Noise in Quantum Information Processing

Christopher Thomas Chubb

A thesis submitted in fulfilment of the requirements for the degree of Doctor of Philosophy

Faculty of Science
School of Physics
University of Sydney

2019

Contents

Abstract

Quantum phenomena such as superposition and entanglement imbue quantum systems with information processing power in excess of their classical counterparts. These properties of quantum states are, however, highly fragile. As we enter the era of noisy intermediate-scale quantum (NISQ) devices, this vulnerability to noise is a major hurdle to the experimental realisation of quantum technologies. In this thesis we explore the role of noise in quantum information processing from two different perspectives.

In Part I we consider noise from the perspective of quantum error correction. Error correcting codes are often analysed with respect to simplified toy models of noise, such as iid depolarising noise. We consider generalising these techniques for analysing codes under more realistic noise models, which include features such as biased or correlated errors. We also consider designing customised codes which not only take into account—but exploit—features of the underlying physical noise processes. Whilst generalist codes may give sufficient performance in the limit of large system sizes, considering such tailored codes will be of particular importance for NISQ applications in which finite-size effects can be significant.

In Part II we apply tools from information theory to study the finite-resource effects which arise in the trade-offs between resource costs and error rates for certain quantum information processing tasks. We start by considering classical communication over quantum channels, providing a refined analysis of the trade-off between communication rate and error in the regime of a finite number of channel uses. We then extend these techniques to the problem of resource interconversion in theories such as quantum entanglement and quantum thermodynamics, studying finite-size effects which arise in resource-error trade-offs. By studying this effect in detail, we also show how detrimental finite-size effects in devices such as thermal engines may be greatly suppressed by carefully engineering the underlying resource interconversion processes on which they are based.

Statement of Student Contribution

This thesis is composed of seven research papers. Four of these papers are published and/or accepted in peer-reviewed journals, while the remaining three are pre-prints currently under consideration for publication. Part I consists of three papers produced with my primary supervisor Prof. Steven T. Flammia, and Part II of four papers produced with my co-supervisor Dr. Marco Tomamichel. For consistency, minimal formatting changes have been made to put each paper into single column format, and each paper has been changed to Australian English.

Two additional papers produced during my PhD which were outside of the scope of this thesis have been omitted:

Hand-waving and Interpretive Dance: An Introductory Course on Tensor Networks
Jacob C. Bridgeman and Christopher T. Chubb
Journal of Physics A: Mathematical and Theoretical **50**, 223001 (2017)

Energy cost of entanglement extraction in complex quantum systems
Cédric Bény, Christopher T. Chubb, Terry Farrelly, and Tobias J. Osborne
Nature Communications **9** (1), 3792 (2018)

The first of these was a review paper produced with a fellow PhD student, which has already appeared in Jacob Bridgeman's PhD thesis. The second of these was a research paper produced with international collaborators.

Chapter 2: Statistical mechanical models for quantum codes with correlated noise

This chapter contains a pre-print paper currently under consideration for publication, co-authored with Steve Flammia.

The research represented in this paper was undertaken by me, with supervision from Steve. The initial ideas originated with Steve, and we both developed these ideas while attending the thematic semester on Analysis in Quantum Information Theory at the Institut Henri Poincaré in Paris. I wrote up the technical sections, produced and executed the numerics, and drafted the paper. Both authors jointly edited the paper.

Chapter 3: Tailoring surface codes for highly-biased noise

This chapter contains a pre-print paper currently under consideration for publication, co-authored with David Tuckett, Sergey Bravyi, Stephen Bartlett and Steve Flammia. This paper is also to appear in David's PhD thesis.

The research represented in this paper was primarily undertaken by David, with supervision from Stephen and Steve. The initial idea of the project arose in discussions between David, Sergey, Stephen and Steve. I contributed some of the initial analytic ideas seen in Sections III and IV, which were then further developed by David and Sergey. The numerics in Sections IV and V were performed by David. The initial draft was produced by David, which all authors contributed to the editing of.

Chapter 4: Approximate symmetries of Hamiltonians

This chapter contains a published paper, which appeared in the Journal of Mathematical Physics, co-authored with Steve Flammia.

The research represented in this paper was undertaken by me, with supervision from Steve Flammia. I proposed the initial idea in discussion with Steve about a previous related paper of his. Steve provided me with invaluable mathematical references which helped me to flesh this idea out. The original technical results were written up by me, with Steve helping me to streamline some of the proofs. I drafted the paper, which both authors then jointly edited.

Chapter 5: Moderate deviation analysis for classical communication over quantum channels

This chapter contains a published paper, which appeared in Communications in Mathematical Physics, co-authored with Marco Tomamichel and Vincent Tan.

The original idea was proposed by Marco. Vincent and Marco both provided me with invaluable references to the relevant classical and quantum literature respectively. I produced both the technical results and initial draft under the supervision of Marco. All authors contributed to subsequent editing.

Chapter 6: Beyond the thermodynamic limit: finite-size corrections to state interconversion rates

This chapter contains a published paper, which appeared in Quantum, co-authored with Kamil Korzekwa and Marco Tomamichel.

The initial ideas arose from discussions among all three authors. I produced the technical results, and performed the numerics, both under the supervision of Marco and Kamil. Marco and Kamil drafted the non-technical parts of the paper, and all authors contributed to the applications discussed. All authors jointly edited the paper.

Chapter 7: Moderate deviation analysis of majorisation-based resource interconversion

This chapter contains a pre-print paper currently under consideration for publication, co-authored with Kamil Korzekwa and Marco Tomamichel.

All authors contributed to the initial ideas. I produced the technical results, under the supervision of Marco and Kamil. Marco and Kamil drafted the non-technical parts, and all authors jointly edited the paper.

Chapter 8: Avoiding irreversibility: engineering resonant conversions of quantum resources

This chapter contains a pre-print paper accepted for publication in Physical Review Letters, co-authored with Kamil Korzekwa and Marco Tomamichel.

All authors contributed to the initial ideas. The applications considered arose from discussion between all authors, and the numerics were performed by Kamil and myself. Marco and Kamil drafted the paper, which all authors then jointly edited.

Statement of Originality

This is to certify that to the best of my knowledge, the content of this thesis is my own work. This thesis has not been submitted for any degree or other purposes.

I certify that the intellectual content of this thesis is the product of my own work and that all the assistance received in preparing this thesis and sources have been acknowledged.

Student: Christopher T. Chubb **Supervisor:** Steven T. Flammia

Acknowledgements

I would like to thank a great many people for their support during my PhD. Firstly I'd like to thank both my supervisor Steve Flammia, as well as my co-supervisor Marco Tomamichel, for their time, effort and encouragement. I would also like to thank all my other co-authors, especially Jacob Bridgeman and Kamil Korzekwa.

My time within the University of Sydney Quantum Information group has highly formative. Specifically I would like to thank Raf Alexander, Daniel Barter, Benji Brown, Adrian Chapman, Zach Christina, Andrew Darmawan, Sam Elman, Tim Evans, Arne Grimsmo, Robin Harper, Angela Karanjai, Hakop Pashayan, Sam Roberts, Thomas Smith, David Tuckett and Wei-Wei Zhang for their support.

Lastly special thanks go to my parents, Ken and Jenny, for tolerating and accommodating me. Hopefully now I can finally get a job...

Chapter 1

Introduction

Quantum mechanical systems provide a powerful platform for information processing tasks, including computing, communication and cryptography [1.1]. These protocols rely on features of quantum systems such as superposition and entanglement which do not appear in classical systems. Unfortunately, these quantum properties are typically fragile, and susceptible to environmental noise. This inherent sensitivity therefore forms a significant roadblock to large-scale realisation of such quantum technologies.

A direct approach for dealing with this issue is to improve and refine both the hardware implementation and software control systems to suppress environmental noise. Whilst this is an important first step, the error rates of state-of-the-art experimental devices are far in excess of those necessary for many applications. Moreover, this approach has the disadvantage that the error tolerances would need to tighten as we consider larger and larger information processing tasks.

Instead of simply lowering raw error rates, another way of making quantum protocols robust to noise is through the use of *quantum error correction* [1.1–1.5]. Here the idea is to modify our protocols such that they may tolerate some non-trivial level of noise, which is done by detecting and then correcting errors as they occur. By doing so, we can devise protocols which can perform the same information processing tasks as their noiseless counterparts, but can be performed on realistic noisy hardware, with low overhead.

In this thesis we study the role of noise in quantum systems from two perspectives: quantum error correction, and resource-error trade-offs. In Part I we study the performance of explicit quantum error correcting codes. Specifically we look at how the analysis and design of such codes may be tailored to the details of the underlying noise processes. In Part II we study the fundamental limits on the performance of certain noisy quantum processes. Specifically we look at these resource-error trade-offs for problems in quantum communication, thermodynamics, and entanglement theory.

1.1 Quantum error correction

The goal of error correction is to find ways of encoding logical information into noisy physical degrees of freedom, such that it may be successfully decoded with high probability. This is done by building redundancy into our encoding. Using this, we can perform consistency checks to identify any possible errors which have occurred, which we can then hope to correct.

The efficacy of an error correcting code is highly dependant on the underlying noise model. The idea here is to tailor both the redundancies and corresponding consistency checks so that they capture typical errors in our model. Extending the analysis of codes away from simple toy models of noise, into the regime of more realistic error models, is an important step towards improving the design and analysis of error correcting procedures for near-term applications.

1.1.1 Classical repetition code

The simplest classical code is the *repetition code*, which simply involves storing multiple copies of the encoded information. If we consider encoding a single logical bit into three physical bits, this means that

$$\overline{0} = 000, \tag{1.1}$$
$$\overline{1} = 111, \tag{1.2}$$

where $\overline{0}$ and $\overline{1}$ denote the two code states. By storing at least three copies of our encoded information, we can correct any single-bit error by performing a majority vote between these three copies. More generally if we consider a $(2n+1)$-bit repetition code, then the majority vote decoder will protect our information against any error which flips at most n bits.

Consider a stochastic noise process in which each bit suffers a bit-flip error with probability p, and is unchanged with probability $1-p$. If p is small, this noise process will tend to flip a minority of bits, precisely the errors to which the majority vote decoder is robust. As such, we can see that the repetition code is well suited to decoding this noise model. Indeed it can be shown that, for any $p < 1/2$, the probability of an unsuccessful decoding vanishes exponentially with code size n, as shown in Figure 1.1.

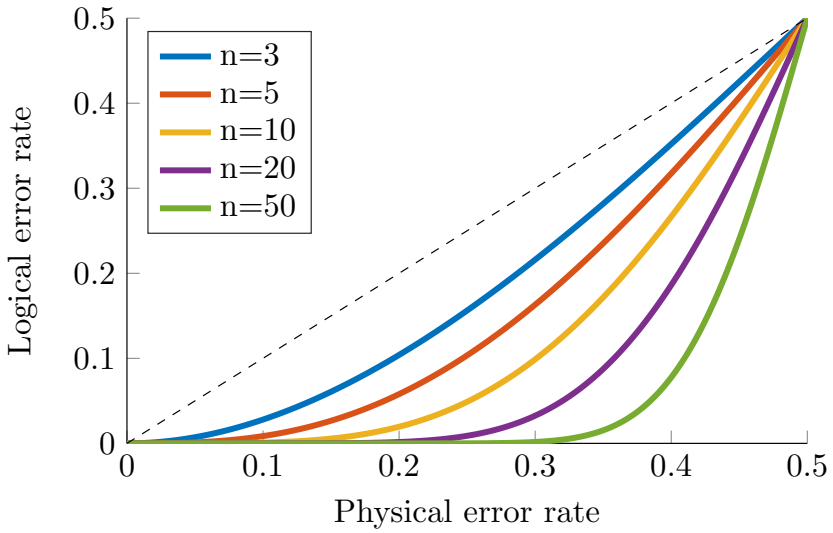


Figure 1.1: The probability of the encoded bit flipping (logical error rate) as a function of the probability of a physical bit flipping (physical error rate) for the n-bit repetition code. The dashed line denotes parity between physical and logical error rates.

One figure of merit for quantifying how well suited a code is to a given error model is its *error threshold*. The error threshold is a critical error rate below which the code may store logical information arbitrarily well by simply growing the code. In this case we can see that the repetition code has a threshold of $p_{\text{th}} = 1/2$ against iid bit-flip noise.

1.1.2 Quantum repetition codes

An obvious starting point for a quantum code would be to simply encode the basis states into a classical error correcting code. For example we could consider a quantum extension of the repetition code,

$$|\overline{0}\rangle = |000\rangle, \tag{1.3}$$
$$|\overline{1}\rangle = |111\rangle. \tag{1.4}$$

Whilst this code is still robust to bit-flip errors $|b\rangle \mapsto |b \oplus 1\rangle$, quantum systems also contain phase information, and this code is *not* robust phase-flip noise $|b\rangle \mapsto (-1)^b|b\rangle$. Specifically we can see that a phase-flip error on any physical qubit induces a phase-flip error on the encoded state, $|\bar{b}\rangle \mapsto (-1)^b|\bar{b}\rangle$. For this reason, this code is often referred to as the *bit-flip code*.

One can also construct a second quantum analogue of the repetition code known as the *phase-flip code*, where information is encoded into the phases of our qubit states as

$$|\overline{0}\rangle = \frac{1}{\sqrt{8}}\Big(|0\rangle + |1\rangle\Big)\Big(|0\rangle + |1\rangle\Big)\Big(|0\rangle + |1\rangle\Big), \tag{1.5}$$
$$|\overline{1}\rangle = \frac{1}{\sqrt{8}}\Big(|0\rangle - |1\rangle\Big)\Big(|0\rangle - |1\rangle\Big)\Big(|0\rangle - |1\rangle\Big). \tag{1.6}$$

Again, whilst this code is protected against single-qubit phase-flip errors, it is however unprotected against any bit-flip errors.

The problem with both codes is that neither is truly a repetition code in the same sense as the classical repetition code. One could imagine a hypothetical code in which

$$|\overline{\psi}\rangle = |\psi\rangle \otimes |\psi\rangle \otimes |\psi\rangle, \tag{1.7}$$

for *any* qubit state $|\psi\rangle$. We note that the above codes are only of this form for $|\psi\rangle$ *in a specific basis*. Unfortunately, such an encoding procedure for *any* $|\psi\rangle$ would violate the *no-cloning principle* [1.6]. For this reason we cannot simply construct quantum error correcting codes based on copying our logical information, as we can in the classical realm.

1.1.3 Shor code

A true quantum-error correcting code will need to not only be robust against bit-flip errors, but arbitrary quantum errors. This might at first seem like a very significant issue: whilst a bit can only suffer an error in a single way, there are a continuous family of qubit errors. How can we possibly hope to correct any of these infinite errors? As we shall see, one way of achieving this is to combine both the bit-flip and phase-flip codes, giving a code known as the *Shor code* [1.2].

The Shor code is an example of a *concatenated code* [1.7], in which we encode our information multiple times. Specifically we first encode our state into the bit-flip code, and then once again into the phase-flip code. This results in the nine-qubit encoding

$$|\bar{0}\rangle = \frac{1}{\sqrt{8}}\Big(|000\rangle + |111\rangle\Big)\Big(|000\rangle + |111\rangle\Big)\Big(|000\rangle + |111\rangle\Big), \tag{1.8}$$

$$|\bar{1}\rangle = \frac{1}{\sqrt{8}}\Big(|000\rangle - |111\rangle\Big)\Big(|000\rangle - |111\rangle\Big)\Big(|000\rangle - |111\rangle\Big). \tag{1.9}$$

By construction this code is separately robust to a bit-flip on any qubit *and* a phase-flip on any qubit, which are described by the Pauli X and Z respectively, where

$$X := |0\rangle\langle 1| + |1\rangle\langle 0|\,, \qquad Z := |0\rangle\langle 0| - |1\rangle\langle 1|\,. \tag{1.10}$$

Fortunately, if a quantum code can correct any two errors, it can also correct any linear superposition thereof [1.1, 1.5], which circumvents the issue of the infinitude of quautum errors mentioned earlier. As a result of this, the ability of the Shor code to correct any single-X *and* any single-Z error implies that it can in fact correct *any* single-qubit error. By considering an analogous construction based on larger repetition codes, we can construct a $(2n+1)^2$-qubit code which can correct any n-qubit error.

1.1.4 Threshold theorem

Above we have considered the simplest model of error correction, in which we attempt to simply store quantum information, with access to ideal encoding and decoding operations. As described earlier, these codes can be used to augment protocols such as quantum computation to make them fault-tolerant. In this model we wish not only to *store* quantum states [1.8, 1.9], but also to perform computations upon them, in a regime in which all operations we may wish to perform are noisy. The *threshold theorem* [1.10–1.13] states that, for any quantum system with operations for which the noise rate is sufficiently low, and errors only mildly correlated, noiseless quantum computations may be arbitrarily well approximated with low overhead.

Critically, the threshold for such a scheme is a constant independent of the size of the computation to be performed. This is a qualitative improvement on simply attempting to drive down raw error rates, which would naturally require tighter error tolerances as the computation grows, as noted earlier.

1.2 Resource-error trade-offs

One way of analysing the role of noise in quantum information processes is, as above, to use error correcting codes to construct specific fault-tolerant protocols, and analyse their performance. Another approach is to, leveraging techniques from quantum Shannon theory [1.14–1.16], study the fundamental trade-off between error rates and the performance for certain quantum tasks. A key advantage of this technique is that it allows for the error tolerance of these tasks to be analysed independent of the specific error correction procedure used.

1.2.1 Channel coding

One of the simplest information theoretic tasks is communication via a noisy channel [1.17,1.18]. Here the goal is to provide encoding and decoding procedures such that messages may be transmitted over the channel, and successfully decoded with high probability. Generally there is a trade-off between the amount of data which may be transmitted per channel use, known as the rate, and the error rate associated with decoding.

Whilst quantum mechanical encodings can be used to transmit quantum information, in the form of qubits, they can also be used to transmit more classical information than classical encodings alone [1.19–1.21]. An example where this can be seen is the bit-flip channel,

$$\mathcal{E}(\rho) = (1-p)\rho + pX\rho X, \tag{1.11}$$

where $X = |0\rangle\langle 1| + |1\rangle\langle 0|$ is the bit-flip operator. If we just encode our messages into classical states, this channel acts as the binary symmetric channel, flipping each of our bits with probability p. As this channel has a non-zero probability of sending any bit-string to any other, we can never hope to noiselessly transmit messages in this way. If however we allow for quantum encodings, we can use a phase encoding,

$$0 \to \frac{|0\rangle + |1\rangle}{\sqrt{2}}, \tag{1.12}$$

$$1 \to \frac{|0\rangle - |1\rangle}{\sqrt{2}}, \tag{1.13}$$

such that our encoded states are now invariant under $\mathcal{E}$. As such, quantum mechanics allows us to *perfectly* transmit one bit per channel use. In this way we can see that the performance of a channel under classical and quantum encodings can vastly differ.

1.2.2 Resource theories

The laws of nature place restrictions on the allowed set of physical processes: nothing can move faster than the speed of light, CPT symmetry must be obeyed, energy may be neither created nor destroyed. Only considering allowed operations which obey these restrictions will imbue certain objects with value, much in the way that scarcity imbues goods with value in economics. The mathematical frameworks for studying such constrained physical systems are known as *resource theories* [1.22, 1.23].

A resource theory is defined by specifying a set of free states and free operations. Any state which can be consumed to allow an operation outside of the free operations is referred to as a *resource*, e.g. a battery being consumed to charge a phone when energy cannot be created. Two important examples of resource theories are those of entanglement [1.24] and quantum thermodynamics [1.25–1.27].

The resource theory of entanglement arises from considering two or more parties who are incapable of performing non-local quantum operations. The free operations are any local operations together with classical communication (LOCC), while the free states are separable states. Supplying the parties with pre-shared entanglement however allows them to perform operations outside of LOCC—it is in this sense that entanglement is a resource in this setting. Indeed in the bipartite case, a maximally entangled state allows for an arbitrary operation to be performed, which is the basis of protocols such as quantum teleportation [1.28] and super-dense coding [1.29].

The resource theory of quantum thermodynamics arises from considering only operations which obey the laws of thermodynamics, as well as exchange with a thermal bath. There exist multiple ways of formalising this resource theory, which we discuss in detail in Chapters 6 and 7. In this theory, identifying resourceful states is not so straightforward, with both athermality and coherence forming quantum thermodynamic resources.

One way of assessing the resource values of given states is through the *state interconversion problem*, which asks which target states may be reached from a initial state using only free operations.

1.2.3 Asymptotics and finite-size effects

The state interconversion problem is often studied in one of two rather extreme regimes: single-shot or asymptotic.

In the single-shot regime we consider a single copy of our input state, and ask which output states we can reach to within some error tolerance ϵ. The resulting conditions are often analytically complicated, as well as difficult to compute with, making these results impractical for all but the smallest systems.

In contrast to single-shot, the asymptotic regime is particularly mathematically convenient. Here we imagine having an arbitrarily large number of copies n of our input state, and transforming these into Rn copies of our target state ϕ to within some error tolerance ϵ, using only free operations. The question here is the trade-off between the rate R, the number of target states *per* input state, and the error tolerance ϵ. The convenience of this asymptotic regime arises the ability to leverage concentration phenomena such as the asymptotic equipartition property.

Whilst both regimes help to illuminate the features of their underlying resource theories, neither analysis is particularly suitable when considering the intermediate regime of a large, but not infinite, number of particles. The single-shot analysis often scales poorly, and the asymptotic analysis neglects finite-size effects. For near-term applications it is therefore imperative to marry both of these approaches, to give analysis which applies in this intermediate regime.

1.3 Thesis overview

In this thesis we will study the roles of noise in quantum information processing, both by studying the design of quantum error correcting codes, as well as fundamental resource-error trade-offs in certain quantum information processing tasks.

In Chapter 2 we extend a mapping between quantum error correcting codes and statistical mechanical models, originally considered in Ref. [1.30], to the regime of correlated noise. Importantly, under this mapping the error correcting threshold of our code is mapped onto the phase transition of the resulting statistical mechanical model. As such, by utilising techniques for locating phase transitions, such as Monte Carlo simulation, this mapping may be used to estimate thresholds of codes, in a decoder-independent fashion. Moreover, using tensor network methods, we can also show that this mapping can provide an efficient algorithm which approximates the optimal decoder, generalising the decoder of Ref. [1.31] to more general codes and correlated error models. By way of example, we apply these methods to numerically study the surface code under a model of correlated bit-flip noise, showing how correlations within the noise model can significantly alter the resulting threshold.

In Chapter 3 we study the performance of the surface code under noise which is biased towards dephasing. By a simple modification, the surface code is known to exhibit an ultra-high threshold to

such noise, and we identify the features of the surface code which cause this. We show that in the limit of pure dephasing the error threshold approaches 50%. We then study the sub-threshold behaviour, showing that careful choice of the system sizes used can dramatically lower the logical error rates. Lastly, we show numerically that all of these effects persist even when the bias is finite.

In Chapter 4 we consider error correcting codes for which only approximate logical operators are known. We show that, as long as they obey approximate commutation relations, these operators still imply the existence of an underlying error correcting code. Following the work of Ref. [1.32], we also discuss how this result can be interpreted from a condensed matter perspective, in terms of approximate anyonic statistics, serving as a certificate of topological order.

In Chapter 5 we consider trade-offs between the classical communication rate and error for communication over quantum channels. This problem is often analysed either in the regime where the error tolerance is fixed, or the transmission rate is fixed. In this work we consider an intermediate regime where both the rate approaches the capacity *and* the error probability vanishes. As well as providing an asymptotic analysis, we also consider finite-size corrections, which are important in the regime of large-but-finite channel uses.

