zhou, and

S. H. Cantú. *Experimental Demonstration of Logical Magic State Distillation*. *Nature* 645 (2025).

[20] B. W. Reichardt. *Quantum Universality from Magic States Distillation Applied to CSS Codes*. *Quantum Information Processing* 4 (2005).

[21] S. Bravyi and J. Haah. *Magic-State Distillation with Low Overhead*. *Physical Review A: Atomic, Molecular, and Optical Physics* 86 (2012).

[22] J. Haah, M. B. Hastings, D. Poulin, and D. Wecker. *Magic State Distillation with Low Space Overhead and Optimal Asymptotic Input Count*. *Quantum* 1 (2017).

[23] M. B. Hastings and J. Haah. *Distillation with Sublogarithmic Overhead*. *Physical Review Letters* 120 (2018).

[24] A. R. Calderbank and P. W. Shor. *Good Quantum Error-Correcting Codes Exist*. *Physical Review A: Atomic, Molecular, and Optical Physics* 54 (1996).

[25] A. Steane. *Multiple-Particle Interference and Quantum Error Correction*. *Proceedings of the Royal Society of London. Series A: Mathematical, Physical and Engineering Sciences* 452 (1996).

[26] D. Litinski. *Magic State Distillation: Not as Costly as You Think*. *Quantum* 3 (2019).

[27] N. Koukoulekidis and D. Jennings. *Constraints on Magic State Protocols from the Statistical Mechanics of Wigner Negativity*. *npj Quantum Information* 8 (2022).

[28] E. Noether. *Invarianten Beliebiger Differentialausdrücke*. *Nachrichten Von Der Gesellschaft Der Wissenschaften Zu Göttingen, Mathematisch-physikalische Klasse* 1918 (1918).

[29] S. Weinberg. *The Quantum Theory of Fields*. 1st ed. Cambridge University Press, 1995.

[30] I. Marvian and R. W. Spekkens. *Extending Noether's Theorem by Quantifying the Asymmetry of Quantum States*. *Nature Communications* 5 (2014).

[31] J. A. Vaccaro, F. Anselmi, H. M. Wiseman, and K. Jacobs. *Tradeoff between Extractable Mechanical Work, Accessible Entanglement, and Ability to Act as a Reference System, under Arbitrary Superselection Rules*. *Physical Review A: Atomic, Molecular, and Optical Physics* 77 (2008).

[32] G. Gour, I. Marvian, and R. W. Spekkens. *Measuring the Quality of a Quantum Reference Frame: The Relative Entropy of Frameness*. *Physical Review A: Atomic, Molecular, and Optical Physics* 80 (2009).

[33] I. Marvian and R. W. Spekkens. *The Theory of Manipulations of Pure State Asymmetry: I. Basic Tools, Equivalence Classes and Single Copy Transformations*. *New Journal of Physics* 15 (2013).

[34] R. Takagi. *Skew Informations from an Operational View via Resource Theory of Asymmetry*. *Scientific Reports* 9 (2019).

[35] I. Marvian. *Coherence Distillation Machines Are Impossible in Quantum Thermodynamics*. *Nature Communications* 11 (2020).

[36] G. Gour, D. Jennings, F. Buscemi, R. Duan, and I. Marvian. *Quantum Majorization and a Complete Set of Entropic Conditions for Quantum Thermodynamics*. *Nature Communications* 9 (2018).

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

[38] A. Peres, ed. *Quantum Theory: Concepts and Methods*. Dordrecht: Springer Netherlands, 2002.

[39] W. F. Stinespring. *Positive Functions on C*-Algebras*. *Proceedings of the American Mathematical Society* 6 (1955).

[40] M.-D. Choi. *Completely Positive Linear Maps on Complex Matrices*. *Linear Algebra and Its Applications* 10 (1975).

[41] A. Jamiolkowski. *Linear Transformations Which Preserve Trace and Positive Semidefiniteness of Operators*. *Reports on Mathematical Physics* 3 (1972).

[42] M. Plenio and S. Virmani. *An Introduction to Entanglement Measures*. *Quantum Information and Computation* 7 (2007).