In Chapters 6 and 7 we consider the trade-offs between error probabilities and the rates of state interconversion in the resource theories of quantum thermodynamics, entanglement, and coherence. In Chapter 6 we analyse the problem in the regime of constant error, which we then refine in Chapter 7 to also account for the possibility of sub-exponentially decreasing error.

In Chapter 8 we explore an interesting physical consequence of the results in Chapters 6 and 7, a phenomenon we refer to as *resource resonance*. We show that, by carefully engineering the resource interconversion process, this resonance allows one to greatly mitigate potentially detrimental finite-size effects. We verify this analytic prediction by performing exact numerical operations to observe this phenomenon even in relatively small systems. As well as this, we explore several applications, such as to small quantum heat engines.

Finally in Chapter 9 we provide an overview of the thesis, discuss possible extensions of this work, and give concluding remarks.

Each of the Chapters 2-8 are all self-contained papers, and all begin with reviews of the relevant literature and notation.

Bibliography

1.1 M. A. Nielsen and I. L. Chuang, Quantum Computation and Quantum Information 10th Anniversary Edition. Cambridge University Press, Cambridge, (2011).

1.2 P. W. Shor, "Scheme for reducing decoherence in quantum computer memory," *Physical Review A*, **52**, pp. R2493–R2496, (1995).

1.3 A. M. Steane, "Error Correcting Codes in Quantum Theory," *Physical Review Letters*, **77**, pp. 793–797, (1996).

1.4 D. Gottesman, "An Introduction to Quantum Error Correction and Fault-Tolerant Quantum Computation," arXiv:0904.2557, (2009).

1.5 S. J. Devitt, W. J. Munro, and K. Nemoto, "Quantum error correction for beginners," *Reports on Progress in Physics*, **76**, p. 076001, arXiv:0905.2794, (2013).

1.6 W. K. Wootters and W. H. Zurek, "A single quantum cannot be cloned," *Nature*, **299**, pp. 802–803, (1982).

1.7 E. Knill and R. Laflamme, "Concatenated Quantum Codes," arXiv:quant-ph/9608012, (1996).

1.8 A. Y. Kitaev, "Quantum Error Correction with Imperfect Gates," in *Quantum Communication, Computing, and Measurement*, pp. 181–188, Boston, MA: Springer US, , Boston, MA(1997).

1.9 E. Knill, R. Laflamme, and W. Zurek, "Threshold Accuracy for Quantum Computation," arXiv:quant-ph/9610011, (1996).

1.10 E. Knill, R. Laflamme, and W. H. Zurek, "Resilient Quantum Computation: Error Models and Thresholds," *Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences*, **454**, pp. 365–384, arXiv:quant-ph/9702058, (1997).

1.11 D. Aharonov and M. Ben-Or, "Fault-Tolerant Quantum Computation With Constant Error Rate," pp. 1–63, arXiv:quant-ph/9906129, (1999).

1.12 D. Aharonov, A. Kitaev, and J. Preskill, "Fault-Tolerant Quantum Computation with Long-Range Correlated Noise," *Physical Review Letters*, **96**, p. 050504, arXiv:quant-ph/0510231, (2006).

1.13 J. Preskill, "Sufficient condition on noise correlations for scalable quantum computing," *Quantum Information & Computation*, **13**, 3-4, pp. 181–194, arXiv:1207.6131, (2013).

1.14 M. Hayashi, Quantum Information Theory. Graduate Texts in Physics, Berlin, Heidelberg: Springer Berlin Heidelberg, , Berlin, Heidelberg(2017).

1.15 M. M. Wilde, "Preface to the Second Edition," in *Quantum Information Theory*, pp. xi–xii, Cambridge: Cambridge University Press, , Cambridge(2011), arXiv:1106.1445.

1.16 J. Preskill, "Quantum Shannon Theory," arXiv:1604.07450, (2016).

1.17 C. E. Shannon, "A Mathematical Theory of Communication," *Bell System Technical Journal*, **27**, pp. 379–423, (1948).

1.18 N. Gisin and R. Thew, "Quantum communication," *Nature Photonics*, **1**, pp. 165–171, arXiv:quant-ph/0703255, (2007).

1.19 A. Holevo, "The capacity of the quantum channel with general signal states," *IEEE Transactions on Information Theory*, **44**, pp. 269–273, arXiv:quant-ph/961102, (1998).

1.20 A. S. Holevo, "Bounds for the Quantity of Information Transmitted by a Quantum Communication Channel," *Problems of Information Transmission*, **9**, 3, pp. 177–183, (1973).

1.21 B. Schumacher, "Sending entanglement through noisy quantum channels," *Physical Review A*, **54**, pp. 2614–2628, arXiv:quant-ph/9604023v1, (1996).

1.22 B. Coecke, T. Fritz, and R. W. Spekkens, "A mathematical theory of resources," *Information and Computation*, **250**, pp. 59–86, arXiv:arXiv:1409.5531v3, (2016).

1.23 M. Horodecki and J. Oppenheim, "(Quantumness in the context of) Resource Theories," *International Journal of Modern Physics B*, **27**, p. 1345019, arXiv:1209.2162, (2012).

1.24 R. Horodecki, P. Horodecki, M. Horodecki, and K. Horodecki, "Quantum Entanglement," *Review of Modern Physics*, **81**, pp. 865–942, (2009).

1.25 D. Janzing, P. Wocjan, R. Zeier, R. Geiss, and T. Beth, "The thermodynamic cost of reliability and low temperatures: Tightening Landauer's principle and the Second Law," *International Journal of Theoretical Physics*, **39**, pp. 2717–2753, arXiv:quant-ph/0002048, (2000).

1.26 M. Horodecki and J. Oppenheim, "Fundamental limitations for quantum and nanoscale thermodynamics," *Nature Communications*, **4**, p. 2059, arXiv:1111.3834, (2013).

1.27 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**, p. 250404, arXiv:1111.3882, (2013).

1.28 C. H. Bennett, G. Brassard, C. Crépeau, R. Jozsa, A. Peres, and W. K. Wootters, "Teleporting an unknown quantum state via dual classical and Einstein-Podolsky-Rosen channels," *Physical Review Letters*, **70**, pp. 1895–1899, (1993).

1.29 C. H. Bennett and S. J. Wiesner, "Communication via one- and two-particle operators on Einstein-Podolsky-Rosen states," *Physical Review Letters*, **69**, 2881, (1992).

1.30 E. Dennis, A. Kitaev, A. Landahl, and J. Preskill, "Topological quantum memory," *Journal of Mathematical Physics*, **43**, pp. 4452–4505, arXiv:quant-ph/0110143, (2001).

1.31 S. Bravyi, M. Suchara, and A. Vargo, "Efficient algorithms for maximum likelihood decoding in the surface code," *Physical Review A*, **90**, p. 032326, arXiv:1405.4883, (2014).

1.32 J. C. Bridgeman, S. T. Flammia, and D. Poulin, "Detecting topological order with ribbon operators," *Physical Review B*, **94**, p. 205123, arXiv:1603.02275, (2016).

Part I

Quantum error correction

Chapter 2

Statistical mechanical models for quantum codes with correlated noise

Christopher T. Chubb[1] and Steven T. Flammia[1,2]
[1] *Centre for Engineered Quantum Systems, University of Sydney, Sydney, NSW, Australia*
[2] *Yale Quantum Institute, Yale University, New Haven, Connecticut 06520, USA*

ArXiv: 1809.10704

Abstract

We give a broad generalisation of the mapping, originally due to Dennis, Kitaev, Landahl and Preskill, from quantum error correcting codes to statistical mechanical models. We show how the mapping can be extended to arbitrary stabiliser or subsystem codes subject to correlated Pauli noise models, including models of fault tolerance. This mapping connects the error correction threshold of the quantum code to a phase transition in the statistical mechanical model. Thus, any existing method for finding phase transitions, such as Monte Carlo simulations, can be applied to approximate the threshold of any such code, without having to perform optimal decoding. By way of example, we numerically study the threshold of the surface code under mildly correlated bit-flip noise, showing that noise with bunching correlations causes the threshold to drop to $p_{\text{corr}} = 10.04(6)\%$, from its known iid value of $p_{\text{iid}} = 10.917(3)\%$. Complementing this, we show that the mapping also allows us to utilise any algorithm which can calculate/approximate partition functions of classical statistical mechanical models to perform optimal/approximately optimal decoding. Specifically, for 2D codes subject to locally correlated noise, we give a linear-time tensor network-based algorithm for approximate optimal decoding which extends the MPS decoder of Bravyi, Suchara and Vargo.

Quantum mechanical systems are inherently sensitive to noise. The inability to completely suppress environmental noise and perform noiseless quantum operations therefore provides a significant barrier to scalable quantum information processing. To mitigate this, quantum error correcting codes [2.1,2.2] were developed that encode quantum information into a larger system whose redundant degrees of freedom provide protection from physical noise. For a given code family and noise process, quantum information can be encoded and decoded arbitrarily well below a critical noise strength known as the *threshold*. Whilst the threshold is defined with respect to the optimal decoder, it is also often also studied for the case of specific sub-optimal decoding procedures.

The most commonly studied model of noise in quantum codes is that of *iid Pauli noise*, in which each qubit is subjected to an independent, identical Pauli noise process. This is a mathematically convenient model, but it does not account for any possible correlations between errors. Whilst this model can provide a proof-of-principle that an error correction procedure can successfully withstand errors, many of the physical architectures in which we might hope to implement quantum error correction are known to experience correlated noise. Examples include proximity effects [2.3, 2.4] and bosonic couplings [2.4–2.11] in solid state systems. Correlations can also arise when modelling non-Markovian noise processes [2.12, 2.13]. As the threshold depends on the error model, including the presence and magnitude of correlations, taking these factors into account is important when attempting to specify physically relevant thresholds.

One of the most important correlated noise models is that of *circuit-based noise* [2.14–2.19], in which elementary gates and measurement are taken to be noisy. Even if we assume the noise introduced in each operation is independent, the operations themselves tend to propagate and accumulate errors, giving an overall correlated noise model. Importantly, there exist codes which have a threshold under this correlated noise model, which is known as the *fault-tolerant threshold* [2.20, 2.21]. Moreover, it can even be shown that there exist error correction procedures which allow not only for fault tolerant storage of quantum information, but also fault tolerant quantum computation, a result known as the *quantum threshold theorem* [2.22–2.25].

In the context of correlated noise, little is known about optimal decoding or fault tolerance procedures. For this reason, most Monte Carlo estimates of thresholds for correlated noise are given with respect to sub-optimal decoders and fault tolerance schemes [2.26–2.36]. Indeed, to our knowledge, no optimal thresholds are known for any interesting quantum code families with a non-trivially correlated noise models.

Remarkably, there exists a method for computing code thresholds with respect to iid Pauli noise called the *statistical mechanical mapping* [2.37]. In this technique, a classical statistical mechanical system is constructed from a quantum code with the noise model manifesting as a quenched disorder. This mapping is designed in such a way that the thermodynamic properties of the statistical mechanical system relate to the error correction properties of the quantum code, under the optimal decoder. This method has been used [2.38] to compute thresholds [2.39–2.43], including fault-tolerance thresholds [2.44–2.49], for a wide variety of code families based on topological codes [2.47], i.e. codes with spatially local stabiliser generators on a lattice in Euclidean space.

Importantly, this link implies that the error correction threshold in the code manifests as a phase transition in the resulting stat mech model. This insight implies that one may bring to bear the various numerical and analytical techniques for determining the phases of stat mech systems to indirectly estimate the threshold of our code, *without* having to implement optimal decoding.

Summary of main results

In this manuscript we give a broad generalisation of the stat mech mapping to the case of correlated Pauli noise models acting on any stabiliser or subsystem quantum code. The original mapping for independent bit-flip noise [2.37], and subsequent generalisation for independent Pauli noise [2.39], works by showing that when a special condition known as the Nishimori condition [2.50] holds, a certain stat mech model with quenched disorder has a partition function that maps directly onto the probability of a logical class given the syndrome of a code. Our first result is that this fundamental

theorem continues to hold in much more generality for *correlated* noise using our more general stat mech mapping.

The specific notion of correlation allowed by our theorem is very general. We first show that any distribution arising from a *factor graph* admits such a mapping, which generalises independent noise as follows. In particular, it works whenever any cluster of errors that is sufficiently far apart is *conditionally* independent given the neighbouring spins of the clusters. This is the well-known spatial Markov condition, and it provides a systematic relaxation of the notion of independence, controlled by the length scale at which disjoint clusters become conditionally independent. When this length scale is zero, there are no intermediate spins on which to condition, and we recover the case of strictly independent noise. We make this notion precise in Definition 2.3 below.

We further generalise this to the case of spatio-temporally correlated noise by mapping to a system of one higher spatial dimension. This allows us to include the most relevant type of noise for quantum computing, namely circuit-based noise thresholds.

By performing Monte Carlo simulations of the resulting stat mech system, we then use this correspondence to approximate the threshold of the toric code subject to a correlated model of bit-flip errors, and quantify how a certain family of positively correlated errors affect the threshold.

Finally we will show how this mapping can also be used to give an efficient approximation of the maximum likelihood decoder. A consequence of this result is a generalisation of the tensor network decoder of Ref. [2.51] to any 2D surface code with spatially local noise correlations. The tensor network that yields the decoder can be approximately contracted in linear time in n, the number of qubits. The contraction sequence allows a systematic approximation of the maximum likelihood decoder by increasing the bond dimension cutoff of an intermediate matrix product state representation.

The paper is organised as follows. After introducing mathematical preliminaries and definitions in Section 2.1, we continue in Section 2.2 by reviewing the stat mech mapping in the case of independent noise. In Section 2.3 we then extend the mapping to account for noise models with spatial correlations. In Section 2.4 we apply this mapping to correlated bit-flip noise in the toric code, and perform Monte Carlo simulations to estimate the threshold. In Section 2.5 we extend our construction to spatio-temporal correlations and circuit-based noise. In Section 2.6 we prove that the phase boundary and the threshold coincide, and using this we show in Section 2.7 how tensor network methods yield efficient approximations of the optimal decoder. We conclude in Section 2.8, provide some background on correlated noise in Section 2.A, and details of our numerical simulations in Section 2.B.

2.1 Preliminaries

Fix a local dimension $d \in \mathbb{N}$. Let ω denote the fundamental dth root of unity, and let $\mathcal{P}$ denote the set of Pauli operators *modulo phase*. We will associate a copy of the Paulis with each site i for $1 \leq i \leq n$, denoting the Paulis supported on site i by $\mathcal{P}_i$, and the set of global Paulis by $\mathcal{P}^n := \bigtimes_{i=1}^n \mathcal{P}_i$.

Pauli operators always commute up to phase. As such, we can capture the non-commutativity of the Paulis via the *scalar commutator* $[\cdot,\cdot] : \mathcal{P} \times \mathcal{P} \to \mathbb{C}$, defined by the relation $AB = [A,B]\,BA$. It can also be seen as the normalised trace of the group commutator,

$$[A,B] := \frac{1}{d} \operatorname{Tr}[A,B], \tag{2.1}$$

where $[A, B] := ABA^{-1}B^{-1}$. We can see that the scalar commutator is a well-defined function on $\mathcal{P} \times \mathcal{P}$, as the group commutator is invariant under phases, $\left[e^{i\theta}A, e^{i\phi}B\right] = [A, B]$. In general the group commutator is not multiplicative[1], but conveniently the scalar commutator of Paulis *is*, in the sense that $[A, BC] = [A, B] \cdot [A, C]$. This also implies that Paulis commute under the scalar commutator, in the sense that $[A, BC] = [A, CB]$, which will be important for considering subsystem codes.

2.1.1 Stabiliser and Subsystem codes

The codes we will be considering in this work are Pauli stabiliser codes [2.52, 2.53]. A stabiliser code is defined by a subgroup $\mathcal{S}$ of the Paulis acting on n qu*d*its. This subgroup must be Abelian and have trivial overlap with the centre, $\mathcal{S} \cap \langle \omega I \rangle = \{I\}$. This group is typically specified in the form of a set of generators $\{S_k\}_k$. The associated code space is given by all states which are stabilised by every element of the stabiliser group, i.e. all $|\psi\rangle$ such that $s|\psi\rangle = |\psi\rangle$ for all $s \in \mathcal{S}$.

The logical (Pauli) operators of such a code are given by the centraliser of the stabiliser group, that is the set of all $L \in \mathcal{P}^n$ such that $[s, L] = +1$ for all stabilisers $s \in \mathcal{S}$. As the stabilisers act trivially on the code space, logicals which differ just by a stabiliser have identical action on the code space, and we will refer to them as *logically equivalent*.

Suppose that we start in a code state $|\psi\rangle$, and our system suffers an error $E \in \mathcal{P}^n$, leaving it in state $E|\psi\rangle$. The first step in the error correction procedure begins by measuring each of the stabiliser generators $\{S_k\}_k$. The Pauli error model has the key feature that any given error yields a deterministic outcome. Specifically, measuring the kth stabiliser generator deterministically gives the outcome $[S_k, E]$, and therefore does not disturb the state. Importantly, this outcome does not depend on the initial code state $|\psi\rangle$, depending only on the error E. We refer to the collection of all such measurement outcomes as the *syndrome*.

A decoder for such a code is an algorithm which takes as input the syndrome, and outputs a decoding Pauli D. We then apply D^{-1} to the state, in the hope that this corrects the error. This decoding successfully restores the system back to the original code state if and only if D was logically equivalent to the true error E, i.e. $D^{-1}E \in \mathcal{S}$.

A more general notion of quantum code is that of a subsystem code [2.54]. Here, some fraction of the logical qubits are sacrificed to become so-called gauge qubits. Gauge qubits then provide a workspace that simplifies some of the measurements, since instead of measuring a stabiliser directly one can measure combined gauge and stabiliser operators that might have lower weight. In a subsystem code, the gauge group is generated by all of the stabiliser and the gauge operators, and the stabiliser group is the centre of the gauge group (modulo phase). Thus, measuring enough gauge generators to reconstruct the stabilisers is sufficient to perform quantum error correction in a subsystem code. This is true even though general elements of the gauge group don't commute, since the stabiliser elements *do* commute by virtue of living in the centre. We refer the reader to Ref. [2.54] for a more detailed discussion of subsystem codes.

In what follows, we never use the fact that the stabiliser generators commute. We only use the fact that Pauli errors form an abelian error algebra, since they commute modulo the scalar commutator. That is, given two Pauli operators P and Q, the accumulated error on a state is the same if one applies PQ or if one applies QP. Because of this fact, everything that we derive below also applies to

[1] For a general group, $[A, BC] = [A, B]\, B\, [A, C]\, B^{-1}$.

subsystem codes where gauge generators are used in place of stabiliser generators. With the exception of Section 2.5, we will use the simpler language of stabiliser codes throughout, but with the understanding that the results can easily be generalised to subsystem codes.

2.2 The statistical mechanical mapping

Before considering correlated noise, we start by reviewing the case of independent noise, as first considered in Ref. [2.37]. Although this material is review, our goal is to write the derivation of the independent case in such a way that the correlated case falls out as naturally as possible.

For notational convenience, we will restrict out attention to stabiliser codes for this section, but we note that the below construction can be naturally extended to subsystem codes by replacing stabiliser generators by gauge generators.

Consider a code given by a set of stabiliser generators $\{S_k\}_k$. In this section we will be considering an independent Pauli noise model: let $\{p_i\}_i$ be probability distributions on $\mathcal{P}_i$ which describe the probability of a Pauli error (independently) occurring at each site i. The probability of an overall error $E \in \mathcal{P}^n$ is therefore given by

$$\Pr(E) = \prod_{i=1}^{n} p_i(E_i), \tag{2.2}$$

where E_i denotes the action of E within $\mathcal{P}_i$.

2.2.1 Statistical mechanical model

We now want to develop a (classical) spin model whose statistical mechanical properties capture the error correction properties of our quantum code, in such a way that that the threshold in the latter naturally corresponds to a phase transition in the former.

The state space of this model will correspond to the stabiliser group, with the noise model determining the interactions. Specifically, associate a classical spin degree of freedom c_k with each stabiliser generator S_k. We will consider each degree of freedom as a member of $\mathbb{Z}_d$, and so our full state space[2] is given by $\times_k \mathbb{Z}_d$.

The family of Hamiltonians we will consider is defined as follows.

Definition 2.1 (Stat mech Hamiltonian: *independent noise*)**.** For a Pauli $E \in \mathcal{P}^{\times n}$, and coupling strengths $\{J_i : \mathcal{P}_i \to \mathbb{R}\}_i$, our Hamiltonian H_E is defined as

$$H_E(\vec{c}) = -\sum_{i,\sigma\in\mathcal{P}_i} J_i(\sigma)\,[\sigma, E] \prod_k [\sigma, S_k]^{c_k}, \tag{2.3}$$

for any state $\vec{c} \in \times_k \mathbb{Z}_d$, where E forms a (quenched) disorder parameter. Here the sum is over all qudits i and all elements σ in the local Pauli group $\mathcal{P}_i$.

[2] In the case where d is composite, and some of the generators S_k are not of maximal order, this will actually cause an over-counting in Theorem 2.1. Each degree of freedom should take on a number of states equal to the order of S_k. For notational convenience we will assume all of our generators S_k are of maximal order, as is always the case for prime d.

For algebraic manipulation, it will be convenient to note that the above Hamiltonian can also be written as

$$H_E(\vec{c}) = -\sum_{i,\sigma\in\mathcal{P}_i} J_i(\sigma)\left[\sigma, E\prod_k S_k^{c_k}\right], \tag{2.4}$$

due to the multiplicativity of the scalar commutator.

Physically, these Hamiltonians correspond to random bond Ising-type models for $d = 2$. This can be seen by putting the Hamiltonian in terms of the degrees of freedom $s_k := (-1)^{c_k}$,

$$H_E(\vec{s}) = -\sum_{i,\sigma\in\mathcal{P}_i} \overbrace{J_i(\sigma)}^{\text{Strength}}\ \overbrace{[\sigma, E]}^{\text{Disorder}}\ \overbrace{\prod_{k:[\sigma,S_k]=-1} s_k}^{\text{Interactions}}. \tag{2.5}$$

Similarly for $d > 2$ this can be seen as a type of random-bond vector Potts model (or clock model).

Readers familiar with the prior work on stat mech Hamiltonians (beginning with [2.37]) might be puzzled as to why we write such a convoluted form of H_E for our definition rather than the straightforward Ising-type model in Equation (2.5). The answer is that, as we will see below, this formulation gives the simplest path to generalising these results to the case of correlated noise.

We note that if our stabiliser generators are local, in the sense that each site is only acted upon non-trivially by a finite number of stabiliser generators, then this Hamiltonian is also local, in the sense that each interaction only touches a finite number of sites.

In this construction, each of our interactions corresponded to a site i and a $\sigma \in \mathcal{P}_i$, i.e. a single-site Pauli. As we will see in Section 2.3, this is intimately linked with the restriction to independent noise, and that including larger range interactions will allow us to account for correlated error models.

Before considering specific coupling strengths, we note that this model has been constructed such that it is symmetric under multiplying the disorder E by a stabiliser generator S_k and adding to the corresponding degree of freedom c_k. Specifically,

$$H_{ES_k}(\vec{c}) = H_E(\vec{c} + \hat{k}). \tag{2.6}$$

This can be seen by noting that our Hamiltonian only depends on each c_k via $ES_k^{c_k}$, which itself has this symmetry.

2.2.2 Nishimori condition

We now want to consider the coupling strengths required such that the above model reproduces the statistical properties of our quantum code and noise model. Specifically we want to choose these couplings such that

$$Z_E = \Pr(\overline{E}), \tag{2.7}$$

where Z_E is the partition function of the Hamiltonian H_E with quenched disorder E, $\Pr(\overline{E})$ denotes the probability of a logical class of errors, and $\overline{E}$ denotes the set of errors which are logically equivalent to E. A sufficient condition for this is given by the *Nishimori conditions*.

Definition 2.2 (Nishimori conditions: *independent noise*)**.** An inverse temperature β and coupling strengths $\{J_i\}_i$ satisfy the *Nishmori conditions* with respect to distributions $\{p_i\}_i$ if

$$\beta J_i(\sigma) = \frac{1}{|\mathcal{P}|} \sum_{\tau \in \mathcal{P}_i} \log p_i(\tau) \left[\sigma, \tau^{-1}\right], \tag{2.8}$$

holds for all i and $\sigma \in \mathcal{P}_i$.

Note that Equation (2.8) is not defined if any of the $p_i(\tau)$ are exactly zero. However, in those cases, we can formally manipulate $\log 0$ together with the convention that $\exp(\log 0) = 0$ and obtain sensible answers. A more rigorous treatment taking limits is certainly possible, but would obscure the thrust of the argument, so we neglect these details.

This now allows us to prove the critical property of our stat mech mapping, which we refer to as the *fundamental theorem.*

Theorem 2.1 (Fundamental theorem of stat mech mapping: *independent noise*)**.** *Given the stat mech Hamiltonian (Equation (2.3)) associated to a noise model that satisfies the Nishimori condition (Equation (2.8)), the probability of a logical class of errors is equal to the corresponding partition function,*

$$Z_E = \Pr(\bar{E}). \tag{2.9}$$

Proof. First we note that the Nishimori condition takes the form of an (inverse) Fourier transform of the log-probabilities $\log p_i$ with respect to the Pauli group. Using the orthogonality condition $\sum_\sigma [\sigma, \rho] = |\mathcal{P}|\delta_{\rho,I}$, we can see that this is equivalent to the requirement that the Fourier transform of βJ_i is $\log p_i$, specifically

$$\sum_{\sigma \in \mathcal{P}_i} \beta J_i(\sigma) \left[\sigma, E\right] = \log p_i(E_i). \tag{2.10}$$

Summing over sites i, we see that this means the Gibbs weight in the all-zero state reproduces the error probability,

$$e^{-\beta H_E(\vec{0})} = \exp\left(\sum_{i, \sigma \in \mathcal{P}_i} \beta J_i(\sigma) \left[\sigma, E\right]\right) \tag{2.11a}$$

$$= \exp\left(\sum_i \log p_i(E_i)\right) \tag{2.11b}$$

$$= \prod_i p_i(E_i) = \Pr(E). \tag{2.11c}$$

Lastly, we can use the symmetry property to write the partition function as a sum over logically equivalent disorders, giving the probability of logical class as required,

$$Z_E = \sum_{\vec{c}} e^{-\beta H_E(\vec{c})} = \sum_{S \in \mathcal{S}} e^{-\beta H_{ES}(\vec{0})} = \Pr(\overline{E}). \tag{2.12}$$

□

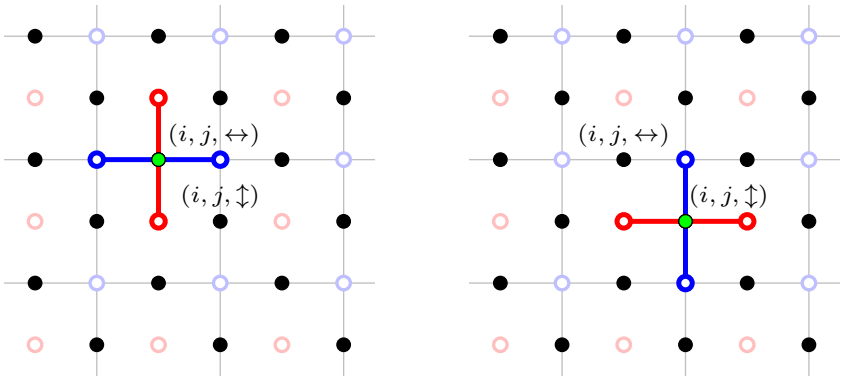


Figure 2.1: Stat mech mapping of toric code with iid noise. Solid circles indicate qubits, and hollow circles indicate the spins of the stat mech model. The lines connecting the spins to the highlighted green qubit indicate the spins which are coupled by the interactions corresponding to that qubit. The interaction terms corresponding to a qubit on a horizontal edge $(i, j, \leftrightarrow)$ give a horizontal Ising coupling on the X-sublattice, a vertical Ising coupling on the Z-sublattice, and a four-body term coupling the two sub-lattices. The vertical qubit $(i, j, \updownarrow)$ gives rotated versions of these couplings. The full Hamiltonian is given in Equations (2.13) and (2.14).

This correspondence between logical class probabilities and partition functions suggests that the regimes in which error correction is and is not possible in our code correspond naturally to phases in our stat mech model. We will show this explicitly in Section 2.6, and see that the error correcting threshold manifests as a quenched phase transition.

In Section 2.7 we will also see how this correspondence, together with tensor network methods for approximating partition functions, can be used to construct a family of efficient tensor network algorithms which approximate maximum likelihood decoding.

2.2.3 Example: Toric code

We now consider several examples of the stat mech systems corresponding to the toric code with iid noise. Let p be the marginal distribution of errors on a single qubit. Given that our stabilisers split into X-type stars and Z-type plaquettes, we will divide the spins into two corresponding sub-lattices, denoting them $\{s_k^X\}_k$, the X-sublattice, or $\{s_k^Z\}_k$, the Z-sublattice, respectively. Applying our mapping to this code gives the Hamiltonian

$$\begin{aligned} H_E = -\sum_e \Big(& J(I) + J(X)\,[X, E_e] \prod_{\partial f \ni e} s_f^Z \\ & + J(Z)\,[Z, E_e] \prod_{v \in \partial e} s_v^X \\ & + J(Y)\,[Y, E_e] \prod_{\partial f \ni e} s_f^Z \prod_{v \in \partial e} s_v^X \Big), \end{aligned} \tag{2.13}$$

where v, e, f denote vertices, edges, and faces in the lattice, ∂e denote the vertices surrounding an edge, and ∂f the edges surrounding a face. We note that the above Hamiltonian is not only valid for the toric code, but in fact *any* homology code.

Writing this Hamiltonian out more explicitly in Cartesian coordinates, as indicated in Figure 2.1, we see that our model corresponds to two copies of the 2D random bond Ising model on each sublattice, with a four-body coupling between them, specifically

$$\begin{aligned} H_E = -\sum_{i,j}\Big(&2J(I)+J(X)\,[X,E_{i,j,\leftrightarrow}]\,s^Z_{i,j}s^Z_{i,j+1} \\ &+J(X)\,\left[X,E_{i,j,\updownarrow}\right]\,s^Z_{i,j}s^Z_{i+1,j} \\ &+J(Z)\,[Z,E_{i,j,\leftrightarrow}]\,s^X_{i,j}s^X_{i-1,j} \\ &+J(Z)\,\left[Z,E_{i,j,\updownarrow}\right]\,s^X_{i,j}s^X_{i,j-1} \\ &+J(Y)\,[Y,E_{i,j,\leftrightarrow}]\,s^Z_{i,j}s^Z_{i,j+1}s^X_{i,j}s^X_{i-1,j} \\ &+J(Y)\,\left[Y,E_{i,j,\updownarrow}\right]s^Z_{i,j}s^Z_{i+1,j}s^X_{i,j}s^X_{i,j-1}\Big), \end{aligned} \tag{2.14}$$

where $E_{i,j,\leftrightarrow}$ and $E_{i,j,\updownarrow}$ indicate the action of the error E on the (i,j)th horizontal and vertical qubits (see Figure 2.1). The terms inside the parentheses are the interactions corresponding to the qubits labelled $(i,j,\leftrightarrow)$ and $(i,j,\updownarrow)$ as indicated in Figure 2.1.

The Nishimori conditions give coupling strengths of the form

$$J(I) = \frac{1}{4\beta}\log p(I)p(X)p(Y)p(Z), \tag{2.15a}$$

$$J(X) = \frac{1}{4\beta}\log\frac{p(I)p(X)}{p(Y)p(Z)}, \tag{2.15b}$$

$$J(Z) = \frac{1}{4\beta}\log\frac{p(I)p(Z)}{p(X)p(Y)}, \tag{2.15c}$$

$$J(Y) = \frac{1}{4\beta}\log\frac{p(I)p(Y)}{p(X)p(Z)}. \tag{2.15d}$$

We now consider this model for specific iid error models of interest.

Depolarising noise (Bombin et.al.)

If we consider the depolarising channel, with $p(I) = 1-p$ and $p(X) = p(Y) = p(Z) = p/3$, then our model reduces to

$$\begin{aligned} H_E = -\sum_{i,j}\Big(&-K + J\,[X,E_{i,j,\leftrightarrow}]\,s^Z_{i,j}s^Z_{i,j+1} \\ &+J\,\left[X,E_{i,j,\updownarrow}\right]\,s^Z_{i,j}s^Z_{i+1,j} \\ &+J\,[Z,E_{i,j,\leftrightarrow}]\,s^X_{i,j}s^X_{i-1,j} \\ &+J\,\left[Z,E_{i,j,\updownarrow}\right]\,s^X_{i,j}s^X_{i,j-1} \\ &+J\,[Y,E_{i,j,\leftrightarrow}]\,s^Z_{i,j}s^Z_{i,j+1}s^X_{i,j}s^X_{i-1,j} \\ &+J\,\left[Y,E_{i,j,\updownarrow}\right]\,s^Z_{i,j}s^Z_{i+1,j}s^X_{i,j}s^X_{i,j-1}\Big), \end{aligned} \tag{2.16}$$

with Nishimori conditions

$$\beta K = \frac{1}{2}\log\frac{27}{p^3(1-p)}\,, \tag{2.17a}$$

$$\beta J = \frac{1}{4}\log\frac{3(1-p)}{p}\,. \tag{2.17b}$$

This corresponds to a disordered variant of the eight-vertex model [2.55–2.57], which was considered in Ref. [2.39].

Independent X and Z

Consider now a model with independent X and Z errors. Specifically let p_X, p_Z denote the probability of each *generator*, such that

$$p(I) = (1-p_X)(1-p_Z), \tag{2.18a}$$

$$p(X) = p_X(1-p_Z), \tag{2.18b}$$

$$p(Z) = (1-p_X)p_Z, \tag{2.18c}$$

$$p(Y) = p_X p_Z. \tag{2.18d}$$

Importantly, this means that $p(X)p(Z) = p(I)p(Y)$, and therefore $J(Y) = 0$. This has the effect of decoupling the two sub-lattices into two non-interacting random-bond Ising models, such that the Hamiltonian can be decomposed $H_E = H_E^X + H_E^Z$, where

$$H_E^X = -\sum_{i,j}\Big(-K^X + J^X\left[Z, E_{i,j,\leftrightarrow}\right] s^X_{i,j}s^X_{i-1,j} + J^X\left[Z, E_{i,j,\updownarrow}\right] s^X_{i,j}s^X_{i,j-1}\Big), \tag{2.19a}$$

$$H_E^Z = -\sum_{i,j}\Big(-K^Z + J^Z\left[X, E_{i,j,\leftrightarrow}\right] s^Z_{i,j}s^Z_{i,j+1} + J^Z\left[X, E_{i,j,\updownarrow}\right] s^Z_{i,j}s^Z_{i+1,j}\Big), \tag{2.19b}$$

and

$$K^A = \frac{1}{2\beta}\log\frac{1}{p_A(1-p_A)}\,, \tag{2.20a}$$

$$J^A = \frac{1}{2\beta}\log\frac{1-p_A}{p_A}\,, \tag{2.20b}$$

for $A \in \{X, Z\}$. This corresponds to two decoupled copies of the random-bond Ising model, with disorder probabilities p_X and p_Z. The decoupled nature of these two models is generic for CSS codes (i.e., codes whose stabilisers split into separate X and Z type) under independent X and Z noise. A further consequence of this is that the optimal decoder can decode the X and Z errors independently.

Pure bit-flip noise (Dennis et.al.)

As mentioned when we introduced the Nishimori condition above, care must be taken whenever the noise model does not have full support, as our Hamiltonian becomes divergent on certain states. An

important example of such a model is that of iid bit-flip noise, as was considered in the seminal paper Ref. [2.37]. Here we take $p(I) = 1 - p$, $p(X) = p$, and $p(Y) = p(Z) = 0$.

We know that the Nishimori condition implies

$$H_E(\vec{c}) = -\frac{1}{\beta} \log \Pr\left(E \prod_k S_k^{c_k} \right). \tag{2.21}$$

Given that the probability of any error containing any non-trivial Z contribution is zero in this error model, this tells us that any state which is *not* entirely magnetised on the Z-sub-lattice ($s^Z_{i,j} s^Z_{i+1,j} = -1$ or $s^Z_{i,j} s^Z_{i,j+1} = -1$ for some i, j) has infinite energy. We can interpret this as the degrees of freedom in this sub-lattice being *frozen out*, into the one of the two entirely magnetised states. Restricting to these (degenerate) Z-magnetised states ($s^Z_{i,j} s^Z_{i+1,j} = s^Z_{i,j} s^Z_{i,j+1} = +1$ for all i, j), our Hamiltonian therefore reduces to a single copy of the random bond Ising model

$$\begin{aligned} H_E = -\sum_{i,j} \Big(& -2K + J\left[Z, E_{i,j,\leftrightarrow}\right] s^X_{i,j} s^X_{i-1,j} \\ & + J\left[Z, E_{i,j,\updownarrow}\right] s^X_{i,j} s^X_{i,j-1} \Big), \end{aligned} \tag{2.22}$$

where

$$K = \frac{1}{2\beta} \log \frac{1}{p(1-p)}, \tag{2.23a}$$

$$J = \frac{1}{2\beta} \log \frac{1-p}{p}. \tag{2.23b}$$

We can see now that J and K are both finite for any $p \in (0, 1)$, meaning that our Hamiltonian is no longer divergent on the remaining degrees of freedom. Moreover, this is the Hamiltonian that was considered in Ref. [2.37]. This shows that, taking appropriate care, errors models without full support can also be considered using this construction, after considering frozen out degrees of freedom.

2.2.4 Extensions

The stat mech mapping for independent noise can also be generalised to several other noise models.

Noisy measurements The first example is a noise model consisting of independent noise, and independently noisy measurements. This can be modelled by including an ancilla bit for each stabiliser, and replacing stabilisers $S \to S \otimes Z$, where the Z acts on this ancilla. Any bit-flip noise on this ancilla bit will effectively model noisy measurements.

Leakage errors Leakage can be accounted for by explicitly including flag bits. The precise construction however will depend on what model of leakage is being used.

Overlapping independent One simple toy model for spatially correlated errors is a model in which independent noise processes acts on overlapping regions, where the overall error is given by the product of these local errors. An example of this is the 'nearest-neighbor depolarizing' model of Ref. [2.89], in which each nearest-neighbour pair is afflicted uniformly by any non-trivial 2 qubit error with probability $1 - p$. We discuss modifying the stat mech mapping for models such as this in Section 2.C.

2.3 Correlated noise

Now that we have reviewed the stat mech construction for *independent* Pauli noise, we now want to consider extending this to *correlated* noise models.

We will consider a model with local spatial correlations, which forms a natural and systematic way of relaxing the independence condition above. Specifically, we want to consider models in which errors at sufficiently distant locations are conditionally independent, but in which they may be arbitrarily correlated at short range. The conditional dependences described above are naturally represented through *probabilistic graphical models*, such as Markov random fields and Bayesian networks. Below we will focus on a simpler and more mathematically convenient model of *factored distributions*. We discuss the relationship between these models in more detail in Section 2.A.

2.3.1 Factored distributions

For a global error $E \in \mathcal{P}^n$, let E_i denote the action on a single site i. Similarly, for some set of sites $R = \{i_1, i_2, \dots\}$, let E_R denote action of E on sites in R, i.e. $E_R := \otimes_{i \in R} E_i$.

For the previous mapping, we leveraged the fact that independence of two random variables A and B implies their joint distribution factors, $\Pr(A, B) = \Pr(A) \cdot \Pr(B)$. We now want to consider more general distributions which can also be locally factored.

Definition 2.3 (Factored distribution)**.** A distribution *factors* over sets $\{R_j\}_j$ if there exists non-negative functions $\{\phi_j : \mathcal{P}_{R_j} \to \mathbb{R}^+\}_j$ such that

$$\Pr(E) = \prod_j \phi_j(E_{R_j}). \tag{2.24}$$

If the regions $\{R_j\}_j$ are disjoint, then these just correspond to distributions which are independent after appropriate coarse-graining. We, however, *do not* require that these regions are disjoint, which allows us to consider genuinely correlated noise models, which remain correlated even after coarse-graining. In another extreme limit, if there is only a single region R and it contains *all* of the random variables, then we can take ϕ_R to be simply the complete joint probability distribution. This shows that the structure of the individual regions R_j can interpolate between the case of independent noise (when $R_j = \{j\}$) and a general distribution (when $R = \{1, \dots, n\}$). Intermediate cases correspond to probability distributions with differing ranges of correlation. Thus the factored distribution formalism forms a natural ansatz for describing finite-range correlations efficiently in a probabilistic model. We refer the reader to Section 2.A for more discussion of factored distributions.

2.3.2 Correlated statistical mechanical mapping

We now extend the stat mech mapping to the case of correlated noise. Specifically we will consider an error model which factors over $\{R_j\}_j$, with factors $\{\phi_j\}_j$. In the mapping for independence noise, each interaction term corresponding to a single-site Pauli. As hinted at earlier, we can account for correlated noise models by including interaction terms corresponding to multi-site σ, specifically σ living on a single region R_j. Recall that S_k are the stabilisers of the quantum code and c_k are the associated d-level classical degrees of freedom taking values in $\mathbb{Z}_d$. Then the stat mech Hamiltonian takes the following form.

Definition 2.4 (Stat mech Hamiltonian: *correlated noise*)**.** For a Pauli $E \in \mathcal{P}^n$, and coupling strengths $\{J_j : \mathcal{P}_{R_j} \to \mathbb{R}\}_j$, the stat mech Hamiltonian H_E is defined as

$$H_E(\vec{c}) = -\sum_{j,\sigma \in \mathcal{P}_{R_j}} J_j(\sigma)\,[\sigma, E] \prod_k [\sigma, S_k]^{c_k}. \tag{2.25}$$

We notice that we can once again use the multiplicativity of the scalar commutator to rewrite this in the more mathematically convenient form

$$H_E(\vec{c}) = -\sum_{j,\sigma \in \mathcal{P}_{R_j}} J_j(\sigma) \left[\sigma, E \prod_k S_k^{c_k}\right], \tag{2.26}$$

and see that our Hamiltonian retains the symmetry we leveraged in the independent case,

$$H_{ES_k}(\vec{c}) = H_E(\vec{c} + \hat{k}). \tag{2.27}$$

In order for the definition of the stat mech Hamiltonian to connect to the error probabilities of the quantum code for a given noise model, we need to find an analogue of the Nishimori conditions. Fortunately, our formulation of this condition in the independent case gives an immediate generalisation.

Definition 2.5 (Nishimori conditions: *correlated noise*)**.** An inverse temperature β and coupling strengths $\{J_j\}_j$ satisfy the *Nishimori conditions* with respect to factors $\{\phi_j\}_j$ if

$$\beta J_j(\sigma) = \frac{1}{|\mathcal{P}_{R_j}|} \sum_{\tau \in \mathcal{P}_{R_j}} \log \phi_j(\tau) \left[\sigma, \tau^{-1}\right], \tag{2.28}$$

Using a proof analogous to the independent case Theorem 2.1, we also get a *fundamental theorem* for this correlated model.

Theorem 2.2 (Fundamental theorem of stat mech mapping: *correlated noise*)**.** *Given the stat mech Hamiltonian (Equation (2.25)) associated to a noise model that satisfies the Nishimori condition (Equation (2.28)), the probability of a logical class of errors is equal to the corresponding partition function,*

$$Z_E = \Pr(\bar{E}). \tag{2.29}$$

As with the independent case, we will see that this similarly implies that the error-correction threshold manifests as a quenched phase transition (Section 2.6), and that this allows us to construct efficient tensor network approximations to the maximum likelihood decoder (Section 2.7).

2.3.3 Noise Hamiltonian

Consider the case where each factor ϕ_j is strictly positive. Suppose we define the local Hamiltonian

$$\tilde{H}(E) := -\sum_j \log \phi_j(E), \tag{2.30}$$

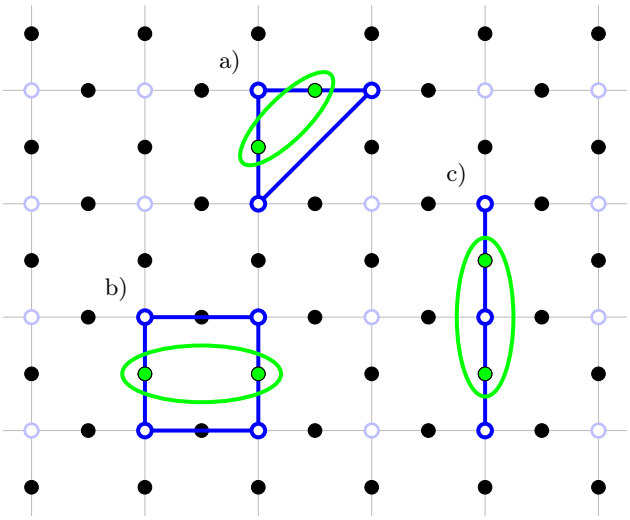


Figure 2.2: Stat mech couplings induced by pairwise correlated bit-flips in the toric code. Solid circles indicate qubits, and hollow circles the spins of our stat mech model, corresponding to stabiliser generators. Green ellipses denote the type of correlations, specifically the regions over which the error model factors. These correspond to the interactions in the noise Hamiltonian Equation (2.33). Blue edges connecting spins indicate the couplings induced in the statistical mechanical model by such correlations. The labelled couplings a), b) and c) respectively correspond to: a) nearest-neighbour correlations induce a two-body diagonal coupling, b) next-nearest neighbour (across-plaquette) correlations induce a four-body face coupling, c) next-nearest-neighbour (across-vertex) induce two-body distance-2 couplings.

which we refer to as the *noise Hamiltonian*. The states of this Hamiltonian are labelled by Pauli errors, elements of $\mathcal{P}^n$. The noise model then corresponds to thermal distribution of this Hamiltonian at inverse-temperature $\beta = 1$,

$$\Pr(E) = e^{-\tilde{H}(E)}. \tag{2.31}$$

Expressed in this way, the Nishimori condition, Definition 2.5, can be seen as a relationship between the stat mech Hamiltonian of Definition 2.4 and the noise Hamiltonian of Equation (2.30). Specifically, it takes the form

$$\beta H_E(\vec{c}) = \tilde{H}\left(E\prod_k S_k^{c_k}\right). \tag{2.32}$$

In this sense we see that, when the Nishimori condition is satisfied, the interactions within the stat mech Hamiltonian H_E naturally correspond to those in the noise Hamiltonian $\tilde{H}$, but with a change in the underlying state space.

2.3.4 Example: Toric code with correlated bit flips

We now consider an example of a correlated noise model: correlated bit-flips in the toric code. Consider a noise model defined by an Ising noise Hamiltonian with coupling $\tilde{J}$ and field strength $\tilde{h}$,

$$\tilde{H} = -\sum_e \tilde{h}\, x_e - \sum_{e\sim e'} \tilde{J}\, x_e x_{e'}, \tag{2.33}$$

where we have chosen the convention that $x_e = -1$ corresponds to $E_e = X$ and $x_e = +1$ to $E_e = I$. Here $\tilde{J}$ controls the magnitude (and sign) of the correlations, with $\tilde{J} = 0$ corresponding to independent errors, $\tilde{J} > 0$ to bunching errors, and $\tilde{J} < 0$ to anti-bunching errors.