[43] J. Aberg. *Quantifying Superposition*. 2006. arXiv: `quant-ph/0612146`.

[44] T. Baumgratz, M. Cramer, and M. B. Plenio. *Quantifying Coherence*. *Physical Review Letters* 113 (2014).

[45] A. Streltsov, G. Adesso, and M. B. Plenio. *Colloquium: Quantum Coherence as a Resource*. *Reviews of Modern Physics* 89 (2017).

[46] F. G. S. L. Brandão, M. Horodecki, J. Oppenheim, J. M. Renes, and R. W. Spekkens. *Resource Theory of Quantum States out of Thermal Equilibrium*. *Physical Review Letters* 111 (2013).

[47] M. Horodecki and J. Oppenheim. *Fundamental Limitations for Quantum and Nanoscale Thermodynamics*. *Nature Communications* 4 (2013).

[48] G. Gour, M. P. Müller, V. Narasimhachar, R. W. Spekkens, and N. Y. Halpern. *The Resource Theory of Informational Nonequilibrium in Thermodynamics*. *Physics Reports* 583 (2015).

[49] F. Albarelli, M. G. Genoni, M. G. A. Paris, and A. Ferraro. *Resource Theory of Quantum Non-Gaussianity and Wigner Negativity*. *Physical Review A: Atomic, Molecular, and Optical Physics* 98 (2018).

[50] R. Takagi and Q. Zhuang. *Convex Resource Theory of Non-Gaussianity*. *Physical Review A: Atomic, Molecular, and Optical Physics* 97 (2018).

[51] M. Howard and E. Campbell. *Application of a Resource Theory for Magic States to Fault-Tolerant Quantum Computing*. *Physical Review Letters* 118 (2017).

[52] G. Gour. *Resources of the Quantum World*. 2024. arXiv: 2402.05474 [quant-ph].

[53] E. T. Campbell, B. M. Terhal, and C. Vuillot. *Roads towards Fault-Tolerant Universal Quantum Computation*. *Nature* 549 (2017).

[54] D. Gottesman. *Surviving as a Quantum Computer in a Classical World*. https://www.cs.umd.edu/class/sp 2024-ch1-15.pdf, 2024.

[55] S. B. Bravyi and A. Y. Kitaev. *Quantum Codes on a Lattice with Boundary*. 1998. arXiv: quant-ph/9811052.

[56] R. Acharya et al. *Quantum Error Correction below the Surface Code Threshold*. *Nature* 638 (2025). arXiv: 2408.13687 [quant-ph].

[57] B. Eastin and E. Knill. *Restrictions on Transversal Encoded Quantum Gate Sets*. *Physical Review Letters* 102 (2009).

[58] G. Nebe, E. M. Rains, and N. J. A. Sloane. *The Invariants of the Clifford Groups*. 2000. arXiv: math/0001038.

[59] G. Nebe, E. M. Rains, and N. J. A. Sloane. *Self-Dual Codes and Invariant Theory*. Vol. 17. Algorithms and Computation in Mathematics. Berlin/Heidelberg: Springer-Verlag, 2006.

[60] D. Gross. *Hudson's Theorem for Finite-Dimensional Quantum Systems*. *Journal of Mathematical Physics* 47 (2006).

[61] X. Wang, M. M. Wilde, and Y. Su. *Quantifying the Magic of Quantum Channels*. *New Journal of Physics* 21 (2019).

[62] E. Wigner. *On the Quantum Correction for Thermodynamic Equilibrium*. *Physical Review* 40 (1932).

[63] V. Veitch, C. Ferrie, D. Gross, and J. Emerson. *Negative Quasi-Probability as a Resource for Quantum Computation*. *New Journal of Physics* 14 (2012).

[64] R. W. Spekkens. *Contextuality for Preparations, Transformations, and Unsharp Measurements*. *Physical Review A: Atomic, Molecular, and Optical Physics* 71 (2005).

[65] D. Schmid, H. Du, J. H. Selby, and M. F. Pusey. *Uniqueness of Noncontextual Models for Stabilizer Subtheories*. *Physical Review Letters* 129 (2022).

[66] A. F. Veinott. *Least D-Majorized Network Flows with Inventory and Statistical Applications*. 17 (1971).