For a given error E (and the corresponding values of the spin variables x_e), the stat mech model is of the form

$$H_E = -\sum_e (h_e x_e) \prod_{v\in\partial e} s_v - \sum_{e\sim e'} (J x_e x_{e'}) \prod_{v\in\partial(ee')} s_v, \tag{2.34}$$

where $\partial(ee')$ denotes the vertices that surround either e or e', but not both. Here the variables s_v form the degrees of freedom, and the x_e form the quenched disorder. The Nishimori conditions for this model reduce to $\beta J = \tilde{J}$ and $\beta h = \tilde{h}$.

As in the independent case, we can see that the $\tilde{h}$ field term in our noise Hamiltonian has induced a 2D random-bond Ising model. The addition of the $\tilde{J}$ term has induced additional longer range couplings. The geometry of these couplings is shown is Figure 2.2 for the case where the noise correlations couple nearest-neighbour and next-nearest-neighbour qubits. The corresponding stat mech Hamiltonian has a similar locality, and contains at most 4-body interactions among the s_v degrees of freedom.

2.4 Numerics

A key advantage of the stat mech mapping is that it allows us to reappropriate techniques for determining the phase diagrams of classical spin systems for approximating the thresholds of quantum codes. By way of example, we consider using Monte Carlo simulations to determine the threshold of the toric code under a correlated model of bit-flip noise.

We shall consider the model of 'across-plaquette' correlated bit-flips described in Section 2.3.4 (see Figure 2.2b). We will restrict our attention to noise which obeys certain natural symmetries, namely, that the correlations are site-independent and symmetric between correlated errors. This is equivalent to saying that the noise Hamiltonian has symmetric, site-independent interaction terms.

Instead of expressing this model in terms of an Ising noise Hamiltonian, it will be convenient to parameterise this model in terms of the marginal error rate. Let p denote the marginal error probability on any site, i.e. $\Pr(E_e = X) = p$ on any edge e. Suppose that the error probability given that a neighbouring error has or has not occurred is $p_\pm$, so that

$$\Pr(E_e = X | E_{e'} = X) = p_+, \tag{2.35a}$$
$$\Pr(E_e = X | E_{e'} = I) = p_-, \tag{2.35b}$$

where e and e' lie on opposite sides of a plaquette. For the marginal probability to be p, these probabilities are subject to the consistency condition $p = pp_+ + (1-p)p_-$.

If $p_- < p < p_+$ the errors tend to bunch together, whereas they tend to anti-bunch if $p_- > p > p_+$. A natural way of parameterising these correlations is in terms of the *correlation parameter*,

$$\eta := p_+/p_-, \tag{2.36}$$

where $\eta = 1$ corresponds to uncorrelated noise, $\eta > 1$ to bunched, and $\eta < 1$ to anti-bunched.

In the limit of infinite bunching ($\eta \to \infty$) the model produces exclusively logical errors, meaning that the threshold vanishes, $p_\mathrm{t} \to 0\%$. Similarly, for infinite anti-bunching ($\eta \to 0$) the model cannot produce non-trivial logical errors, and we expect that $p_\mathrm{t} \to 50\%$ [2.58]. We will be considering the case of mild bunching correlations, $\eta = 2$, which one would expect to lower the threshold.

Applying the stat mech mapping, we get a random-bond Ising-type model on a square lattice, containing 2-body edge terms and 4-body face terms. The Hamiltonian takes the form

$$H_E(\vec{s}) = -\sum_e (J_2 x_e) \prod_{v\in\partial e} s_v - \sum_{e\sim e'} (J_4 x_e x_{e'}) \prod_{v\in\partial(ee')} s_v, \tag{2.37}$$

where $x_e = +1$ if $E_e = I$ and $x_e = -1$ if $E_e = X$, and $e \sim e'$ denotes edges lying across a plaquette from each other.

Normalising our Hamiltonian such that $J_2 \equiv 1$, our system has two Nishimori conditions. When $\eta = 2$, the first of these conditions is

$$J_4 = \frac{\log \frac{1-p}{2}}{\log 4p^2}. \tag{2.38}$$

We will impose this condition, and determine the phase diagram of the system in the two remaining parameters: the error probability p and temperature T. The threshold probability can be found by finding the intersection of this phase boundary with the second remaining Nishimori condition,

$$\beta_\mathrm{Nish} = -\log 4p^2. \tag{2.39}$$

In Figure 2.3 we show the phase diagrams of the $\eta = 1$ uncorrelated model, and our numerical results for the $\eta = 2$ correlated model. The details of our numerical simulation are found in Section 2.B. We find that that under these mild correlations the threshold drops to $p_\mathrm{t}(\eta = 2) = 10.04(6)\%$, from the uncorrelated threshold of $p_\mathrm{t}(\eta = 1) = 10.917(3)\%$ [2.59], confirming the earlier intuition that correlated errors will indeed reduce the threshold.

2.5 Spatio-temporal correlations

In Section 2.3 we considered the stat mech mapping for error models which factor. This model applies in the important case of spatially correlated errors, followed by ideal measurements. We would like to generalise this construction to the case of multiple rounds of syndrome measurements, subject to *spatio-temporal* correlations in the noise. We will do this by constructing a subsystem code which, subject to purely spatially correlated noise, reproduces the measurement statistics of our original stabiliser code

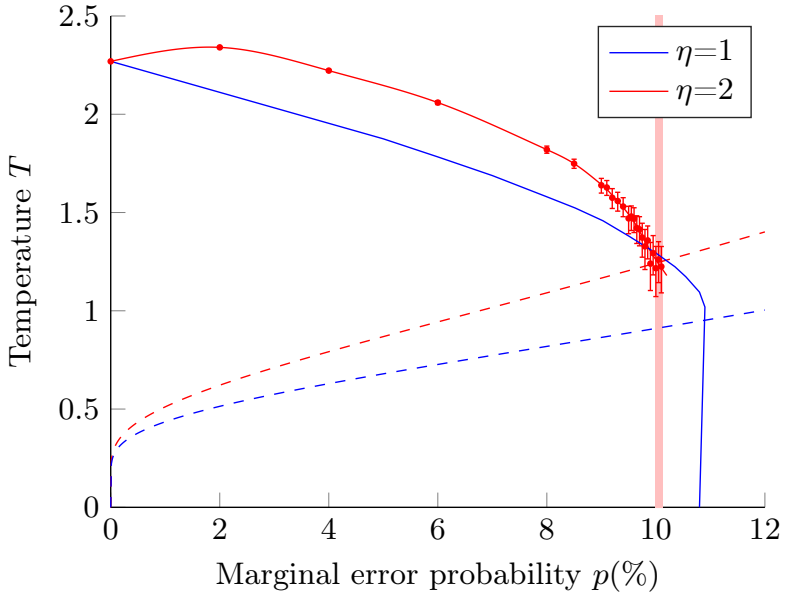


Figure 2.3: Phase boundary of the stat mech models corresponding to uncorrelated ($\eta = 1$) and correlated ($\eta = 2$) bit-flip noise in the toric code. The solid lines indicate the phase boundaries of the two models, with error bar indicating statistical uncertainty in our numerical results. The $\eta = 1$ data is taken from Refs. [2.37,2.59–2.62], and the $\eta = 2$ comes from our numerics, as detailed in Section 2.B. The dashed lines indicate the corresponding Nishimori conditions. The shaded red region indicates our estimate of $p_{\mathrm{t}} = 10.04(6)\%$ for the threshold of the correlated model.

under a spatio-temporally correlated noise model. We refer to this construction as the *history code*. Doing this in such a way that the locality of our code and noise model are preserved, we can then apply the construction of Section 2.3 to the history code to give a corresponding stat mech model.

As well as presenting the construction of the history code, we will also discuss an important family of spatio-temporally correlated noise models which possess a correlation structure which make them amenable to this construction: circuit-based noise. In this model we consider a syndrome measurement procedure composed of faulty gates. The key distinction between this and purely spatial correlations is that errors incurred at an earlier time can be spread around by subsequent measurement circuits. Considering Clifford measurement circuits subject to Pauli noise, this gives a spatio-temporally correlated Pauli noise model. By applying the construction of Section 2.3 to the corresponding history code, we get a stat mech model with a phase transition corresponding to the *fault-tolerant threshold* of the original code.

2.5.1 History code

The idea behind the history code is to convert time into an additional spatial dimension, allowing us to naturally convert spatio-temporal correlations into purely spatial correlations, albeit living in one dimension higher. Specifically, for each site i in our original code, and measurement round t, the history code will have a corresponding site (i, t).

Noise model

The error models we shall consider are those which factor in a spatio-temporal sense. Specifically, let $E^{(t)}$ denote the error to which our code is subjected *prior* to the tth round of measurements, and $E := (E^{(1)}, \ldots, E^{(T)})$ denote the *error history*. We will consider models in which the distribution of error histories factors, i.e.

$$\Pr(E) = \prod_l \phi_l(E), \tag{2.40}$$

where the supports of ϕ_l are local in both space *and* time.

As the name may suggest, we will take our error model on the history code to be given by the error *histories* of our noise with spatio-temporal correlations. Specifically let $\hat{E}^{(t)}$ denote the action of $E^{(t)}$ on the tth layer of the history code, and correspondingly let the action of an error history be denoted $\hat{E} := \bigotimes \hat{E}^{(t)}$. Given this action, the distribution of errors is simply given by the original error model, and so

$$\Pr(\hat{E}) = \prod_l \phi_l(\hat{E}). \tag{2.41}$$

As $\hat{E}$ can be interpreted as a purely spatial error in the history code, instead of an error history, this now constitutes a noise model with purely spatial correlations.

Gauge generators

Let $M_j^{(t)}$ denote the jth Pauli measurement occurring in the tth round of syndrome measurements. For error correction to remain well-defined, the final measurement round will need to consist of ideal

measurements of all the stabiliser generators[3]. As such if we let T be the number of measurement rounds, this assumption means that $\left\{M_j^{(T)}\right\}_j = \{S_k\}_k$. We will require the measurements within each round have disjoint support with the exception of the special final round.

Let $\hat{M}_j^{(t)}$ denote the action of $M_j^{(t)}$ on the corresponding layer of the history code. By construction, we have that the measurement statistics of these operators reproduce those of the measurements in the original code, as

$$\left[\hat{M}_j^{(t)}, \hat{E}\right] = \left[M_j^{(t)}, E^{(t)}\right] \tag{2.42}$$

for any j and t.

We now want to consider the history code itself, which takes the form of a subsystem code. To find the gauge generators, we can consider when two error histories are logically equivalent. As it is only the final accumulated error which determines whether or not a logical error has occurred, any two error histories are logically equivalent if and only if they possess the same syndrome, and do not differ by a non-trivial logical operator in the original code on the final time-slice. This implies that the gauge group is generated locally, and that these generators come in two forms: for $t < T$, the generators are given by the generators of the centraliser of $\hat{M}_j^{(t)}$ on the support of $\hat{M}_j^{(t)}$, and for $t = T$ they are just given by the measurement operators themselves $\{\hat{M}_j^{(T)}\}_j$.

2.5.2 Circuit-based model

We now turn our attention to a spatio-temporally correlated noise model which arises in the study of fault-tolerant error correction—circuit-based noise. In this noise model we explicitly take into account the circuits used to implement our syndrome measurements, and consider the faults within the constituent gates.

Measurement circuits

We will consider measurement circuits of the form shown in Figure 2.4: several ancillae are prepared in the state $|0\rangle$, a Clifford gate is applied to these and the qudits to be measured, and finally a Pauli measurement is performed on the ancillae. For the measurement circuit corresponding to $S_j^{(t)}$, we let $U_j^{(t)}$ denote the Clifford gate applied and $M_j^{(t)}$ the Pauli being measured. We denote the set of code qudits and ancillae involved in this syndrome measurement by

$$C_j^{(t)} := \operatorname{supp}\left(S_j^{(t)}\right), \tag{2.43a}$$

$$A_j^{(t)} := \operatorname{supp}\left(M_j^{(t)}\right), \tag{2.43b}$$

and let $R_j^{(t)} := C_j^{(t)} \cup A_j^{(t)}$ denote the full set of qudits involved, such that $\operatorname{supp}\left(U_j^{(t)}\right) = R_j^{(t)}$. The assumption that our measurements in each round are non-overlapping is equivalent $\left\{C_j^{(t)}\right\}_j$ being non-

[3]This is required to preclude errors occurring after measurements, which are clearly uncorrectable. This assumption is equivalent to only requiring that an error correction procedure correct errors prior to the final measurement round for it to be deemed successful.

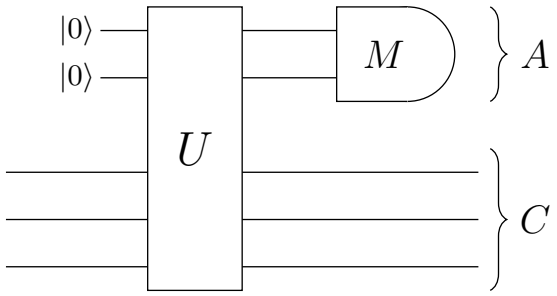


Figure 2.4: The form of measurement circuits we shall be considering. For each stabiliser generator S, several ancillae are prepared in the state $|0\rangle$, an entangling Clifford U is performed, and then a Pauli measurement M is performed on the ancillae. Sets C and A are used to denote the code and ancillary qubits involved in this measurement.

overlapping for each t. The layout of the history code corresponding to such a measurement procedure is shown in Figure 2.6.

As noted earlier, the final error correction round will correspond to an ideal measurement of all of the stabiliser generators. As such we will take $\{M_j^{(T)}\}_j = \{S_k\}_k$.

Noise model

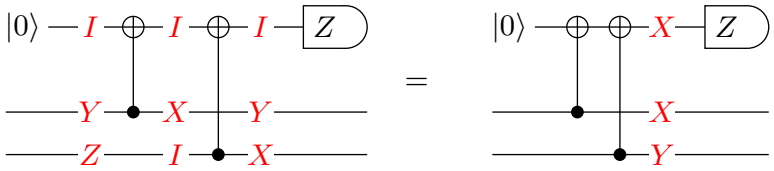


Figure 2.5: General Pauli errors being pushed through a measurement circuit. Notice that the Y error on the upper code qubit spreads onto the measurement qudit.

Consider now performing the above syndrome measurement circuits, subject to Pauli noise. For simplicity we will assume that the errors experienced within different measurement circuits are independent[4], but will allow arbitrary correlations within each circuit. Without loss of generality, we can push the error through the circuit, giving an effective error in $\mathcal{P}_{R_j^{(t)}}$ acting *after* the application of $U_j^{(t)}$ (see Figure 2.5). Since we have assumed that our circuits are Clifford, this updated Pauli error can be computed efficiently. Let $p_j^{(t)}$ be the distribution on $\mathcal{P}_{R_j^{(t)}}$ of these pushed-through errors.

There are two error sources which contribute to $E^{(t)}$: errors which accrued from previous syndrome measurement rounds, and those which occurred in the tth round itself. This means that our errors satisfy the recurrence relation

$$E^{(t)} = \epsilon^{(t)} \cdot U_t E^{(t-1)} U_t^\dagger, \tag{2.44}$$

[4]More generally any model which produces factored spatio-temporally correlated noise could be used.

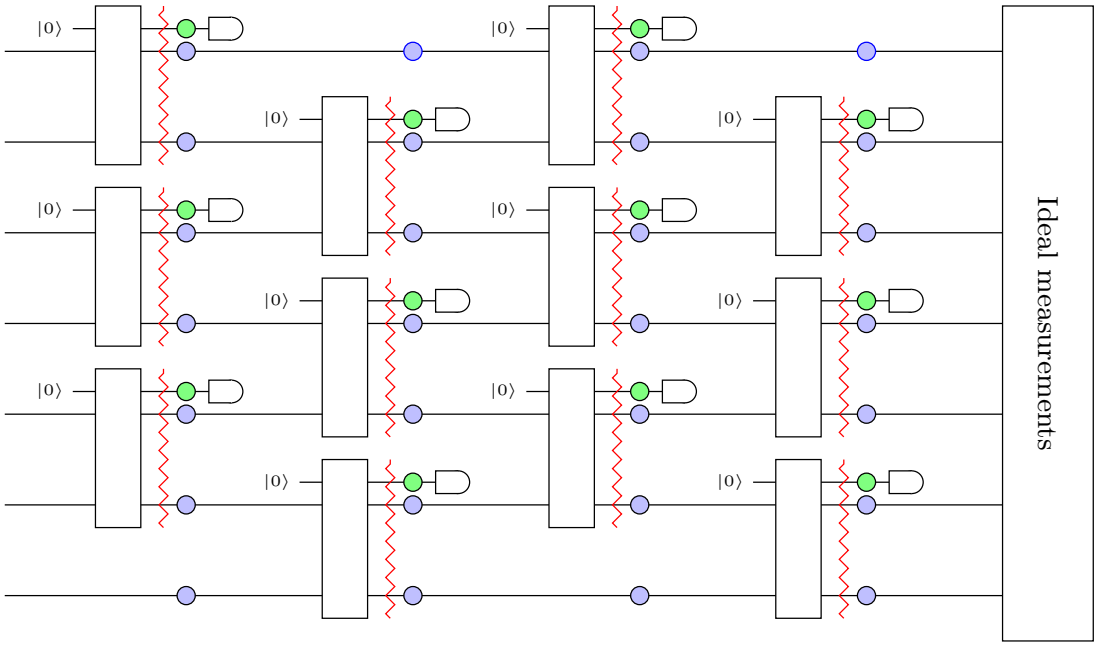


Figure 2.6: Laying out the qubits of the history code on a syndrome extraction circuit. The blue dots indicate code qubits contained in sets $\{C_j^{(t)}\}_{j,t}$, green dots indicate ancilla qubits contained in sets $\{A_j^{(t)}\}_{j,t}$, and red squiggles indicate the location of errors $\{\epsilon_j^{(t)}\}_{j,t}$.

where $\epsilon^{(t)}$ denotes the freshly introduced errors which occurred in the tth round, and $U^{(t)}$ to the unitary action of the syndrome measurement circuits in the tth round.

Decomposing this further into the individual measurement circuits, we have

$$U^{(t)} = \bigotimes_j U_j^{(t)} \quad \text{and} \quad \epsilon^{(t)} = \bigotimes_j \epsilon_j^{(t)}, \tag{2.45}$$

where $\epsilon_j^{(t)}$ and $U_j^{(t)}$ are supported solely on $R_j^{(t)}$, and $\epsilon_j^{(t)}$ is distributed according to $p_j^{(t)}$. As our code is assumed to be error-free prior to the beginning of the syndrome measurements, we have the initial condition $E^{(0)} = I$, and so this recurrence relation entirely describes the error model.

Factorising circuit noise

We start by noticing that in Equation (2.44) the errors at time t are entirely determined by the errors at time $t-1$, and the newly incurred errors $\epsilon^{(t)}$. As we have assumed that these new errors incurred in each round of syndrome measurements are independent, our error model is therefore Markovian, allowing us to factorise our noise model in the temporal direction

$$\Pr(E^{(1)}, \ldots, E^{(T)}) = \prod_t \Pr(E^{(t)} | E^{(t-1)}). \tag{2.46}$$

Next we can decompose this into the individual measurement circuits using Equation (2.45). Specifically we know that the individual new errors $\epsilon_j^{(t)}$ are independently distributed according to $p_j^{(t)}$, which

allows us to express the conditional probability as

$$\Pr\left(E^{(t)}\middle|E^{(t-1)}\right) = \prod_j p_j^{(t)}\left(\epsilon_j^{(t)}\right). \tag{2.47}$$

We now want to put this expression back purely in terms of the accumulated errors $E^{(t)}$, instead of the fresh errors $\epsilon^{(t)}$. Inverting the recurrence relation Equation (2.44), we have that

$$\epsilon^{(t)} = E^{(t)} \cdot U^{(t)} E^{(t-1)\dagger} U^{(t)\dagger}. \tag{2.48}$$

Recalling that $\epsilon^{(t)}$ and $U^{(t)}$ factor across $\{R_{j,t}\}_j$ (see Equation (2.45)), this implies that

$$\epsilon_j^{(t)} = \left[E^{(t)}\right]_{R_j^{(t)}} \cdot U_j^{(t)} \left[E^{(t-1)\dagger}\right]_{C_j^{(t)}} U_j^{(t)\dagger}, \tag{2.49}$$

where $[P]_R$ denotes the restriction of a Pauli P to a region R. Putting this together with Equations (2.46) and (2.47), we have that the distribution on error histories $\Pr(E^{(1)}, \ldots, E^{(T)})$ factors both in space and time, allowing us to apply the history code construction of Section 2.5.1.

2.6 Error correction as a statistical mechanical phase

The fundamental theorem of the stat mech mapping, Theorem 2.2, links the equilibrium thermodynamic properties of our stat mech model and the error correction properties of our code via the disordered partition function and the error class probabilities. In fact, a much stronger connection is true, as has already been noted by previous authors in the case of independent noise, beginning with Dennis *et al.* [2.37]. Previous work has shown that the regions in parameter space for independent noise in which a code can and cannot be decoded are phases in the associated stat mech system [2.37, 2.39, 2.42]. This implies that the error correction threshold of our code manifests as a phase transition in the associated stat mech model. Moreover, there is an explicit order parameter that captures this phase transition, although unfortunately it actually a "disorder" parameter in that it involves the amount of disorder in the stat mech model.

In this section, we review and extend this connection to the general case of correlated noise. We define appropriate notions of "above" and "below" the threshold and prove that the disorder parameter exhibits non-analytic behaviour in the sense that it converges or diverges if and only if the code is above or below the threshold, respectively. This shows a precise sense in which the phase boundary of the stat mech model is exactly the threshold of the corresponding code.

In Section 2.4 we saw how, using this correspondence, numerical techniques to analyse phase transition in stat mech systems can be used to give approximations of code thresholds. We also note that this correspondence also opens the door to applying analytic techniques for studying phase transitions, such as the duality method [2.39, 2.63, 2.64].

2.6.1 Maximum probability and maximum likelihood decoding

A decoder is an algorithm which takes as input the syndrome and attempts to estimate which error occurred. A natural starting point would be a decoder which simply outputs the most likely error in

this error model among those which are consistent with the observed syndrome. We will refer to this as the *maximum probability* (MP) decoder.

For degenerate codes, where several logically equivalent errors can have the same syndrome, the MP decoder is generally sub-optimal. This stems from the fact that successful decoding does not require the decoder to output precisely the error which occurred, but just an error which is logically equivalent. As such, the ideal decoding will not correspond to the single error with the highest probability, but the *error class* with the highest probability. The optimal decoder therefore outputs an error from this most likely class, and we will refer to this as the *maximum likelihood* (ML) decoder.

To see the sub-optimality of the maximum probability decoder, consider the following example. Suppose a syndrome s is measured which is consistent with three errors, E_1, E_2, E_3, which occur with probabilities

$$\Pr(E_1) = 4\%, \quad \Pr(E_2) = 3\%, \quad \Pr(E_3) = 3\%. \tag{2.50}$$

If E_2 and E_3 are logically equivalent, then the error *class* $\overline{E_2}$ is more likely than $\overline{E_1}$, even though E_1 itself is the single most likely error. In this case we can see that the MP decoder will be sub-optimal, and the conditional success probabilities for the two decoders are

$$\Pr\left(\text{MPD success}|s\right) = 40\%, \tag{2.51a}$$
$$\Pr\left(\text{MLD success}|s\right) = 60\%. \tag{2.51b}$$

For each syndrome s, let C_s denote an arbitrary Pauli with syndrome s. Similarly, for each logical l, let L_l denote a Pauli corresponding to logical l. The sets $\{\overline{C_s L_l}\}_{s,l}$ correspond to the logical error classes, and form a partition of the Paulis $\mathcal{P}^n$.

For a decoder to always return the code back to the code space, we require that when a syndrome s is input, it always returns an error with this same syndrome s. As such, without loss of generality, a decoder can be taken to be of the form $s \mapsto C_s L_{\delta(s)}$, where δ is a map from syndromes to logicals. In this notation, the maximum probability decoder takes the form

$$\delta_{\text{MP}}(s) := \arg\max_l \max_{E \in \overline{C_s L_l}} \Pr(E), \tag{2.52}$$

and the maximum likelihood decoder the form

$$\delta_{\text{ML}}(s) := \arg\max_l \Pr(\overline{C_s L_l}). \tag{2.53}$$

2.6.2 Minimum free energy and minimum energy decoding

Let H_E denote the stat mech model corresponding to our code and error model (see Sections 2.2, 2.3 and 2.5), which satisfies the Nishimori conditions at inverse temperature β_{N}.

Similar to the maximum probability and likelihood decoders defined above, consider the *β-minimum free energy* (β-MFE) decoder, which is given by minimising the free energy at inverse temperature β,

$$\delta_{\beta\text{-MFE}}(s) := \arg\min_l F_{C_s L_l}(\beta), \tag{2.54}$$

where $F_E(\beta) := -\frac{1}{\beta} \ln Z_E(\beta)$.

When satisfying the Nishimori conditions, $\beta = \beta_{\mathrm{N}}$, we can apply the fundamental theorem of the stat mech mapping (Theorem 2.2), to give

$$F_E(\beta_{\mathrm{N}}) = -\frac{1}{\beta_{\mathrm{N}}} \ln \Pr(\overline{E}). \tag{2.55}$$

As this is a monotonically decreasing function of the error class probability, this tells us that the β_{N}-MFE decoder is precisely the maximum likelihood decoder,

$$\delta_{\beta_{\mathrm{N}}\text{-MFE}} \equiv \delta_{\mathrm{ML}}. \tag{2.56}$$

Similarly, if we take $\beta \to \infty$, then the free energy reduces to the minimum energy,

$$\lim_{\beta\to\infty} F_E(\beta) = \min_{\vec{s}} H_E(\vec{s}). \tag{2.57}$$

Theorem 2.2 gives us that the energies of our stat mech model correspond to error probabilities, and so

$$\lim_{\beta\to\infty} F_E(\beta) = \min_{E'\in\overline{E}} -\frac{1}{\beta_{\mathrm{N}}} \ln \Pr(E'). \tag{2.58}$$

As this is a monotonically decreasing function of the error probabilities, this tells us that the ∞-MFE decoder—which one could refer to as the *minimum energy* decoder—is precisely the MP decoder,

$$\delta_{\infty\text{-MFE}} \equiv \delta_{\mathrm{MP}}. \tag{2.59}$$

This reduction to MFE decoding implies that any of the plentiful methods for approximating partition functions [2.65, 2.66], or free energy differences [2.67] can be used to implement approximate ML decoding. In Section 2.7 we will expand upon this connection, giving a tensor network algorithm which approximates ML decoding.

2.6.3 The error correction threshold as a phase transition

An important way of quantifying the resilience of a quantum error correction procedure to an error model is the *error threshold*. Specifically, consider a family of quantum codes, with a logical algebra of finite dimension K, and an error model which depends on a parameter $\theta \geq 0$. We define the notion of threshold as follows. A code family has a *threshold* if there exists a $\theta_t > 0$ such that the asymptotic success probability is maximal for $\theta < \theta_t$, and minimal for $\theta > \theta_t$, i.e.

$$\lim_{n\to\infty} \Pr(\text{Decoder success}) = \begin{cases} 1 & \text{if } \theta < \theta_t, \\ 1/K & \text{if } \theta > \theta_t. \end{cases} \tag{2.60}$$

These regimes we refer to as being *below threshold* and *above threshold* respectively.

Clearly this is a rather strong notion of threshold. One can imagine codes or noise models which possess an intermediate regime between being below or above threshold, where the code can be decoded better than random chance, but not perfectly. As an example, having asymmetric rates of independent X and Z noise in the surface code will generate such a gap. Therefore, not all codes necessarily possess

a threshold in this sense. However, the presence of a threshold in our sense is often taken as a desirable property in engineered quantum codes. We also note that for this section we have restricted our attention to codes with finite K. Some codes with growing K, such as finite-rate LDPCs, are also known exhibit such an intermediate regime [2.43, 2.68].

We also use the notion of threshold with respect to a particular choice of decoder, not just the optimal decoder. However, unless otherwise specified, we will be default to considering the threshold with respect to the optimal decoder, and refer to this as *the* threshold.

Given that the threshold corresponds to a dramatic jump in the success probability of decoding, one might naturally suspect that this is precipitated by a corresponding dramatic jump in the error class probabilities. As these probabilities equal the partition function of our stat mech models, this would also suggest a (disordered) phase transition. To see that this is indeed the case, we will consider the disorder parameter given by the free energy cost of a non-trivial logical operator L_m,

$$\Delta_m(E) := F_{EL_m}(\beta_{\mathrm{N}}) - F_E(\beta_{\mathrm{N}}) \tag{2.61a}$$

$$= -\frac{1}{\beta_{\mathrm{N}}} \log Z_{EL_m} + \frac{1}{\beta_{\mathrm{N}}} \log Z_E. \tag{2.61b}$$

We note that for topological codes which have string-like (or sheet-like in higher dimensions) logical operators, this corresponds to the free energy cost of a domain wall, as noted by Dennis *et al.* [2.37]. Let θ_p denote the point at which the stat mech model undergoes a phase transition corresponding to this parameter.

We now want to show that the threshold corresponds to a phase transition in our stat mech model, $\theta_t = \theta_p$. As a first step, we consider the quenched average, $\Delta_m := \langle \Delta_m(E) \rangle_E$. We would expect that $\Delta_m \to \infty$ below threshold and $\Delta_m \to 0$ above. The former limit was shown in Ref. [2.42], which implies that $\theta_p > \theta_t$. We present a simplified proof below.

Lemma 2.1 (Divergence in mean)**.** *If the code is below threshold $\theta < \theta_t$, then the quenched average free energy cost diverges,*

$$\lim_{n\to\infty} \Delta_m = \infty. \tag{2.62}$$

Proof. By recalling that $Z_E = \Pr(\overline{E})$, we can see that $\Delta_m(E)$ only depends on E up to logical equivalence. A convenient set of representatives from each error class are given by the Paulis $\{D_s L_l\}_{s,l}$, where s and l correspond to the syndrome and logical degrees of freedom, and $\{D_s\}_s$ are the decoding Paulis associated with the MLD. As shorthand, we will denote $\Pr(s,l) := \Pr(\overline{D_s L_l})$.

Using this, we can see that Δ_m can be expanded as a Kullback-Leibler divergence,

$$\Delta_m = \frac{1}{\beta_{\mathrm{N}}} \sum_{s,l} \Pr(s,l) \log \frac{\Pr(s,l)}{\Pr(s,l+m)}. \tag{2.63}$$

Applying the log sum inequality to the summation over syndromes, we get a bound in terms of the KL divergence of the marginal $\Pr(l)$,

$$\Delta_m \geq \frac{1}{\beta_{\mathrm{N}}} \sum_{l} \Pr(l) \log \frac{\Pr(l)}{\Pr(l+m)}. \tag{2.64}$$

We note that $\Pr(l=0)$ is the probability of the MLD succeeding. Since we are below threshold we have that $\Pr(l=0) \to 1$, and so $\Delta_m \to \infty$. □

Above the threshold one cannot simply consider the quenched average Δ_m. To see this, consider a code in which a single unlikely syndrome can be perfectly decoded with certainty, and all others provide a uniform logical error. The correctable syndrome gives that $\Delta_m = \infty$, despite the fact that the code is above threshold. As such, we cannot necessarily conclude that $\Delta_m \to 0$ above threshold.

Therefore, to provide a converse statement we must consider more than just the quenched average. Here we will formalise an argument first sketched in Ref. [2.39] that shows that the way to obtain a converse is to change our notion of convergence.

Suppose for the moment that $\Delta_m(E)$ concentrates around a value $\tilde{\Delta}_m$, in the sense that

$$\lim_{n\to\infty} \Pr\Big(|\Delta_m(E) - \tilde{\Delta}_m| \leq \epsilon\Big) = 1 \quad \forall \epsilon > 0. \tag{2.65}$$

A phase transition corresponds to a jump in this 'typical' value $\tilde{\Delta}_m$. Because $\Delta_m(E)$ is not bounded from above, we notice that this typical value $\tilde{\Delta}_m$ need not necessarily correspond to the mean Δ_m. Indeed, in the above counterexample, where $\Delta_m = \infty$ above threshold, we nonetheless see that $\tilde{\Delta}_m \to 0$ as expected. As such, we see that the mean is not necessarily the correct figure-of-merit. If instead we look to this typical value, then we can see that $\tilde{\Delta}_m \to \infty$ and $\tilde{\Delta}_m \to 0$ below *and* above threshold respectively.

Lemma 2.2 (Divergence/convergence in probability [2.39])**.** *The code is below threshold if and only if the free energy cost of every non-trivial logical m diverges in probability,*

$$\lim_{n\to\infty} \Pr\big(\Delta_m(E) \geq t\big) = 1 \quad \forall t, m \neq 0. \tag{2.66}$$

Similarly, the code is above threshold if and only if the free energy cost of every logical converges to zero in probability,

$$\lim_{n\to\infty} \Pr\big(\,|\Delta_m(E)| \leq \epsilon\,\big) = 1 \quad \forall \epsilon > 0, m. \tag{2.67}$$

Proof. The success probability of the MLD corresponds to $\Pr(l = 0)$. We first note that $\Delta_m(E)$ can be written in terms of conditional probabilities,

$$\Delta_m(E) = \frac{1}{\beta_{\mathrm{N}}} \log \frac{\Pr(l|s)}{\Pr(l+m|s)}. \tag{2.68}$$

Next we recall that the maximum likelihood condition $\Pr(l = 0|s) \geq \Pr(m|s)$ holds for all s and m, and as such $\Pr(l = 0) \geq 1/d^K$. The code is below threshold if the probability of successful decoding approaches 1—this is equivalent to the probability of successful decoding *conditioned upon the syndrome* approaching 1 almost surely,

$$\lim_{n\to\infty} \Pr\Big[\Pr(l = 0|s) \geq 1 - \gamma\Big] = 1 \quad \forall \gamma > 0. \tag{2.69}$$

To be clear, in this expression the inner probability is a probability over logicals l and the outer probability is over syndromes. Thus, the below threshold condition says informally that this distribution peaks at one in the limit. This can be seen to be equivalent to being in the ordered phase,

$$\lim_{n\to\infty} \Pr\big(\Delta_m(E) \geq t\big) = 1 \quad \forall t, m \neq 0. \tag{2.70}$$

The forward and reverse directions are given by taking the choices of parameter

$$t := \frac{1}{\beta_{\mathrm{N}}} \log \frac{1-\gamma}{\gamma}, \tag{2.71a}$$

$$\gamma := 1 - \frac{1}{1 + e^{-\beta_{\mathrm{N}} t}(K-1)}, \tag{2.71b}$$

respectively.

In a similar vein, being above threshold implies that the conditional probability of decoding almost surely approaches $1/K$,

$$\lim_{n\to\infty} \Pr\Big[\Pr(l = 0|s) \leq 1/K + \eta\Big] = 1 \quad \forall \eta > 0. \tag{2.72}$$

This is equivalent to being in the disordered phase,

$$\lim_{n\to\infty} \Pr\big(|\Delta_m(E)| \leq \epsilon\big) = 1 \quad \forall \epsilon > 0. \tag{2.73}$$

The forward and reverse directions here follow from the choices of parameter

$$\epsilon := \frac{1}{\beta_{\mathrm{N}}} \log \frac{1+\eta K}{1-\eta K(K-1)}, \tag{2.74a}$$

$$\eta := \frac{1}{1 + (K-1)e^{-\beta_{\mathrm{N}}\epsilon}} - \frac{1}{K}, \tag{2.74b}$$

respectively. □

We add more remark about the notions of above and below threshold that we have adopted here. As noted in the case of the surface code with independent X and Z noise with different strengths, there can be a gap between the cases of above and below threshold. In that case our theorem does not apply. However, it is still the case that the steps of our proof could be followed to establish that there are a sequence of phase transitions in the disorder parameters Δ_m in the case where there are potentially differing thresholds for each logical subalgebras.

2.6.4 Reentrance

Above we have shown that, along the Nishimori line, there exists a phase transition which corresponds to the error correction threshold of the optimal decoder. By considering non-optimal decoders, we can extend this to show the equivalence between phase transitions and thresholds away from the Nishimori line.

By way of example, we consider the β-MFE decoders. We could now consider the phase diagram of our stat mech system in (θ, T)-space, where θ was the parameter of our noise model, and $T = 1/\beta$ is the temperature.

Lemma 2.3 (Reentrance of the ordered phase)**.** *For any β such that the β-MFE decoder has a threshold, this threshold $\theta_t(\beta)$ is always less than or equal to that seen on the Nishimori line,*

$$\theta_t(\beta) \leq \theta_t(\beta_N). \tag{2.75}$$

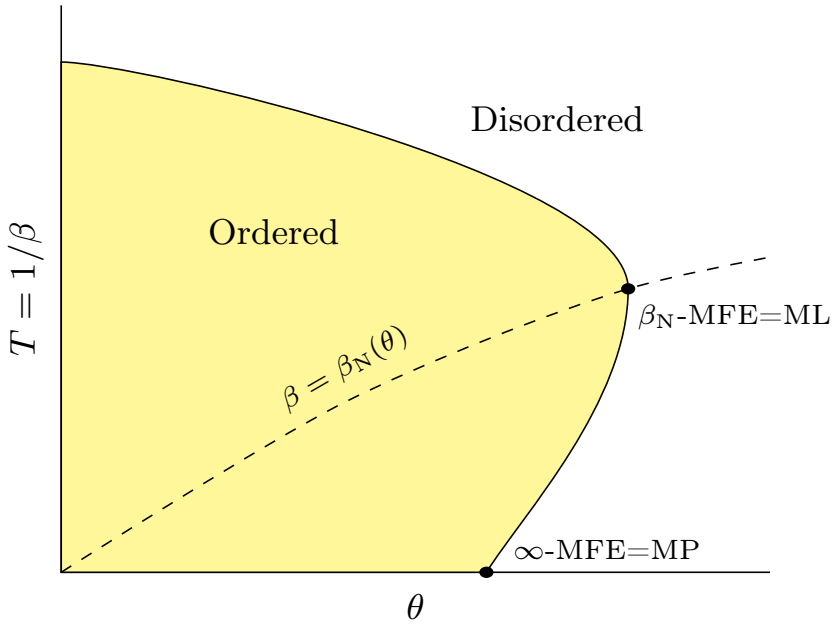


Figure 2.7: A sketch of the generic phase diagram for the stat mech model. The solid line indicates the phase boundary, and the shaded region the phase in which error correction is possible. The dashed line indicates the Nishimori condition, along which the MFE decoder corresponds to the ML decoder. The thresholds of the MP and ML decoder are both indicated by black dots. An explicit example, for the case of bit-flip noise in the toric code, is presented in Figure 2.3.

Proof. As the β-MFE decoder is directly defined in terms of minimising the free energy, Lemmas 2.1 and 2.2 both naturally extend to show that the stat mech model has a phase transition at inverse temperature β and $\theta = \theta_t(\beta)$.

As the ML decoder is optimal, it necessarily has the highest threshold of any decoder. Recalling that the ML decoder is precisely the β_N-MFE decoder, where β_N is the inverse Nishimori temperature, this implies that $\theta_t(\beta) \leq \theta_t(\beta_N)$ for all β. □

2.7 Tensor network maximum likelihood decoding

In Section 2.6 we showed that the problem of maximum likelihood decoding can be reduced to calculating partition functions/free energy differences. Specifically it reduces to calculating

$$\delta_{\mathrm{ML}}(s) := \arg\max_l Z_{C_s L_l}(\beta_N), \tag{2.76}$$

where β_N is the inverse Nishimori temperature. This relationship was first discussed in Ref. [2.37].

The problem of calculating the partition function of a stat mech model can be naturally expressed as the contraction of a tensor network [2.69, 2.70], as we will review below. This suggests a general method for approximating the ML decoder by finding a tractable (possibly approximate) contraction sequence for the tensor network associated to the partition function.

The idea of a such a *tensor network decoder* is not new. In Ref. [2.51], Bravyi, Suchara and Vargo (BSV) consider an explicit tensor network for the ML decoder for the surface code with independent Pauli noise, and numerically study the threshold for some parameter choices. It is clear from their paper that this method generalises to other codes with independent Pauli noise models [2.71], though the specific methods they use for (approximately) contracting the tensor network would need to be modified if the code were not a planar code. Other authors have also considered tensor network decoders which can account for non-Pauli noise [2.36, 2.72].

In fact, the tensor network considered by BSV [2.51] is exactly the natural tensor network that one obtains from the partition function using the mapping of Refs. [2.69, 2.70]. Thus, the tensor network decoders that we define below are generalisations of the BSV decoder that, by virtue of Theorem 2.2, apply even in the case of correlated noise models on arbitrary codes. If the code has low-weight stabiliser (or gauge) generators, then storing these tensors is efficient. Contracting a general tensor network is unfortunately #P-hard [2.73,2.74], however, we will discuss efficient approximate contraction strategies that could be employed in the interesting cases of spatially local codes.

2.7.1 Tensor network algorithm for approximate maximum likelihood decoding

To see how calculating the partition function of a stat mech model can be naturally expressed as the contraction of a tensor network, we begin with an arbitrary Hamiltonian

$$h(\vec{s}) = \sum_j h_j(\vec{s}), \tag{2.77}$$

with a corresponding partition function

$$Z = \sum_s e^{-\beta h(\vec{s})} = \sum_s \prod_j e^{-\beta h_j(\vec{s})}. \tag{2.78}$$

This expression almost takes the form of a tensor contraction. To make this into a tensor network, consider indices $\{\alpha_{i,j}\}_{i,j}$, labelled by both spins and interactions. We want to require that $\alpha_{i,j} = s_i$ for all j, which we can impose by including Kronecker delta factors. In terms of these indices, the partition function can be expressed as

$$Z = \sum_\alpha \left[\prod_i \delta\left(\{\alpha_{i,j}\}_j\right) \prod_j e^{-\beta h_j(\{\alpha_{i,j}\}_i)} \right]. \tag{2.79}$$

The virtue of this expression is that each $\{\alpha_{i,j}\}_{i,j}$ now occurs precisely twice in each term above, and as such the above expression forms a tensor network. Obviously, for local Hamiltonians in which h_j only depends non-trivially on a finite set of sites i, we can drop any index $\alpha_{i,j}$ for which interaction h_j is independent of s_i. After doing so, the tensor network simplifies to a network which has the same connectivity as the underlying Hamiltonian. Specifically, there is a tensor corresponding to each site (δ) and each bond (Gibbs weight), with the natural connectivity. In Figure 2.8 we show the corresponding diagram for the surface code under iid noise, which reduces to that found in Ref. [2.51].

Applying this construction to the stat mech mappings of Sections 2.2, 2.3 and 2.5, we see that contracting this tensor network gives an algorithm to compute the ML decoding. Because contracting

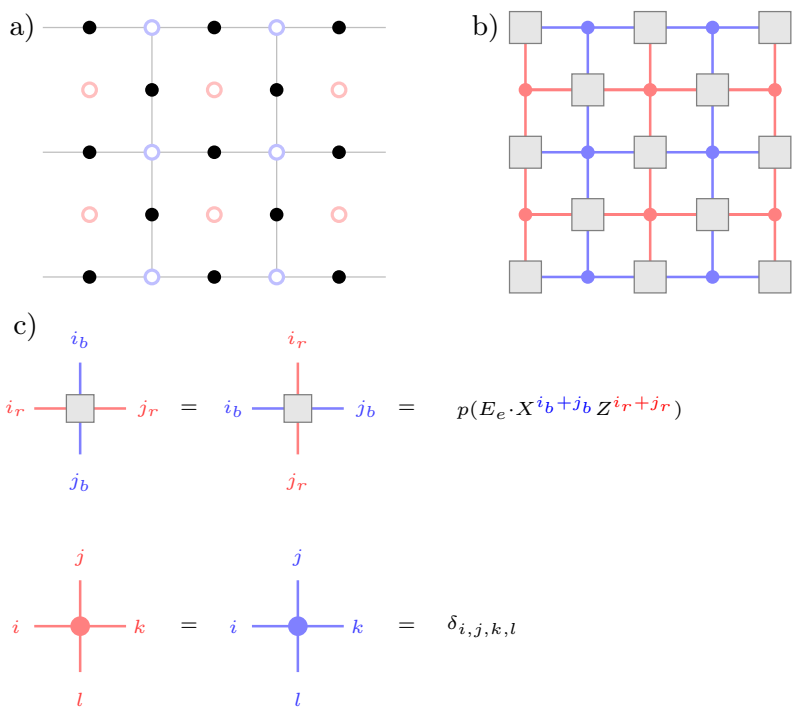


Figure 2.8: Tensor network ML decoder of the surface code under iid noise. a) surface code and the associated statistical mechanical model. Black circles indicate qubits, and the red/blue open circles the Z/X-type spins. See Figure 2.1 for the interactions within this model. b) Tensor network for the probability of an error class $\Pr(\overline{E})$, or equivalently for the partition function Z_E, for iid noise distributed according to $p : \mathcal{P} \to \mathbb{R}$. We note that this matches Fig. 6 of Ref. [2.51]. c) Values of the two tensors involved. Red/blue dots are δ tensors, corresponding to the Z/X-type spins, and grey tensors correspond to Gibbs weights. Once again this matches the tensors in Equations. (39–41) of Ref. [2.51].

tensor networks is generally hard, we will also need *approximate* tensor contraction schemes that provide parameterised approximations.

In the case of 2D topological codes under spatially correlated noise, three approximate contraction schemes suggest themselves:

- MPS-MPO contraction [2.75, 2.76]. For the toric code, this method exactly reproduces the MPS decoder of Ref. [2.51] (see Figure 2.8).
- Transfer matrix [2.77] and corner transfer matrix methods [2.78, 2.79].
- Methods that involve renormalisation on the virtual level, e.g. tensor network renormalisation (TNR) [2.80, 2.81] or (higher-order) tensor renormalisation group (HOTRG) [2.82, 2.83]. These examples remain efficient when extended to more than two dimensions.

Each of these methods contain an approximation parameter in the form of the bond dimension. For a fixed bond dimension, these methods all provide polynomial time approximations to the ML decoder, though we do not know of any general results which control the approximation as a function of the

bond dimension in any nontrivial way. Actually implementing these decoders (beyond what was done in Refs. [2.51, 2.58]) or finding provable guarantees on the approximation to the ML decoder both remain open problems.

2.8 Conclusions

In this paper we have provided a broad extension of the stat mech mapping of Ref. [2.37] to consider arbitrary stabiliser and subsystem codes, subject to correlated noise, including circuit-based noise. This class of noise models encompasses noise where distant spins are only conditionally independent, and allows for much more realistic noise modelling. As an application, we applied Monte Carlo simulation to this construction to show how positive correlations can push down the bit-flip threshold of the toric code. Finally, we showed how the stat mech mapping gives a natural family of efficient tensor network algorithms which approximate maximum likelihood decoding, generalising the decoder of Ref. [2.51].