[67] D. Blackwell. *Equivalent Comparisons of Experiments*. *Annals of Mathematical Statistics* 24 (1953).

[68] E. Ruch and A. Mead. *The Principle of Increasing Mixing Character and Some of Its Consequences*. *Theoretica Chimica Acta* 41 (1976).

[69] M. Lostaglio. *An Introductory Review of the Resource Theory Approach to Thermodynamics*. 82 (2019).

[70] R. Smith, Z. Papić, and A. Hallam. *Nonstabilizerness in Kinetically Constrained Rydberg Atom Arrays*. *Physical Review Letters* 111 (2025).

[71] D. M. Appleby. *Symmetric Informationally Complete–Positive Operator Valued Measures and the Extended Clifford Group*. *Journal of Mathematical Physics* 46 (2005).

[72] L. Catani and D. E. Browne. *Spekkens' Toy Model in All Dimensions and Its Relationship with Stabiliser Quantum Mechanics*. *New Journal of Physics* 19 (2017).

[73] N. Delfosse, P. Allard Guerin, J. Bian, and R. Raussendorf. *Wigner Function Negativity and Contextuality in Quantum Computation on Rebits*. *Physical Review X* 5 (2015).

[74] L. Catani and D. E. Browne. *State-Injection Schemes of Quantum Computation in Spekkens' Toy Theory*. *Physical Review A: Atomic, Molecular, and Optical Physics* 98 (2018).

[75] A. W. Marshall, I. Olkin, and B. C. Arnold. *Inequalities: Theory of Majorization and Its Applications*. 2nd ed. Springer Series in Statistics. New York: Springer Science+Business Media, LLC, 2011.

[76] T. van Erven and P. Harremos. *Rényi Divergence and Kullback-Leibler Divergence*. *IEEE Transactions on Information Theory* 60 (2014).

[77] X. Wang, M. M. Wilde, and Y. Su. *Efficiently Computable Bounds for Magic State Distillation*. *Physical Review Letters* 124 (2020).

[78] K. Fang and Z.-W. Liu. *No-Go Theorems for Quantum Resource Purification*. *Physical Review Letters* 125 (2020).

[79] K. Fang and Z.-W. Liu. *No-Go Theorems for Quantum Resource Purification: New Approach and Channel Theory*. *PRX Quantum* 3 (2022).

[80] R. Babbush, C. Gidney, D. W. Berry, N. Wiebe, J. McClean, A. Paler, A. Fowler, and H. Neven. *Encoding Electronic Spectra in Quantum Circuits with Linear T Complexity*. *Physical Review X* 8 (2018).

[81] N. Rengaswamy, R. Calderbank, M. Newman, and H. D. Pfister. *On Optimality of CSS Codes for Transversal T*. *IEEE Journal on Selected Areas in Information Theory* 1 (2020).

[82] E. T. Campbell and D. E. Browne. *On the Structure of Protocols for Magic State Distillation*. In: *Theory of Quantum Computation, Communication, and Cryptography*. Ed. by D. Hutchison, T. Kanade, J. Kittler, J. M. Kleinberg, F. Mattern, J. C. Mitchell, M. Naor, O. Nierstrasz, C. Pandu Rangan, B. Steffen, M. Sudan, D. Terzopoulos, D. Tygar, M. Y. Vardi, G. Weikum, A. Childs, and M. Mosca. Vol. 5906. Berlin, Heidelberg: Springer Berlin Heidelberg, 2009.

[83] E. Knill, R. Laflamme, and W. Zurek. *Threshold Accuracy for Quantum Computation*. 1996. arXiv: `quant-ph/9610011`.

[84] A. M. Steane. *Quantum Reed-Muller Codes*. *IEEE Transactions on Information Theory* 45 (1999).

[85] G. Vidal and R. Tarrach. *Robustness of Entanglement*. *Physical Review A: Atomic, Molecular, and Optical Physics* 59 (1999).

[86] J. R. Seddon, B. Regula, H. Pashayan, Y. Ouyang, and E. T. Campbell. *Quantifying Quantum Speedups: Improved Classical Simulation from Tighter Magic Monotones*. *PRX Quantum* 2 (2021).

[87] B. Regula. *Probabilistic Transformations of Quantum Resources*. *Physical Review Letters* 128 (2022).

[88] N. Datta. *Relative Entropies and Entanglement Monotones*. In: *Mathematical Results in Quantum Physics*.

[89] R. L. Frank and E. H. Lieb. *Monotonicity of a Relative Rényi Entropy*. *Journal of Mathematical Physics* 54 (2013).

[90] M. Mosonyi. *Convexity Properties of the Quantum Rényi Divergences, with Applications to the Quantum Stein's Lemma*. In: *9th Conference on the Theory of Quantum Computation, Communication and Cryptography, May 21-23, 2014, Singapore*. Ed. by S. T. Flammia and A. W. Harrow. Vol. 27. LIPIcs. Schloss Dagstuhl - Leibniz-Zentrum für Informatik, 2014.

[91] M. M. Wilde, A. Winter, and D. Yang. *Strong Converse for the Classical Capacity of Entanglement-Breaking and Hadamard Channels via a Sandwiched Rényi Relative Entropy*. *Communications in Mathematical Physics* 331 (2014).

[92] R. Raussendorf, J. Bermejo-Vega, E. Tyhurst, C. Okay, and M. Zurel. *Phase-Space-Simulation Method for Quantum Computation with Magic States on Qubits*. *Physical Review A: Atomic, Molecular, and Optical Physics* 101 (2020).

[93] A. Giovagnoli and H. P. Wynn. *Cyclic Majorization and Smoothing Operators*. *Linear Algebra and Its Applications* 239 (1996).

[94] A. Giovagnoli and H. P. Wynn. *G-Majorization with Applications to Matrix Orderings*. *Linear Algebra and Its Applications* 67 (1985).

[95] A. G. M. Steerneman. *G-Majorization, Group-Induced Cone Orderings, and Reflection Groups*. *Linear Algebra and Its Applications* 127 (1990).

[96] I. Marvian and R. W. Spekkens. *How to Quantify Coherence: Distinguishing Speakable and Unspeakable Notions*. *Physical Review A: Atomic, Molecular, and Optical Physics* 94 (2016).

[97] Y. Aharonov and L. Susskind. *Observability of the sign change of spinors under 2π rotations*. *Physical Review* 158 (1967).

[98] G. Chiribella, G. M. D'Ariano, P. Perinotti, and M. F. Sacchi. *Efficient Use of Quantum Resources for the Transmission of a Reference Frame*. *Physical Review Letters* 93 (2004).

[99] S. J. Jones, H. M. Wiseman, S. D. Bartlett, J. A. Vaccaro, and D. T. Pope. *Entanglement and Symmetry: A Case Study in Superselection Rules, Reference Frames, and Beyond*. *Physical Review A: Atomic, Molecular, and Optical Physics* 74 (2006).

[100] G. Gour and R. W. Spekkens. *The Resource Theory of Quantum Reference Frames: Manipulations and Monotones*. *New Journal of Physics* 10 (2008).

[101] M. J. W. Hall and H. M. Wiseman. *Does Nonlinear Metrology Offer Improved Resolution? Answers from Quantum Information Theory*. *Physical Review X* 2 (2012).

[102] V. Giovannetti, S. Lloyd, and L. Maccone. *Quantum Metrology*. *Physical Review Letters* 96 (2006).

[103] M. P. Woods and Á. M. Alhambra. *Continuous Groups of Transversal Gates for Quantum Error Correcting Codes from Finite Clock Reference Frames*. *Quantum* 4 (2020).

[104] R. Alexander. *A New Approximate Eastin-Knill Theorem*. 2025. arXiv: 2505.00427 [quant-ph].

[105] P. Faist, S. Nezami, V. V. Albert, G. Salton, F. Pastawski, P. Hayden, and J. Preskill. *Continuous Symmetries and Approximate Quantum Error Correction*. *Physical Review X* 10 (2020).

[106] Y. Yang, Y. Mo, J. M. Renes, G. Chiribella, and M. P. Woods. *Optimal Universal Quantum Error Correction via Bounded Reference Frames*. *Physical Review Research* 4 (2022).