There are several natural avenues for further inquiry. For example, it is clear that our method should apply to the setting of continuous variables quantum codes such as GKP codes [2.84], however there are analytical issues that must be addressed owing to the infinite dimensional Hilbert space. It would be interesting to understand these conditions in detail as these codes continue to gain experimental relevance [2.85].

Even more interesting would be a formalism for deriving stat mech models that can handle non-Pauli errors such as amplitude damping or coherent errors, or for quantum codes that are outside the stabiliser code formalism such as commuting projector codes. Since the errors in the most general such models do not have an abelian action on the codewords, this raises the possibility that one would need a *quantum* stat mech model to accurately capture the threshold in these cases. An interesting test case would be to extend our work to the semion code model of Ref. [2.86], since the anyons in the double semion model are abelian, but the check operators are not simple products of Paulis.

Acknowledgements

We thank Michael Beverland, Jacob Bridgeman, Benjamin Brown, Marius Junge and Thomas Smith for helpful discussions. Special thanks go to Robin Harper for assistance with the numerics. The authors acknowledge the University of Sydney HPC service at The University of Sydney for providing HPC resources that have contributed to the research results reported within this paper. This work was initiated during the thematic semester on Analysis in Quantum Information Theory at the Institut Henri Poincaré, and the authors thank the Institut and the organisers for their hospitality during that time.

This research was supported by the Australian Research Council Centre of Excellence for Engineered Quantum Systems (Project ID CE170100009). CTC also acknowledges support Sydney Nano Postgraduate Scholarship (John Makepeace Bennett Gift).

2.A Graphical models

An important family of correlated probability distributions is that of *probabilistic graphical models* [2.87]. In this appendix we briefly review the relationship between certain classes of graphical

models and the factored distributions we consider in Section 2.3. Specifically we will see that both undirected graphical models (Markov random fields) and directed acyclic graphical models (Bayesian networks) are both special examples of factored distributions. As such, the statistical mechanical mapping given in Section 2.3 also naturally extends to noise models of these forms as well.

2.A.1 Markov random field

We start by considering undirected graphical models, known as Markov random fields.

Definition 2.6 (Markov random field)**.** A distribution $\Pr(E) = \Pr(E_1, \ldots, E_n)$ is a *Markov random field* (MRF) with respect to an undirected graph G, if it satisfies the local Markov condition

$$\Pr(E_i | E_{G\setminus i}) = \Pr(E_i | E_{\partial i}), \tag{2.80}$$

for all i, where ∂i denotes the neighbours of i within G.

For such a graphical model, Hammersley-Clifford theorem tells us that it can be expressed as a factored distribution.

Lemma 2.4 (Hammersley-Clifford theorem)**.** *A strictly positive distribution which is a MRF with respect to G, can be factored over the set of all maximal cliques of G.*

Proof sketch. For any subset $Y \subseteq G$, let

$$\phi_Y(E_Y) := \prod_{X : X \subseteq Y} \Pr(E_X \otimes I_{G\setminus X})^{(-1)^{|Y\setminus X|}}, \tag{2.81}$$

where I was chosen as an arbitrary, but fixed, reference value of E. Noting that $\mu(X, Y) = (-1)^{|Y\setminus X|}$ forms a Möbius function, we can apply Möbius inversion theorem to give

$$\Pr\left(E_Z \otimes I_{G\setminus Z}\right) = \prod_{Y \subseteq Z} \phi_Y(E_Y), \tag{2.82}$$

and more specifically

$$\Pr\left(E\right) = \prod_{Y \subseteq G} \phi_Y(E_Y). \tag{2.83}$$

Applying the local Markov property, we find that $\phi_Y \equiv 1$ for any Y which is *not* a clique. As such we find that

$$\Pr(E) = \prod_j \phi_{C_j}(E_{C_j}), \tag{2.84}$$

where $\{C_j\}_j$ are the set of *all* cliques in G. By grouping together factors, this can straightforwardly reduced to a factorisation over just *maximal* cliques. □

2.A.2 Bayesian Networks

We now consider directed cyclic graphical models, known as Bayesian Networks.

Definition 2.7 (Bayesian network)**.** A distribution $\Pr(E) = \Pr(E_1, \ldots, E_n)$ is a *Bayesian network* (BN) with respect to an directed acyclic graph G, if it satisfies the local Markov condition

$$\Pr(E_i|E_{G\setminus\partial^- i}) = \Pr(E_i|E_{\partial^+ i}), \tag{2.85}$$

for all i, where $\partial^- i$ and $\partial^+ i$ denote the descendants and parents of i within G respectively.

As with Markov random fields, we see that Bayesian networks also naturally factorise.

Lemma 2.5 (Factorisation of Bayesian Networks)**.** *A distribution* $\Pr(\cdot)$ *which is a Bayesian network with respect to a directed acyclic graph* G*, can be factored over* $\{\{i\} \cup \partial^+ i\}_i$*. Specifically,*

$$\Pr(E) = \prod_{i=1}^{n} \Pr(E_i|E_{\partial^+ i}). \tag{2.86}$$

Proof. As our directed graph G is acyclic, it possesses a topological order. Consider labelling our indices according to that ordering, such that $\partial^+ i > i > \partial^- i$ element-wise. Next we consider expanding our joint distribution using the chain rule of conditional probabilities,

$$\Pr(E_1, \ldots, E_n) = \prod_{i=1}^{n} \Pr(E_i|E_{i+1}, \ldots, E_n). \tag{2.87}$$

Applying the local Markov condition, we get the desired factorisation,

$$\Pr(E) = \prod_{i=1}^{n} \Pr(E_i|E_{\partial^+ i}). \tag{2.88}$$

□

2.B Numerical simulation details

In this appendix we cover the details of how the statistical mechanical simulations of Section 2.4 were performed, specifically how the data presented in Figure 2.3 was collected.

The simulations we present follow closely the techniques used in Ref. [2.39] to study the toric code under a depolarising noise model.

2.B.1 Order parameter

As discussed in Section 2.4, the system we are studying corresponds to a random-bond Ising model, with an addition four-body coupling. In the case of zero disorder, $p = 0$, this system simply limits to the standard square-lattice Ising model. Noticing this, a convenient order parameter for this system is simply given by the average magnetisation,

$$m := \frac{1}{L^2} \sum_i s_i, \tag{2.89}$$

where L is the linear size of our system, such that we have L^2 sites.

2.B.2 Finite-size scaling

For very large system sizes, we could sample the magnetisation directly, and find the phase transition by detecting when the system spontaneously magnetises. This naïve method would however require prohibitively large system sizes, especially for large disorder. Instead we will utilise this order parameter indirectly, by considering the finite-size scaling of the corresponding correlation length.

Following Ref. [2.39], we define the wave-vector-dependent magnetic susceptibility

$$\chi(\vec{k}) := \frac{1}{L^2}\left\langle \left| \sum_i s_i e^{i\vec{k}\cdot\vec{r_i}} \right|^2 \right\rangle, \tag{2.90}$$

where $\vec{r_i}$ denotes the spatial position of site i, and $\langle\cdot\rangle$ denotes the thermal and disorder average,

$$\langle X \rangle := \sum_E \Pr(E) \cdot \sum_s \frac{e^{-\beta H_E(\vec{s})}}{Z_E} X(\vec{s}). \tag{2.91}$$

Using this, we can now define the two-point finite-size correlation function

$$\xi := \frac{1}{2\sin(k_{\min}/2)} \sqrt{\frac{\chi(\vec{0})}{\chi(\vec{k}_{\min})} - 1}, \tag{2.92}$$

where $\vec{k}_{\min} = (2\pi/L, 0)$ is the minimal non-zero wave-vector. Near the phase transition this correlation length is expected to have the finite-size scaling [2.88]

$$\xi/L \approx f\left[L^{1/\nu}(T - T_c)\right], \tag{2.93}$$

where f is a dimensionless scaling function.

At the critical temperature $T = T_c$, the normalised correlation length ξ/L becomes independent of temperature. We will determine the temperature of the phase transition by plotting ξ/L as a function of T for several different system sizes L, and fitting to Equation (2.93). If these curves do not cross, then this will be taken as indication that no phase transition is present.

2.B.3 Simulation parameters

We simulate our system for system sizes of $L \in \{12, 14, 18, 24, 28, 36\}$. Equilibration of our system is found by logarithmically binning the correlation length, and requiring that three successive bins agree to within error bars. To save time, the equilibration testing was only done on $L = 14, 24, 36$, and these times were reused for $L = 12, 18, 28$. The times and temperature ranges are given in Section 2.B.3.

For $L = 12, 14$ the correlation length is recorded for 5000 disorder samples, 1000 samples are taken for $L = 18, 24$, and 500 for $L = 28, 36$. The temperature ranges are swept over with a resolution of $\Delta T = 0.025$. Error bars are generated for all the data is given by the 95% confidence intervals given by bootstrapping, using 10,000 resamples. An example crossing plot is given for $p = 6\%$ in Figure 2.9. By performing this analysis for each p, we can estimate the phase boundary, as seen in Figure 2.3. All the simulations took a total of 1.1×10^6 CPU-hours, or 120 CPU-years.

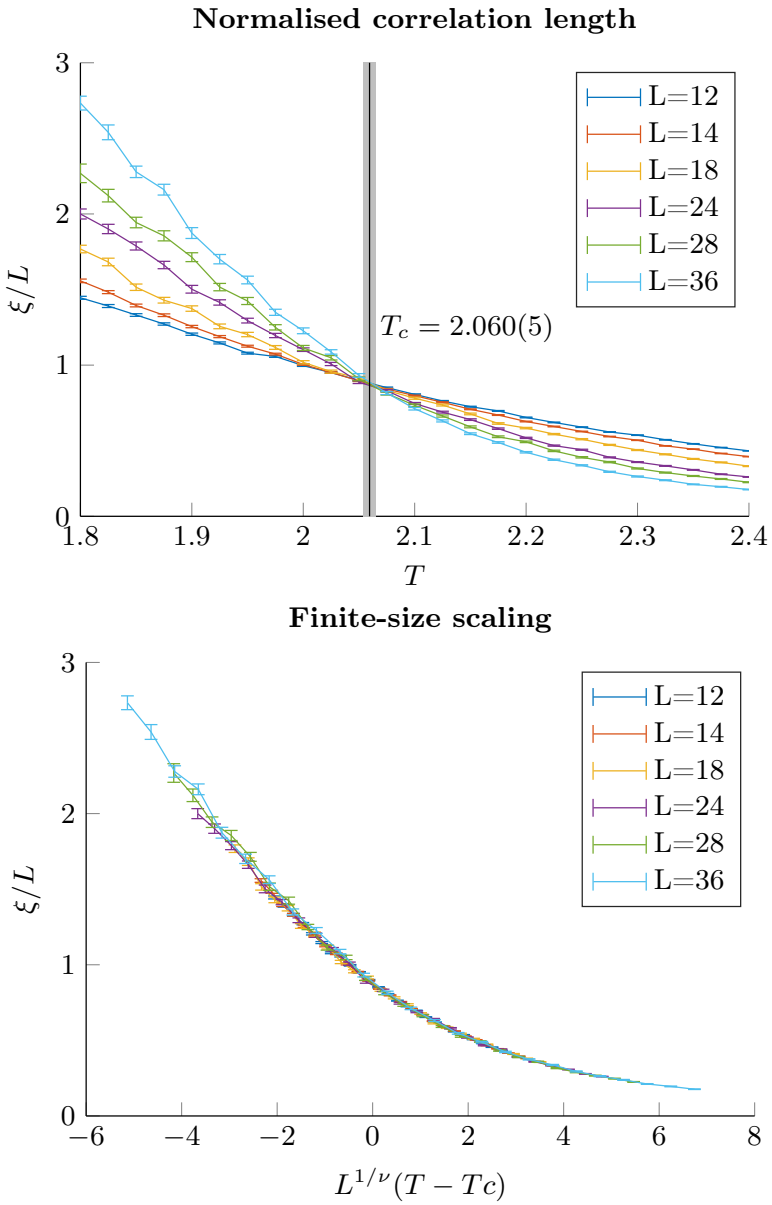


Figure 2.9: Crossing diagrams for $p = 6\%$. a) Normalised correlation length versus temperature, for several system sizes, with critical temperature indicated. b) Finite-size scaling of normalised correlation length. All error bars indicated are 95% confidence intervals given by bootstrapping.

p (%)	T	τ_{14}	τ_{24}	τ_{36}
2.00	2.00 – 2.50	14	16	18
4.00	2.00 – 2.50	14	16	18
6.00	1.80 – 2.40	14	16	18
8.00	1.60 – 2.30	16	18	22
8.50	1.50 – 2.10	17	21	23
9.00 – 9.10	1.40 – 2.00	18	22	24
9.20 – 9.50	1.30 – 2.00	18	22	25
9.55 – 10.20	1.25 – 2.00	19	23	26
10.30 – 10.50	1.25 – 2.00	20	24	27

Table 2.1: Temperature ranges and equilibration times. An equilibration time of τ corresponds to 2^τ Metropolis-Hastings steps (a number of single spin updates equal to the system size) before the correlation lengths recorded in the three last logarithmic bins are statistically consistent.

2.C Overlapping independent errors

In this appendix we will consider a correlated noise model which, at least superficially, appears to have a similar structure to the factored noise considered in Section 2.3. We will show that a statistical mechanical mapping for these noise models can be given by considering an appropriately enlarged code, and applying the mapping given for independent noise in Section 2.2.

Consider a model which involves independent noise processes acting on overlapping regions, such that the overall noise is given by the product of these constituent errors. Specifically we have overlapping regions $\{R_j\}$, each subject to independent noise given by distributions $\{p_j : \mathcal{P}_{R_j} \to \mathbb{R}^+\}$. For each error $E_j \sim p_j$, the overall error is simply $E = \prod_j E_j$. When these regions are non-overlapping each E_j simply corresponds to the restriction $E|_{R_j}$, but for overlapping regions multiple constituent errors can give rise to the same overall error, e.g. both pairs of errors $\{XYI, IIZ\}$ and $\{XII, IYZ\}$ give the overall error $E = XYZ$. Due to this redundancy, these noise models take the more complicated form

$$\Pr(E) = \sum_{\{E_j\}:\prod_j E_j = E} \prod_j p_j(E_j). \tag{2.94}$$

The approach here is similar to a stripped-down version of the history code construction considered in Section 2.5. The idea here is to consider a code now acting on multiple copies of the original lattice, which we will refer to as the *enlarged code*. This added redundancy will allow previously overlapping errors to act on disjoint copies, giving a truly independent noise model. This added redundancy will in turn be compensated for by introducing gauge generators between the layers.

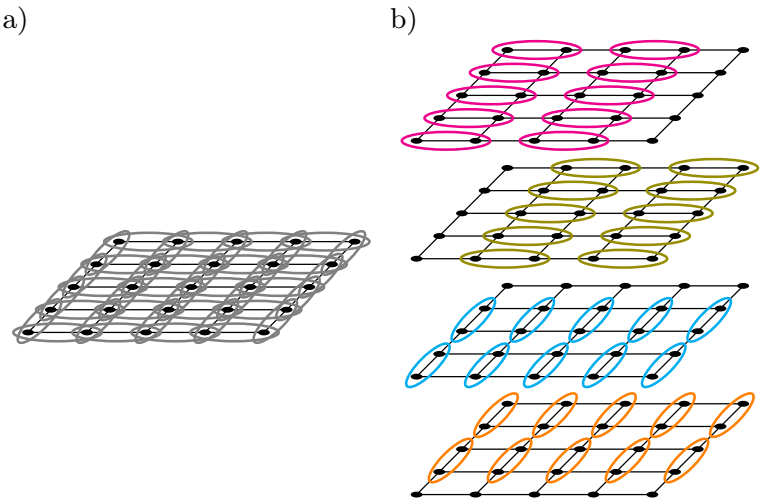


Figure 2.10: Overlapping-independent errors effecting nearest neighbour pairs on a square lattice, such as the 'NN-depolarizing' model considered in Ref. [2.89]. a) Overlapping regions R_j on the original code. b) Non-overlapping regions $\tilde{R}_j$ on $C = 4$ copies of the original code which compose the enlarged code.

Enlarged code

Specifically, consider C copies of our code, such that each independent noise process can act disjointly on precisely one copy. Specifically we will consider assigning each R_j to one of these copies; we will denote by $\{\tilde{R}_j\}_j$ these now-non-overlapping regions in the enlarged code (see Figure 2.10). If we denote the error on each copy c of our code by $E^{(c)}$, then the corresponding error on the original and enlarged codes are given by $E = \prod_c E^{(c)}$ and $\tilde{E} := \otimes_c E^{(c)}$ respectively. The enlarged error $\tilde{E}$ can be considered a specific manifestation of the overall error E.

Due to the non-overlapping nature of $\{\tilde{R}_j\}_j$, the enlarged errors $\tilde{E}$ are independently distributed,

$$\Pr(\tilde{E}) = \prod_j p_j \left(\tilde{E}\Big|_{\tilde{R}_j} \right). \tag{2.95}$$

This added redundancy means that the relationship between the errors in the original code E and the enlarged code $\tilde{E}$ is, as noted earlier, one-to-many. To account for this, we need to introduce gauge generators acting between copies,

$$\{\tilde{G}_l\}_l := \{X_{i,c} \otimes X_{i,c+1}^{-1}, Z_{i,c} \otimes Z_{i,c+1}^{-1}\}_{i,c}, \tag{2.96}$$

where $P_{i,c}$ denotes Pauli P acting on the ith site in the cth copy of our code. By construction, the set of all enlarged errors $\tilde{E}$ which correspond to the same overall error E is given by cosets of $\langle G_l \rangle_l$.

Next we define our stabilisers in the enlarged code, $\tilde{S}_k$, to simply be the original stabilisers S_k acting on the the first copy of our code,

$$\tilde{S}_k := S_k \otimes I \otimes \cdots \otimes I. \tag{2.97}$$

We now define the corresponding stat mech mapping analogous to that in Section 2.2, as

$$H_{\tilde{E}}(\vec{c},\vec{g}) = -\sum_{j,\sigma\in\mathcal{P}_{\tilde{R}_j}} J_j(\sigma)\left[\sigma,\tilde{E}\prod_k \tilde{S}_k^{c_k}\prod_l \tilde{G}_l^{g_l}\right], \tag{2.98}$$

with corresponding Nishimori condition

$$\beta J_j(\sigma) = \frac{1}{|\mathcal{P}_{\tilde{R}_j}|}\sum_{\tau\in\mathcal{P}_{\tilde{R}_j}} \log p_j(\tau)\left[\sigma,\tau^{-1}\right]. \tag{2.99}$$

Recalling that the gauge generators $\{G_l\}_l$ map between different manifestations $\tilde{E}$ of the same overall error E, and $\{\tilde{S}_k\}_k$ between stabiliser-equivalent errors, we have that

$$e^{-\beta H_{\tilde{E}}(\vec{0},\vec{0})} = \Pr(\tilde{E}), \tag{2.100}$$

$$\sum_{\vec{g}} e^{-\beta H_{\tilde{E}}(\vec{0},\vec{g})} = \Pr(E), \tag{2.101}$$

$$Z_{\tilde{E}} := \sum_{\vec{c},\vec{g}} e^{-\beta H_{\tilde{E}}(\vec{c},\vec{g})} = \Pr(\overline{E}), \tag{2.102}$$

in analogy to Theorem 2.1.

Bibliography

2.1 P. W. Shor, "Scheme for reducing decoherence in quantum computer memory," *Physical Review A*, **52**, p. R2493–R2496, (1995).

2.2 A. M. Steane, "Error Correcting Codes in Quantum Theory," *Physical Review Letters*, **77**, p. 793–797, (1996).

2.3 A. G. Fowler and J. M. Martinis, "Quantifying the effects of local many-qubit errors and nonlocal two-qubit errors on the surface code," *Physical Review A*, **89**, p. 032316, arXiv:1401.2466, (2014).

2.4 E. Novais, A. J. Stanforth, and E. R. Mucciolo, "Surface code fidelity at finite temperatures," *Physical Review A*, **95**, p. 042339, arXiv:1606.03163, (2017).

2.5 B. J. Brown, D. Loss, J. K. Pachos, C. N. Self, and J. R. Wootton, "Quantum memories at finite temperature," *Reviews of Modern Physics*, **88**, p. 045005, arXiv:1411.6643, (2016).

2.6 R. Alicki, M. Fannes, and M. Horodecki, "On thermalization in Kitaev's 2D model," *Journal of Physics A: Mathematical and Theoretical*, **42**, p. 065303, arXiv:0810.4584, (2009).

2.7 S. Chesi, B. Röthlisberger, and D. Loss, "Self-correcting quantum memory in a thermal environment," *Physical Review A*, **82**, p. 022305, arXiv:0908.4264, (2010).

2.8 E. Novais and E. R. Mucciolo, "Surface Code Threshold in the Presence of Correlated Errors," *Physical Review Letters*, **110**, p. 010502, arXiv:1209.2157, (2013).

2.9 P. Jouzdani, E. Novais, and E. R. Mucciolo, "Fidelity of the surface code in the presence of a bosonic bath," *Physical Review A*, **88**, p. 012336, arXiv:1304.2975, (2013).

2.10 C. D. Freeman, C. M. Herdman, D. J. Gorman, and K. B. Whaley, "Relaxation dynamics of the toric code in contact with a thermal reservoir: Finite-size scaling in a low-temperature regime," *Physical Review B*, **90**, p. 134302, arXiv:1405.2315, (2014).

2.11 D. A. López-Delgado, E. Novais, E. R. Mucciolo, and A. O. Caldeira, "Long-time efficacy of the surface code in the presence of a super-Ohmic environment," *Physical Review A*, **95**, p. 062328, arXiv:1612.06460, (2017).

2.12 H. K. Ng and J. Preskill, "Fault-tolerant quantum computation versus Gaussian noise," *Physical Review A*, **79**, p. 032318, arXiv:0810.4953, (2009).

2.13 D. P. S. McCutcheon, N. H. Lindner, and T. Rudolph, "Error Distributions on Large Entangled States with Non-Markovian Dynamics," *Physical Review Letters*, **113**, p. 260503, arXiv:1403.4956, (2014).

2.14 R. Raussendorf, J. Harrington, and K. Goyal, "A fault-tolerant one-way quantum computer," *Annals of Physics*, **321**, p. 2242–2270, arXiv:quant-ph/0510135, (2006).

2.15 A. G. Fowler, A. M. Stephens, and P. Groszkowski, "High-threshold universal quantum computation on the surface code," *Physical Review A*, **80**, p. 052312, arXiv:0803.0272, (2009).

2.16 N. H. Nickerson, Y. Li, and S. C. Benjamin, "Topological quantum computing with a very noisy network and local error rates approaching one percent," *Nature Communications*, **4**, p. 1756, arXiv:1211.2217, (2013).

2.17 Y. Tomita and K. M. Svore, "Low-distance surface codes under realistic quantum noise," *Physical Review A*, **90**, p. 062320, arXiv:1404.3747, (2014).

2.18 A. G. Fowler, A. C. Whiteside, A. L. McInnes, and A. Rabbani, "Topological Code Autotune," *Physical Review X*, **2**, p. 041003, arXiv:1202.6111, (2012).

2.19 B. J. Brown, N. H. Nickerson, and D. E. Browne, "Fault-tolerant error correction with the gauge color code," *Nature Communications*, **7**, p. 12302, arXiv:1503.08217, (2016).

2.20 A. Y. Kitaev, "Quantum Error Correction with Imperfect Gates," in *Quantum Communication, Computing, and Measurement*, p. 181–188, Boston, MA, Springer US, (1997).