[107] A. Almheiri, X. Dong, and D. Harlow. *Bulk Locality and Quantum Error Correction in AdS/CFT*. *Journal of High Energy Physics* 2015 (2015).

[108] F. Pastawski, B. Yoshida, D. Harlow, and J. Preskill. *Holographic Quantum Error-Correcting Codes: Toy Models for the Bulk/Boundary Correspondence*. *Journal of High Energy Physics* 2015 (2015).

[109] M. Gschwendtner, A. Bluhm, and A. Winter. *Programmability of Covariant Quantum Channels*. *Quantum* 5 (2021).

[110] C. Cirstoiu, K. Korzekwa, and D. Jennings. *Robustness of Noether's Principle: Maximal Disconnects between Conservation Laws and Symmetries in Quantum Theory*. *Physical Review X* 10 (2020).

[111] G. Chiribella, E. Aurell, and K. Życzkowski. *Symmetries of Quantum Evolutions*. *Physical Review Research* 3 (2021).

[112] M. Lostaglio, D. Jennings, and T. Rudolph. *Description of Quantum Coherence in Thermodynamic Processes Requires Constraints beyond Free Energy*. *Nature Communications* 6 (2015).

[113] M. Lostaglio, K. Korzekwa, D. Jennings, and T. Rudolph. *Quantum Coherence, Time-Translation Symmetry, and Thermodynamics*. *Physical Review X* 5 (2015).

[114] E. P. Wigner. *Die Messung Quantenmechanischer Operatoren*. *Zeitschrift fur Physik* 133 (1952).

[115] H. Araki and M. M. Yanase. *Measurement of Quantum Mechanical Operators*. *Physical Review* 120 (1960).

[116] M. M. Yanase. *Optimal Measuring Apparatus*. *Physical Review* 123 (1961).

[117] I. Marvian and R. W. Spekkens. *An Information-Theoretic Account of the Wigner-Araki-Yanase Theorem*. 2012. arXiv: 1212.3378 [quant-ph].

[118] M. Ahmadi, D. Jennings, and T. Rudolph. *The Wigner-Araki-Yanase Theorem and the Quantum Resource Theory of Asymmetry*. *New Journal of Physics* 15 (2013).

[119] B. Yadin and V. Vedral. *General Framework for Quantum Macroscopicity in Terms of Coherence*. *Physical Review A: Atomic, Molecular, and Optical Physics* 93 (2016).

[120] I. Marvian, R. W. Spekkens, and P. Zanardi. *Quantum Speed Limits, Coherence, and Asymmetry*. *Physical Review A: Atomic, Molecular, and Optical Physics* 93 (2016).

[121] F. Giacomini, E. Castro-Ruiz, and Č. Brukner. *Quantum Mechanics and the Covariance of Physical Laws in Quantum Reference Frames*. *Nature Communications* 10 (2019).

[122] A. Vanrietvelde, P. A. Hoehn, F. Giacomini, and E. Castro-Ruiz. *A Change of Perspective: Switching Quantum Reference Frames via a Perspective-Neutral Framework*. *Quantum* 4 (2020).

[123] R. Bousso. *Robust Singularity Theorem*. *Physical Review Letters* 135 (2025).

[124] M. Keyl and R. F. Werner. *Optimal Cloning of Pure States, Testing Single Clones*. *Journal of Mathematical Physics* 40 (1999).

[125] I. Marvian Mashhad. *Symmetry, Asymmetry and Quantum Information*. PhD thesis. University of Waterloo, 2012.

[126] R. Renner. *Security of Quantum Key Distribution*. PhD thesis. ETH Zurich, 2005.

[127] S. D. Bartlett, T. Rudolph, R. W. Spekkens, and P. S. Turner. *Quantum Communication Using a Bounded-Size Quantum Reference Frame*. *New Journal of Physics* 11 (2009).

[128] L. Loveridge, P. Busch, and T. Miyadera. *Relativity of Quantum States and Observables*. *EPL (Europhysics Letters)* 117 (2017).

[129] L. Loveridge, T. Miyadera, and P. Busch. *Symmetry, Reference Frames, and Relational Quantities in Quantum Mechanics*. *Foundations of Physics* 48 (2018).

[130] L. Loveridge. *A Relational Perspective on the Wigner-Araki-Yanase Theorem*. *Journal of Physics: Conference Series* 1638 (2020).

[131] R. Konig, R. Renner, and C. Schaffner. *The Operational Meaning of Min-and Max-Entropy*. *IEEE Transactions on Information Theory* 55 (2009).

[132] D. N. Page and W. K. Wootters. *Evolution without Evolution: Dynamics Described by Stationary Observables*. *Physical Review D: Particles and Fields* 27 (1983).

[133] K. Korzekwa, M. Lostaglio, J. Oppenheim, and D. Jennings. *The Extraction of Work from Quantum Coherence*. *New Journal of Physics* 18 (2016).

[134] I. Bengtsson and K. Życzkowski. *Geometry of Quantum States: An Introduction to Quantum Entanglement*. 2nd ed. Cambridge, UK: Cambridge University Press, 2017.

[135] K. Yamaguchi and H. Tajima. *Beyond i.i.d. in the Resource Theory of Asymmetry: An Information-Spectrum Approach for Quantum Fisher Information*. *Physical Review Letters* 131 (2023).

[136] M. Hayashi. *Quantum Information*. Springer Berlin Heidelberg, 2006.

[137] I. Marvian and R. W. Spekkens. *Modes of Asymmetry: The Application of Harmonic Analysis to Symmetric Quantum Dynamics and Quantum Reference Frames*. *Physical Review A: Atomic, Molecular, and Optical Physics* 90 (2014).

[138] P. Hausladen and W. K. Wootters. *A 'Pretty Good' Measurement for Distinguishing Quantum States*. *Journal of Modern Optics* 41 (1994).

[139] B. C. Hall. *Lie Groups, Lie Algebras, and Representations: An Elementary Introduction*. 2 ed. 2015. Graduate Texts in Mathematics 222. Cham: Springer, 2015.

[140] C. Cirstoiu and D. Jennings. *Global and Local Gauge Symmetries beyond Lagrangian Formulations*. 2018. arXiv: 1707.09826 [quant-ph].

[141] M. Tomamichel, R. Colbeck, and R. Renner. *Duality between Smooth Min- and Max-Entropies*. *IEEE Transactions on Information Theory* 56 (2010).

[142] T. Rudolph and L. Grover. *A 2 Rebit Gate Universal for Quantum Computing*. 2002. arXiv: quant-ph/0210187.

[143] P. Allard Guérin. *Wigner Function Negativity and Contextuality in Quantum Computation with Rebits*. MA thesis. The University of British Columbia, 2015.

[144] C.-Y. Lai, Y.-C. Zheng, and T. A. Brun. *Fault-Tolerant Preparation of Stabilizer States for Quantum Calderbank-Shor-Steane Codes by Classical Error-Correcting Codes*. *Physical Review A: Atomic, Molecular, and Optical Physics* 95 (2017).

[145] D. Gross. *Computational Power of Quantum Many-Body States and Some Results on Discrete Phase Spaces*. PhD thesis. Imperial College London, 2005.

[146] R. Bhatia. *Matrix Analysis*. Vol. 169. Springer Science & Business Media, 2013.

[147] M. P. Woods and M. Horodecki. *Autonomous Quantum Devices: When Are They Realizable without Additional Thermodynamic Costs?* *Physical Review X* 13 (2023).

[148] M. Tomamichel. *A Framework for Non-Asymptotic Quantum Information Theory*. PhD thesis. ETH Zurich, 2012.

[149] R. A. Horn and C. R. Johnson. *Matrix Analysis*. 2 ed, corrected reprint. New York, NY: Cambridge University Press, 2017.

[150] T. Tkocz. *An Introduction to Convex and Discrete Geometry*. Lecture Notes. 2019.

[151] M. Ledoux and M. Talagrand. *Probability in Banach Spaces : Isoperimetry and Processes*. Ergebnisse Der Mathematik Und Ihrer Grenzgebiete ; 3. Folge, Band 23. Springer-Verlag, 1991.

[152] B. Poonen. *Real Representations*. Notes.