2.21 E. Knill, R. Laflamme, and L. Viola, "Theory of Quantum Error Correction for General Noise," *Physical Review Letters*, **84**, p. 2525–2528, arXiv:quant-ph/9604034, (2000).

2.22 E. Knill, R. Laflamme, and W. H. Zurek, "Resilient Quantum Computation: Error Models and Thresholds," *Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences*, **454**, p. 365–384, arXiv:quant-ph/9702058, (1997).

2.23 D. Aharonov and M. Ben-Or, "Fault-Tolerant Quantum Computation With Constant Error Rate," p. 1–63, arXiv:quant-ph/9906129, (1999).

2.24 D. Aharonov, A. Kitaev, and J. Preskill, "Fault-Tolerant Quantum Computation with Long-Range Correlated Noise," *Physical Review Letters*, **96**, p. 050504, arXiv:quant-ph/0510231, (2006).

2.25 J. Preskill, "Sufficient condition on noise correlations for scalable quantum computing," *Quantum Information & Computation*, **13**, 3-4, p. 181–194, arXiv:1207.6131, (2013).

2.26 P. Baireuther, T. E. O'Brien, B. Tarasinski, and C. W. J. Beenakker, "Machine-learning-assisted correction of correlated qubit errors in a topological code," *Quantum*, **2**, p. 48, arXiv:1705.07855, (2018).

2.27 S. Krastanov and L. Jiang, "Deep Neural Network Probabilistic Decoder for Stabilizer Codes," *Scientific Reports*, **7**, p. 11003, arXiv:1705.09334, (2017).

2.28 J. R. Wootton and D. Loss, "High Threshold Error Correction for the Surface Code," *Physical Review Letters*, **109**, p. 160503, arXiv:1202.4316, (2012).

2.29 B. Criger and I. Ashraf, "Multi-path Summation for Decoding 2D Topological Codes," *Quantum*, **2**, 102, arXiv:1709.02154, (2017).

2.30 S. Varsamopoulos, B. Criger, and K. Bertels, "Decoding small surface codes with feedforward neural networks," *Quantum Science and Technology*, **3**, p. 015004, arXiv:1705.00857, (2018).

2.31 N. Delfosse and J.-P. Tillich, "A decoding algorithm for CSS codes using the X/Z correlations," in *2014 IEEE International Symposium on Information Theory*, i, p. 1071–1075, IEEE, (2014), arXiv:1401.6975.

2.32 A. Hutter and D. Loss, "Breakdown of surface-code error correction due to coupling to a bosonic bath," *Physical Review A*, **89**, p. 042334, arXiv:1402.3108, (2014).

2.33 G. Duclos-Cianci and D. Poulin, "Fast Decoders for Topological Quantum Codes," *Physical Review Letters*, **104**, p. 050504, arXiv:0911.0581, (2010).

2.34 N. H. Nickerson and B. J. Brown, "Analysing correlated noise on the surface code using adaptive decoding algorithms," arXiv:1712.00502, (2017).

2.35 N. Maskara, A. Kubica, and T. Jochym-O'Connor, "Advantages of versatile neural-network decoding for topological codes," arXiv:1802.08680, (2018).

2.36 A. S. Darmawan and D. Poulin, "Linear-time general decoding algorithm for the surface code," *Physical Review E*, **97**, 5, p. 1–5, arXiv:1801.01879, (2018).

2.37 E. Dennis, A. Kitaev, A. Landahl, and J. Preskill, "Topological quantum memory," *Journal of Mathematical Physics*, **43**, p. 4452–4505, arXiv:quant-ph/0110143, (2001).

2.38 K. Fujii, "Quantum information and statistical mechanics: an introduction to frontier," *Interdisciplinary Information Sciences*, **19**, 1, p. 1–15, arXiv:1306.6757, (2013).

2.39 H. Bombin, R. S. Andrist, M. Ohzeki, H. G. Katzgraber, and M. A. Martin-Delgado, "Strong Resilience of Topological Codes to Depolarization," *Physical Review X*, **2**, p. 021004, arXiv:1202.1852, (2012).

2.40 A. A. Kovalev and L. P. Pryadko, "Spin glass reflection of the decoding transition for quantum error correcting codes," arXiv:1311.7688, (2013).

2.41 H. G. Katzgraber, H. Bombin, and M. A. Martin-Delgado, "Error Threshold for Color Codes and Random Three-Body Ising Models," *Physical Review Letters*, **103**, p. 090501, arXiv:0902.4845, (2009).

2.42 A. Kubica, M. E. Beverland, F. Brandao, J. Preskill, and K. M. Svore, "Three-dimensional color code thresholds via statistical-mechanical mapping," arXiv:1708.07131, (2017).

2.43 A. A. Kovalev, S. Prabhakar, I. Dumer, and L. P. Pryadko, "Numerical and analytical bounds on threshold error rates for hypergraph-product codes," *Physical Review A*, **97**, 6, p. 1–14, arXiv:1804.01950, (2018).

2.44 C. Wang, J. Harrington, and J. Preskill, "Confinement-Higgs transition in a disordered gauge theory and the accuracy threshold for quantum memory," *Annals of Physics*, **303**, p. 31–58, arXiv:quant-ph/0207088, (2003).

2.45 R. S. Andrist, H. G. Katzgraber, H. Bombin, and M. A. Martin-Delgado, "Tricolored lattice gauge theory with randomness: Fault tolerance in topological color codes," *New Journal of Physics*, **13**, 2, arXiv:1005.0777, (2011).

2.46 I. Dumer, A. A. Kovalev, and L. P. Pryadko, "Thresholds for Correcting Errors, Erasures, and Faulty Syndrome Measurements in Degenerate Quantum Codes," *Physical Review Letters*, **115**, 5, p. 1–10, arXiv:1412.6172, (2015).

2.47 H. Bombin, "An Introduction to Topological Quantum Codes," arXiv:1311.0277, (2013).

2.48 R. S. Andrist, "Understanding topological quantum error-correction codes using classical spin models," (2012).

2.49 T. Ohno, G. Arakawa, I. Ichinose, and T. Matsui, "Phase structure of the random-plaquette gauge model: accuracy threshold for a toric quantum memory," *Nuclear Physics B*, **697**, p. 462–480, arXiv:quant-ph/0401101, (2004).

2.50 H. Nishimori, "Internal Energy, Specific Heat and Correlation Function of the Bond-Random Ising Model," *Progress of Theoretical Physics*, **66**, p. 1169–1181, (1981).

2.51 S. Bravyi, M. Suchara, and A. Vargo, "Efficient algorithms for maximum likelihood decoding in the surface code," *Physical Review A*, **90**, p. 032326, arXiv:1405.4883, (2014).

2.52 D. Gottesman, "Class of quantum error-correcting codes saturating the quantum Hamming bound," *Physical Review A*, **54**, p. 1862–1868, (1996).

2.53 D. Gottesman, "Stabilizer Codes and Quantum Error Correction," arXiv:quant-ph/9705052, (1997).

2.54 D. Kribs, R. Laflamme, and D. Poulin, "Unified and Generalized Approach to Quantum Error Correction," *Physical Review Letters*, **94**, p. 180501, arXiv:quant-ph/0412076, (2005).

2.55 B. Sutherland, "Two-dimensional hydrogen bonded crystals without the ice rule," *Journal of Mathematical Physics*, **11**, p. 3183–3186, (1970).

2.56 C. Fan and F. Wu, "General Lattice Model of Phase Transitions," *Physical Review B*, **2**, 3, p. 723–733, (1970).

2.57 R. J. Baxter, "Eight-vertex model in lattice statistics," *Physical Review Letters*, **26**, p. 832–833, (1971).

2.58 D. K. Tuckett, C. T. Chubb, and B. J. Brown, *In preparation*, (2018).

2.59 F. Parisen Toldin, A. Pelissetto, and E. Vicari, "Strong-Disorder Paramagnetic-Ferromagnetic Fixed Point in the Square-Lattice ±J Ising Model," *Journal of Statistical Physics*, **135**, p. 1039–1061, (2009).

2.60 A. Honecker, M. Picco, and P. Pujol, "Universality Class of the Nishimori Point in the 2D ±J Random-Bond Ising Model," *Physical Review Letters*, **87**, p. 047201, (2001).

2.61 C. Amoruso and A. K. Hartmann, "Domain-wall energies and magnetization of the two-dimensional random-bond Ising model," *Physical Review B*, **70**, p. 134425, (2004).

2.62 F. Merz and J. T. Chalker, "Two-dimensional random-bond Ising model, free fermions, and the network model," *Physical Review B*, **65**, p. 054425, (2002).

2.63 H. Nishimori, "Duality in Finite-Dimensional Spin Glasses," *Journal of Statistical Physics*, **126**, p. 977–986, arXiv:cond-mat/0602453, (2007).

2.64 M. Ohzeki, "Locations of multicritical points for spin glasses on regular lattices," *Physical Review E*, **79**, p. 021129, arXiv:0811.0464, (2009).

2.65 N. Metropolis, A. W. Rosenbluth, M. N. Rosenbluth, A. H. Teller, and E. Teller, "Equation of State Calculations by Fast Computing Machines," *The Journal of Chemical Physics*, **21**, p. 1087–1092, (1953).

2.66 W. K. Hastings, "Monte Carlo sampling methods using Markov chains and their applications," *Biometrika*, **57**, p. 97–109, (1970).

2.67 C. H. Bennett, "Efficient estimation of free energy differences from Monte Carlo data," *Journal of Computational Physics*, **22**, p. 245–268, (1976).

2.68 Y. Jiang, I. Dumer, A. A. Kovalev, and L. P. Pryadko, "Duality and free energy analyticity bounds for few-body Ising models with extensive homology rank", arXiv:1805.00644, (2018).

2.69 F. Verstraete, M. M. Wolf, D. Perez-Garcia, and J. I. Cirac, "Criticality, the Area Law, and the Computational Power of Projected Entangled Pair States," *Physical Review Letters*, **96**, p. 220601, arXiv:quant-ph/0601075, (2006).

2.70 J. C. Bridgeman and C. T. Chubb, "Hand-waving and interpretive dance: an introductory course on tensor networks," *Journal of Physics A: Mathematical and Theoretical*, **50**, p. 223001, arXiv:1603.03039, (2017).

2.71 D. K. Tuckett, S. D. Bartlett, and S. T. Flammia, "Ultrahigh Error Threshold for Surface Codes with Biased Noise," *Physical Review Letters*, **120**, p. 050505, arXiv:1708.08474, (2018).

2.72 A. S. Darmawan and D. Poulin, "Tensor-Network Simulations of the Surface Code under Realistic Noise," *Physical Review Letters*, **119**, 4, arXiv:1607.06460, (2017).

2.73 I. Arad and Z. Landau, "Quantum computation and the evaluation of tensor networks," *SIAM Journal on Computing*, **39**, 7, p. 3089–3121, arXiv:0805.0040, (2008).

2.74 N. Schuch, M. Wolf, F. Verstraete, and J. Cirac, "Entropy Scaling and Simulability by Matrix Product States," *Physical Review Letters*, **100**, p. 030504, arXiv:0705.0292, (2008).

2.75 U. Schollwöck, "The density-matrix renormalization group in the age of matrix product states," *Annals of Physics*, **326**, p. 96–192, arXiv:1008.3477, (2011).

2.76 F. Verstraete and J. I. Cirac, "Renormalization algorithms for Quantum-Many Body Systems in two and higher dimensions," arXiv:cond-mat/0407066, (2004).

2.77 L. Onsager, "Crystal Statistics. I. A Two-Dimensional Model with an Order-Disorder Transition," *Physical Review*, **65**, p. 117–149, (1944).

2.78 R. Baxter, "Corner transfer matrices," *Physica A: Statistical Mechanics and its Applications*, **106**, p. 18–27, (1981).

2.79 R. J. Baxter, "Corner transfer matrices in statistical mechanics," *Journal of Physics A: Mathematical and Theoretical*, **40**, p. 12577–12588, arXiv:cond-mat/0611167, (2007).

2.80 G. Evenbly and G. Vidal, "Tensor Network Renormalization," *Physical Review Letters*, **115**, p. 180405, arXiv:1412.0732, (2015).

2.81 M. Bal, M. Mariën, J. Haegeman, and F. Verstraete, "Renormalization Group Flows of Hamiltonians Using Tensor Networks," *Physical Review Letters*, **118**, p. 250602, arXiv:1703.00365, (2017).

2.82 M. Levin and C. P. Nave, "Tensor Renormalization Group Approach to Two-Dimensional Classical Lattice Models," *Physical Review Letters*, **99**, p. 120601, arXiv:cond-mat/0611687, (2007).

2.83 Z. Y. Xie, J. Chen, M. P. Qin, J. W. Zhu, L. P. Yang, and T. Xiang, "Coarse-graining renormalization by higher-order singular value decomposition," *Physical Review B*, **86**, p. 045139, arXiv:1201.1144, (2012).

2.84 D. Gottesman, A. Kitaev, and J. Preskill, "Encoding a qubit in an oscillator," *Physical Review A*, **64**, p. 012310, arXiv:quant-ph/0008040, (2000).

2.85 N. Ofek, A. Petrenko, R. Heeres, P. Reinhold, Z. Leghtas, B. Vlastakis, Y. Liu, L. Frunzio, S.M. Girvin, L. Jiang, M. Mirrahimi, M.H. Devoret, and R.J. Schoelkopf, "Demonstrating Quantum Error Correction that Extends the Lifetime of Quantum Information", *Nature*, **536**, p. 441–445, arXiv:1602.04768, (2016).

2.86 G. Dauphinais, L. Ortiz, S. Varona, and M.A. Martin-Delgado, "Quantum Error Correction with the Semion Code", *arXiv:1810.08204*, (2018).

2.87 L. E. Sucar, Probabilistic Graphical Models. Advances in Computer Vision and Pattern Recognition, Springer London (2015).

2.88 M. Palassini and S. Caracciolo, "Universal Finite-Size Scaling Functions in the 3D Ising Spin Glass," *Physical Review Letters*, **82**, p. 5128–5131, arXiv:cond-mat/9904246, (1999).

2.89 N. Maskara, A. Kubica, and T. Jochym-O'Connor, "Advantages of versatile neural-network decoding for topological codes," arXiv:1802.08680, (2018)

Chapter 3

Tailoring surface codes for highly-biased noise

David K. Tuckett[1], Christopher T. Chubb[1], Stephen D. Bartlett[1], Steven T. Flammia[1,2], and Sergey Bravyi[3]

[1] *Centre for Engineered Quantum Systems, School of Physics, The University of Sydney, Sydney, NSW 2006, Australia*

[2] *Yale Quantum Institute, Yale University, New Haven, Connecticut 06520, USA*

[3] *IBM T.J. Watson Research Center, Yorktown Heights, NY 10598, USA*

ArXiv: 1812.08186

Abstract

The surface code, with a simple modification, exhibits ultra-high error correction thresholds when the noise is biased towards dephasing. Here we identify features of the surface code responsible for these ultra-high thresholds and show how to exploit them to achieve significant improvement in logical failure rate. First, we consider the infinite bias limit, meaning pure dephasing. We prove that the error threshold of the modified surface code for pure dephasing noise is 50%, i.e., that all qubits are fully dephased, and this threshold can be achieved by a polynomial time decoding algorithm. The sub-threshold behavior of the code is controlled by the parameter $g = \gcd(j, k)$, where j and k are dimensions of the surface code lattice. We demonstrate a significant improvement in logical failure rate with pure dephasing for co-prime codes that have $g = 1$. The effect is dramatic: the same logical failure rate achievable with a standard surface code and n physical qubits can be obtained with a co-prime surface code using only $O(\sqrt{n})$ physical qubits. Finally, we use an approximate optimal decoder to demonstrate that this improvement persists for a general Pauli noise biased towards dephasing. In particular, we observe a significant improvement in logical failure rate against biased noise using a *smaller* co-prime $(j-1)\times j$ surface code compared with a square $j\times j$ surface code.

3.1 Introduction

Quantum error correcting codes are expected to play a fundamental role in enabling quantum computers to operate at large scale in the presence of noise. The surface code [3.1], an example of a

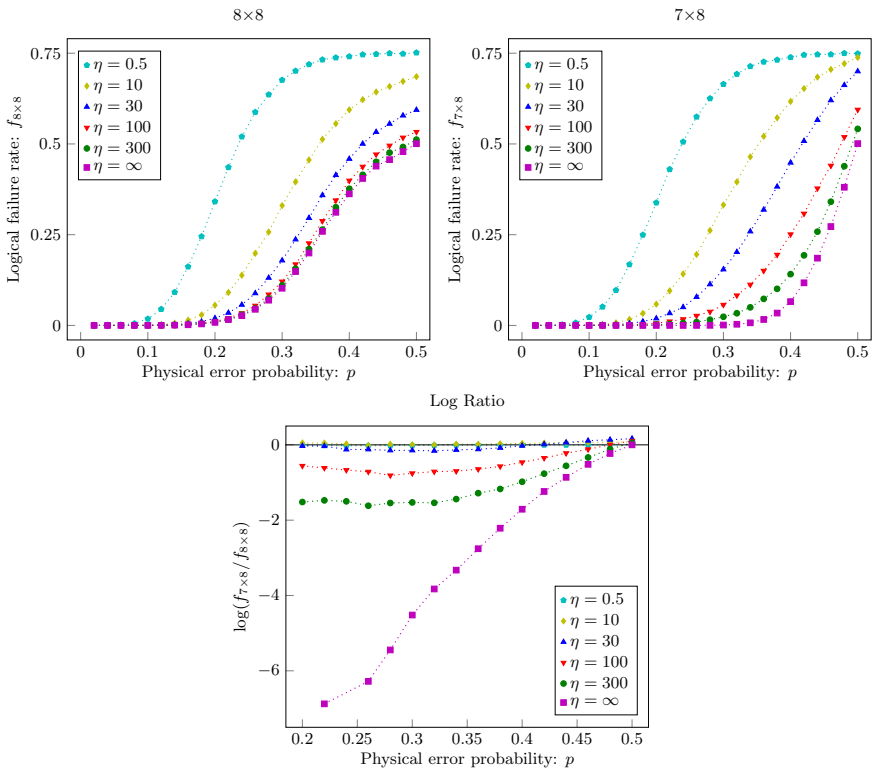


Figure 3.1: Logical failure rates $f_{8\times8}$ and $f_{7\times8}$ as a function of physical error probability p for small comparable square 8×8 and co-prime 7×8 codes, and the logarithm of the ratio of logical failure rates $\log(f_{7\times8}/f_{8\times8})$ with noise biases $\eta \in \{0.5, 10, 30, 100, 300, \infty\}$. Data points are sample means over 60 000 runs using the BSV decoder [3.8] with approximation parameter $\chi = 96$. Dotted lines connect successive data points for a given η.

topological stabiliser code [3.2], is one of the most studied and promising candidates, giving excellence performance for error correction while only requiring check operators (stabilisers) acting on a small number of neighbouring qubits [3.3].

The error correction threshold of a code family, which denotes the physical error rate below which the logical failure rate can be made arbitrarily small by increasing the code size, is strongly dependent on the noise model. The most commonly studied noise model is uniform depolarisation of all qubits, where independent single-qubit Pauli X, Y, and Z errors occur at equal rates. However, in many quantum architectures the computational basis corresponds to states of non-degenerate energy levels, which causes dephasing noise, whereby Z error are more likely to occur than other errors. Examples of systems known to experience such Z-biased noise include superconducting qubits [3.4], quantum dots [3.5], and trapped ions [3.6], among others. Recently, it was shown that, with a simple modification, the surface code exhibits ultra-high thresholds with such Z-biased noise [3.7].

In this paper, we identify and characterise the features of the modified surface code that contribute to its ultra-high thresholds with Z-biased noise and demonstrate a further significant improvement in logical failure rate. We note that the modification of the surface code, described in Ref. [3.7], simply exchanges the roles of Z and Y operators in stabiliser and logical operator definitions. Therefore,

results for the modified surface code with Z-biased noise can equivalently be expressed in terms of the standard surface code with Y-biased noise, where Y errors occur more frequently than X or Z errors. In order to frame our analysis in the context of the familiar standard surface code, and to simplify comparison with other codes, we consider pure Y noise and Y-biased noise on the standard surface code throughout this paper. However, we emphasise that our results apply equally to the modified surface code with pure Z noise or the Z-biased noise prevalent in many quantum architectures.

Our main analytical result is a structural theorem that reveals a hidden concatenated form of the surface code. We shotw that, in the limit of pure Y noise, the surface code can be viewed as a classical concatenated code with two concatenation levels. The top level contains the so-called cycle code whose parity checks correspond to cycles in the complete graph. The bottom level contains several copies of the repetition code. We prove that the cycle code has the error threshold of 50% and give an efficient decoding algorithm that achieves this threshold. As a corollary, we show that the threshold of the surface code with pure Y noise is 50% thus answering an open question posed in Ref. [3.7]. The concatenated structure described above is controlled by the parameter $g = \gcd(j, k)$, where j and k are dimensions of the surface code lattice. In particular, the top-level cycle code has length $O(g^2)$ while the bottom-level repetition codes have length $O(jk/g^2)$. Two important special cases are *co-prime codes* and *square codes* that have $g = 1$ and $g = j = k$ respectively. Informally, a co-prime surface code can be viewed as a repetition code, whereas a square surface code can be viewed as a cycle code (in the limit of pure Y noise). Although the repetition and the cycle codes both have 50% error threshold, we argue that the former performs much better in the sub-threshold regime. This suggests that co-prime surface codes may have an intrinsic advantage in correcting strongly biased noise.

We present further insights into the origins of the ultra-high threshold by investigating the form of logical operators. We show that logical operators consistent with pure Y noise are much rarer and heavier than those consistent with pure X or Z noise, and their structure depends strongly on the parameter g. In particular, there are 2^{g-1} Y-type logical operators of which the minimum weight is $(2g-1)(jk/g^2)$. This compares to $2^{j(k-1)}$ X-type logical operators of which the minimum weight is j. In the case of co-prime codes there is only one Y-type logical operator and its weight is jk. Hence the distance of co-prime codes to pure Y noise is $O(n)$ whereas for square codes it is $O(\sqrt{n})$.

Based on the intuitions behind these results, we define an exact optimal maximum likelihood Y-noise decoder that is efficient for co-prime codes and tractable for moderate-sized square codes. We use this decoder to confirm in numerical simulations the 50% threshold for the surface code with pure Y noise and demonstrate a significant reduction in logical failure rate for co-prime codes compared to square codes with pure Y noise. In particular, we demonstrate that the logical failure rate decays exponentially with the distance to pure Y noise such that a target logical failure rate may be achieved with quadratically fewer physical qubits by using co-prime codes.

Finally, we demonstrate a remarkable property of surface codes: by *removing* physical qubits appropriately from square surface codes to yield one with linear dimensions that are co-prime, we observe a significant reduction in logical failure rate with biased noise. Specifically, we use the tensor network decoder of Ref. [3.8], which is an approximate maximum likelihood decoder, to demonstrate the aforementioned significant reduction in logical failure rate against biased noise that is achieved using a smaller co-prime $(j-1)\times j$ code compared with a square $j\times j$ code. Fig. 3.1 summarises this result, comparing logical failure rate as a function of physical error probability for a co-prime 7×8 code and a square 8×8 code across a range of biases, where bias η is the ratio of the probability of a Y error to the probability of an X or Z error occurring. We see that the advantage of the co-prime

code over the square code is greatest in the limit of pure Y noise ($\eta = \infty$) and remains significant down to more modest intermediate bias, $\eta = 100$ (where Y errors are 100 times more likely than both X and Z errors). We further argue that the relative advantage of co-prime codes over square codes increases with code size, motivating the search for efficient near-optimal biased-noise decoders for co-prime codes.

Note this performance with biased noise is not shared by all topological codes; in stark contrast, the triangular 6.6.6 colour code [3.9] exhibits a decrease in threshold with bias; see the Appendix.

The paper is structured as follows. Section 3.2 provides some definitions used throughout the paper. Our main analytical results for surface codes with pure Y noise are in Section 3.3. Section 3.4 presents our numerical results for surface codes with pure Y noise. The demonstration of the advantage of using co-prime codes with biased noise is detailed in Section 3.5. Section 3.6 includes a discussion of our results in the context of prior work and raises some open questions for future work.

3.2 Definitions

Surface code.— We consider $j \times k$ surface codes [3.1] on a square lattice with "smooth" top/bottom boundaries and "rough" left/right boundaries. Physical qubits are associated with edges on the lattice. Following the usual convention, stabiliser generators consist of X operators on edges around vertices, $A_v = \prod_{e \in v} X_e$, and Z operators on edges around plaquettes, $B_p = \prod_{e \in p} Z_e$. The stabiliser group is, therefore, $\mathcal{G} = \langle A_v, B_p \rangle$. The $\bar{X}$ ($\bar{Z}$) logical operator consists of X (Z) operators along the left (top) edge, such that $X, Z \in \mathcal{C}(\mathcal{G}) \setminus \mathcal{G}$ and $\bar{X}\bar{Z} = -\bar{Z}\bar{X}$, where $\mathcal{C}(\mathcal{G}) = \{f \in \mathcal{P} : fg = gf \ \forall \ g \in \mathcal{G}\}$ is the centraliser of $\mathcal{G}$ and $\mathcal{P}$ is the group of n-qubit Paulis. As such, a $j \times k$ surface code encodes one logical qubit into $n = 2jk - j - k + 1$ physical qubits with distance $d = \min(j, k)$. Fig. 3.2 illustrates a 4×5 surface code.

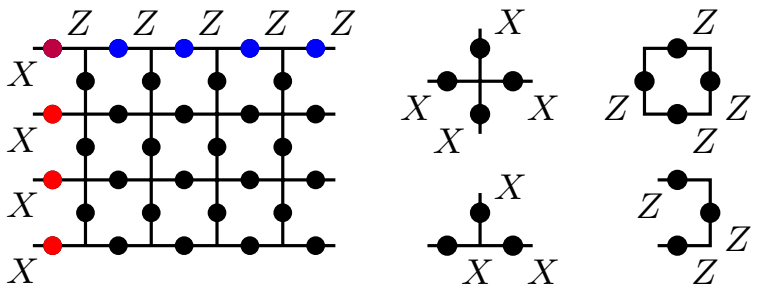


Figure 3.2: Standard 4×5 surface code, with logical operators given by a product of X along the left edge and a product of Z along the top edge. Stabiliser generators are shown at right.

Surface code families — For $j \times k$ surface codes, we define the following code families: *square* where $j = k$; $\gcd(j, k) = g$ constant; and *co-prime* where g=1 (special case of g constant).

Y-type stabilisers and logical operators.— We define a Y-type stabiliser to be any operator on a code that is in the stabiliser group $\mathcal{G}$ and consists only of Y and identity single-qubit Paulis. We define a Y-type logical operator to be any operator on a code that is in $\mathcal{C}(\mathcal{G}) \setminus \mathcal{G}$ and consists only of Y and identity single-qubit Paulis. We define X- and Z-type stabilisers and logical operators analogously.

As usual, the weight of an operator is the number of non-identity single-qubit Paulis applied by the operator.

Y-distance.—We define Y-distance, or distance d_Y to pure Y noise, of a code as the weight of the minimum-weight Y-type logical operator. X- and Z-distance are defined analogously. The overall distance of the code is defined in the usual way and is upper-bounded by $\min(d_X, d_Y, d_Z)$.

Y-biased noise.— Several conventions have previously been used to define biased Pauli noise models [3.4, 3.7, 3.10–3.22]. We adapt the approach of Ref. [3.7] to Y-biased noise, by considering an independent, identically distributed Pauli noise model defined by an array $\boldsymbol{p} = (1-p, p_X, p_Y, p_Z)$ corresponding to the probabilities of each single-qubit Pauli I (no error), X, Y, and Z, respectively, such that the probability of any error on a single-qubit is $p = p_X + p_Y + p_Z$. We define bias η to be the ratio of the probability of a Y error to the probability of a non-Y error such that $\eta = p_Y/(p_X + p_Z)$. For simplicity, we restrict to the case $p_X = p_Z$. With this definition $\eta = 1/2$ corresponds to standard depolarising noise with $p_X = p_Y = p_Z = p/3$ and the limit $\eta \to \infty$ corresponds to pure Y noise, i.e., only Y errors with probability p. We define X- and Z-biased noise analogously.

3.3 Features of surface codes with pure Y noise

In this section, we present our analytical results for surface codes with pure Y noise. We start by highlighting the specificities of syndromes of pure Y noise. Our main result reveals that error correction with the surface code with pure Y noise is equivalent to a concatenation of two classical codes: the repetition code at the bottom level and the cycle code at the top level. As a corollary, we show that the surface code with pure Y noise has a threshold of 50%. We also highlight that for $j \times k$ surface codes with small $g = \gcd(j, k)$, the more effective repetition code dominates the performance of the code. Finally, we give explicit formulas for the minimum weight and count of Y-type logical operators. These results explain the ultra-high thresholds of the surface code with Y-biased noise, as shown in Ref. [3.7], and the lower logical failure rates seen with co-prime surface codes, presented in Sections 3.4 and 3.5.

3.3.1 Syndromes of pure Y noise

An obvious feature of Y noise on the surface code is that Y errors anticommute with both X- and Z-type stabiliser generators, providing additional bits of syndrome information. For comparison, Fig. 3.3 shows a sample of Y-error configurations alongside identically placed X- and Z-error configurations with corresponding anticommuting syndrome locations for each error type. In each case, we see that Y-error strings anticommute with more syndrome locations than X- or Z-error strings, providing the decoder with more information about the location of errors to be corrected.

We remark that the displacement between the X- and Z-type stabiliser generators appears to be significant. For example, the colour 6.6.6 code has co-located X- and Z-type stabiliser generators, so that, even if Y errors anticommute with more stabiliser generators, the number of distinct syndrome locations triggered by Y errors is no greater than for X or Z errors.

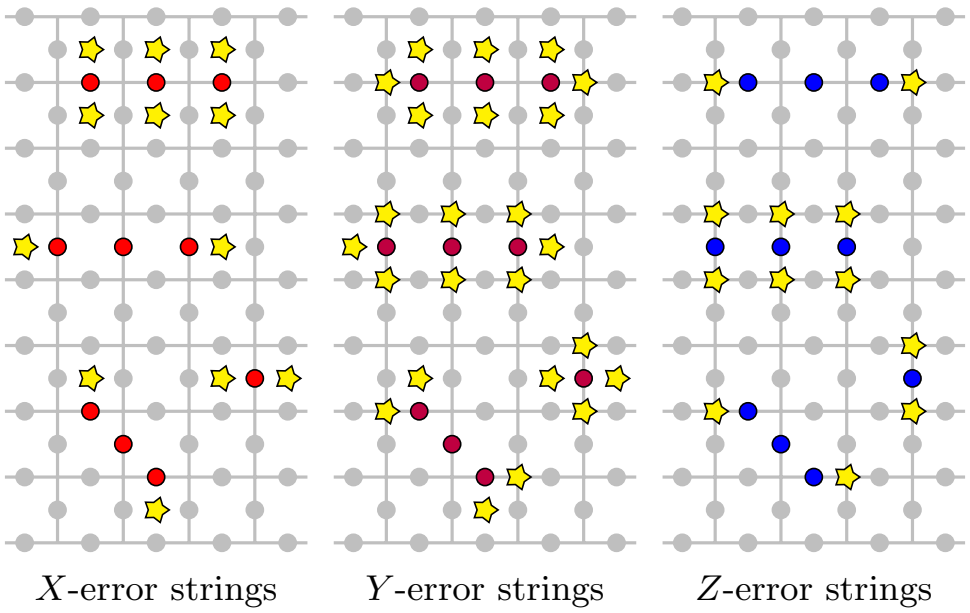


Figure 3.3: A sample of X-, Y- and Z-error strings, indicated by coloured circles, with corresponding anti-commuting syndrome locations, indicated by yellow stars.

3.3.2 Structure of surface codes with pure Y noise

In this section we consider surface codes subject to pure Y noise. We describe a polynomial time decoding algorithm and prove that it achieves an error threshold of 50%. We also derive an exponential upper bound on the probability of logical errors in the sub-threshold regime. Our main result is a structural theorem that reveals a hidden concatenated structure of the surface code and highlights the role of the parameter $g = \gcd(j, k)$. The theorem implies that error correction with the surface code subject to Y noise can be viewed as a concatenation of two classical codes: the repetition code at the bottom level and the so-called cycle code at the top level. Both codes admit efficient decoding algorithms and have an error threshold of 50%, although the repetition code scores much better in terms of the logical error probability. We show that for a fixed number of qubits, the size of each code can vary drastically depending on the value of g. Loosely speaking, the error correction workload is shared between the two codes such that for small g the dominant contribution comes from the more effective repetition code. This explains the enhanced performance of co-prime surface codes ($g = 1$) observed in the numerics.

We will demonstrate this concatenated structure by studying the restricted patterns of syndrome locations which can arise from pure-Y errors. Specifically we will see that this syndrome information is highly redundant. We will then show that, by considering a code involving concatenations of the repetition code, we can remove this redundancy to reveal the underlying effective classical code: the cycle code.

Concatenated structure

Consider a Pauli error

$$P(y) \equiv Y_1^{y_1} \otimes Y_2^{y_2} \otimes \cdots \otimes Y_n^{y_n}, \tag{3.1}$$

where $y \in \{0,1\}^n$. As described in Section 3.3.1, the syndrome of $P(y)$ is given by

$$a_v(y) = \sum_{e \in v} y_e \qquad \text{and} \qquad b_p(y) = \sum_{e \in p} y_e \tag{3.2}$$

where v and p run over all vertices and all plaquettes of the lattice and the sums are modulo two. A decoding algorithm takes as input the error syndrome and outputs a candidate recovery operator $P(y')$ that agrees with the observed syndrome. The decoding succeeds if $y' = y$ and fails otherwise. (More generally, the decoder only needs to identify the equivalence class of errors that contains $P(y)$, where the equivalence is defined modulo stabilisers of the surface code.)

Consider a classical linear code of length n defined by the parity checks $a_v(y) = 0$ and $b_p(y) = 0$ for all v, p. We shall refer to this code as a *Y-code*. As described above, error correction for the surface code subject to Y-noise is equivalent to error correction for the Y-code subject to classical bit-flip errors. We shall now establish the structure of the Y-code. For any integer $m \geq 3$, let K_m be the complete graph with m vertices and $e = m(m-1)/2$ edges. Consider bit strings $x \in \{0,1\}^e$ such that bits of x are associated with edges of the graph K_m. Let $x_{i,j}$ be the bit associated with an edge (i,j). Here it is understood that $x_{i,j} = x_{j,i}$. Define a *cycle code* $\mathcal{C}_m$ of order m that encodes $m-1$ bits into e bits with parity checks

$$x_{i,j} \oplus x_{j,k} \oplus x_{i,k} = 0 \qquad \text{for all } 1 \leq i < j < k \leq m. \tag{3.3}$$

Thus parity checks of $\mathcal{C}_m$ correspond to cycles (triangles) in the graph K_m. Note that Eq. (3.3) defines a redundant set of parity checks. It is well-known that any connected graph with m vertices and e edges has $e - m + 1$ independent cycles. Thus $\mathcal{C}_m$ has $e - (m-1)$ independent parity checks. The number of encoded bits is $m - 1$. Note that $\mathcal{C}_2$ is a trivial code (it has no parity checks). Let $\mathrm{REP}(m)$ be the repetition code that encodes one bit into m bits. We can now describe the structure of the Y-code.

Theorem 3.1 (Y-code structure)**.** *The Y-code is a concatenation of the cycle code $\mathcal{C}_{g+1}$ at the top level and $g(g+1)/2$ repetition codes at the bottom level. The latter consists of repetition codes* $\mathrm{REP}(jk/g^2)$, $\mathrm{REP}(2jk/g^2)$, *and* $\mathrm{REP}(4jk/g^2)$ *with multiplicities* 1, $2(g-1)$, *and* $g(g+1)/2 - 2g + 1$ *respectively.*

An important corollary of the theorem is that a decoding algorithm for the cycle code can be directly applied to correcting Y errors in the surface code. Indeed, a decoder for the Y-code can be constructed in a level-by-level fashion such that the bottom level repetition codes are decoded first and the top level cycle code is decoded afterwards.

For example, Theorem 3.1 implies that, with pure Y noise, a co-prime ($g = 1$) surface code is essentially a single repetition code of size growing linearly with n, whereas a square surface code is equivalent to the concatenation of bottom-level fixed-size repetition codes REP(1), REP(2), REP(4) and a top-level cycle code of size growing linearly with n, where n is the number of physical qubits in the surface code.

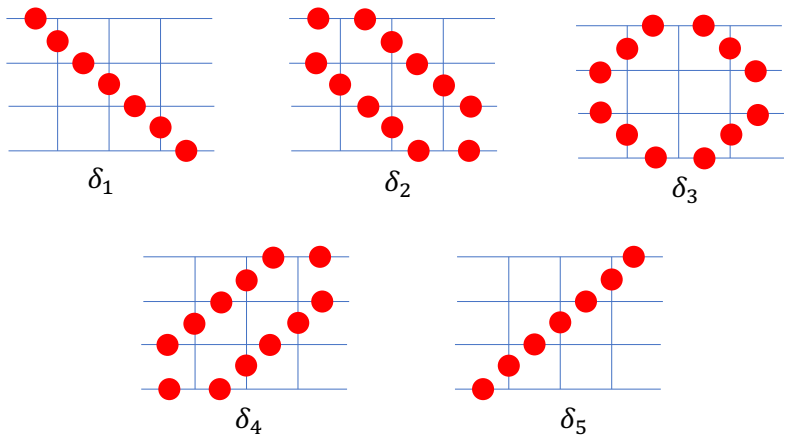


Figure 3.4: Diagonals δ^i for the 4×4 surface code. We consider the symmetry group $\mathcal{R}$ generated by reflections of the lattice against δ^1 and δ^5. Note that any diagonal δ^i is symmetric under reflections from $\mathcal{R}$.

Proof. Let us first prove the theorem in the special case of square surface codes, $j = k = g$. Let $\mathcal{G} \subset \{0,1\}^n$ be the codespace of the Y-code. We shall use a particular basis set of codewords called *diagonals*. The $j \times j$ lattice has $j+1$ diagonals denoted $\delta^1, \delta^2, \ldots, \delta^{j+1} \in \mathcal{G}$; see Fig. 3.4. Given a codeword $y \in \mathcal{G}$, let $\partial y \in \{0,1\}^j$ be the restriction of y onto the top horizontal row of edges in the surface code lattice. We claim that y is uniquely determined by ∂y. Indeed let $H_1, \ldots, H_j$ be the rows of horizontal edges (counting from the top). Let $V_2, \ldots, V_j$ be the rows of vertical edges (counting from the top). By definition, the restriction of y onto H_1 coincides with ∂y. Suppose the restriction of y onto $H_1 V_2 \ldots H_p$ is already determined (initially $p = 1$). Vertex parity checks $a_v(y) = 0$ located at the row H_p then determine the restriction of y onto V_{p+1}. Likewise, suppose the restriction of y onto $H_1 V_2 \ldots H_p V_p$ is already determined. Plaquette parity checks $b_p(y) = 0$ located at the row V_p then determine the restriction of y onto H_{p+1}. Proceeding inductively shows that any codeword $y \in \mathcal{G}$ is uniquely determined by ∂y.

Define bit strings

$$e^1 = 100\ldots0, \quad e^2 = 010\ldots0, \quad e^3 = 001\ldots0 \quad \text{etc.}$$

Then $\partial\delta^1 = e^1$, $\partial\delta^i = e^{i-1} + e^i$ for $2 \le i \le j$, and $\partial\delta^{j+1} = e^j$; see Fig. 3.4. It follows that $\partial\delta^1, \ldots, \partial\delta^j$ span the binary space $\{0,1\}^j$. Accordingly, the diagonals $\delta^1, \ldots, \delta^j$ span the codespace $\mathcal{G}$ and

$$\delta^{j+1} = \delta^1 \oplus \delta^2 \oplus \ldots \oplus \delta^j.$$

In particular, $\dim(\mathcal{G}) = j$, that is, the Y-code encodes j bits into n bits.

Let $\mathcal{R} \cong \mathbb{Z}_2 \times \mathbb{Z}_2$ be a group generated by reflections of the lattice against the diagonals δ^1 and δ^{j+1}. Note that any diagonal δ^i is invariant under reflections from $\mathcal{R}$, see Fig. 3.4. Suppose f is an edge of the surface code lattice. Let $\mathcal{R}(f)$ be the orbit of f under the action of $\mathcal{R}$. The above shows that any diagonal δ^i is constant on orbits of $\mathcal{R}$, that is, $\mathcal{R}(f) = \mathcal{R}(g)$ implies that $\delta^i_f = \delta^i_g$. Since the diagonals δ^i span the full codespace $\mathcal{G}$, we conclude that any codeword $y \in \mathcal{G}$ is constant on orbits of $\mathcal{R}$, that is, $\mathcal{R}(f) = \mathcal{R}(g)$ implies that $y_f = y_g$. Equivalently, each orbit of $\mathcal{R}$ of size m gives rise to

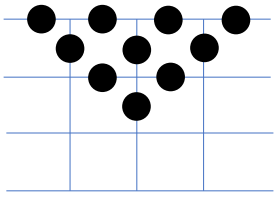


Figure 3.5: A set of qubits $\mathcal{O}$ such that each orbit of $\mathcal{R}$ contains exactly one qubit from $\mathcal{O}$. In this example the group $\mathcal{R}$ has 10 orbits of size $1, 2, 4$.

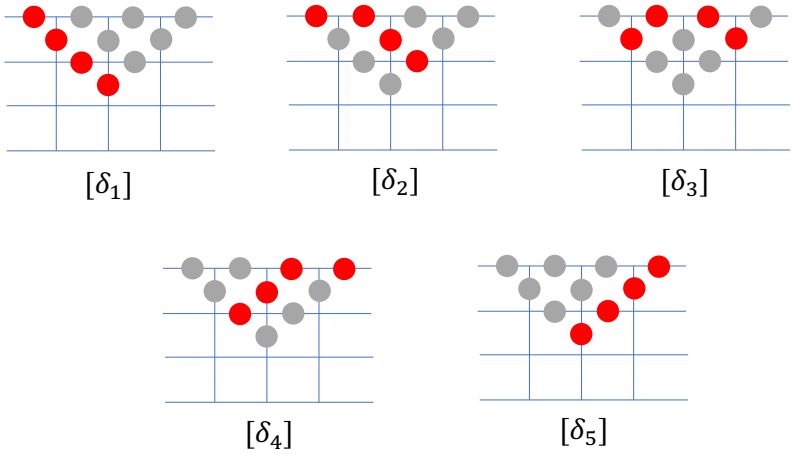


Figure 3.6: Restrictions of the diagonal δ^i onto $\mathcal{O}$ define a basis set of codewords for the top level code.

the repetition code REP(m). A simple counting shows that $\mathcal{R}$ has a single orbit of size 1 (the central vertical edge), $2(j-1)$ orbits of size 2 (pairs of qubits located on the diagonals δ^1 and δ^{j+1}), whereas all remaining orbits have size 4. This proves the last statement of the theorem (in the special case $j = k$).

Fix a set of qubits $\mathcal{O}$ such that each orbit of $\mathcal{R}$ contains exactly one qubit from $\mathcal{O}$. In other words, $\mathcal{O}$ is a set of orbit representatives. We shall choose $\mathcal{O}$ as shown on Fig. 3.5. A simple counting shows that $|\mathcal{O}| = j(j+1)/2$. Consider a codeword $y \in \mathcal{G}$ and let $[y] \in \{0,1\}^{|\mathcal{O}|}$ be a vector obtained by restricting y onto $\mathcal{O}$. We define the top-level code as a linear subspace $\mathcal{L} \subseteq \{0,1\}^{|\mathcal{O}|}$ spanned by vectors $[y]$ with $y \in \mathcal{G}$. Equivalently, $\mathcal{L}$ is spanned by vectors $[\delta^i]$ with $i = 1, \ldots, j+1$. A direct inspection shows that each qubit $e \in \mathcal{O}$ belongs to exactly two vectors $[\delta^i]$ and $[\delta^k]$ for some $i \neq k$, see Fig. 3.6 for an example. Thus one can identify $\mathcal{O}$ with the set of edges of the complete graph K_{j+1}, whereas the vectors $[\delta^i]$ can be identified with "vertex stabilisers" in K_{j+1}. In other words, the support of each vector $[\delta^i]$ coincides with the set of edges incident to some vertex of K_{j+1}. We conclude that parity checks of $\mathcal{L}$ corresponds to closed loops in K_{j+1}. Thus the top-level code coincides with the cycle code $\mathcal{C}_{j+1}$.

The above proves the theorem in the special case $j = k$. Consider now the general case $j \neq k$. Let us tile the surface code lattice by $t = jk/g^2$ tiles of size $g \times g$ as shown on Fig. 3.7. Note that each horizontal edge is fully contained in some tile. Let us say that a vertical edge is a boundary edge

if it overlaps with the boundary of some adjacent tiles. If one ignores the boundary edges, each tile contains a single copy of the $g \times g$ surface code. For each tile define the diagonals $\delta^1, \delta^2, \ldots, \delta^{g+1}$ as

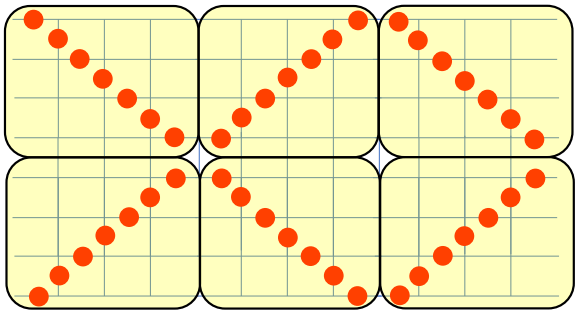


Figure 3.7: Partition of the 8×12 surface code into 4×4 tiles. Solid red circles: the extended diagonal Δ^1 alternating between δ^1 and δ^5, see Fig. 3.4.

above. Let $\mathcal{G}$ be the codespace of the Y-code for the full $j \times k$ lattice. Recall that any codeword $y \in \mathcal{G}$ is is fully determined by its projection ∂y onto the top horizontal row of edges. Using this property one can easily verify that the codespace $\mathcal{G}$ is spanned by "extended diagonals" Δ^i such that the restriction of Δ^i onto the top left tile coincides with δ^i and Δ^i alternates between δ^i and δ^{g+2-i} in a checkerboard fashion, see Fig. 3.8. An example of the extended diagonal Δ^1 is shown on Fig. 3.7. By definition,

$$\Delta^i = \begin{array}{|c|c|c|} \hline \delta^i & \delta^{g+2-i} & \delta^i \\ \hline \delta^{g+2-i} & \delta^i & \delta^{g+2-i} \\ \hline \end{array}$$

Figure 3.8: Extended diagonal Δ^i.

Δ^i has no support on the boundary edges. This implies that the Y-code has a weight-1 parity check for each boundary edge. Ignoring such weight-1 checks, each codeword Δ^i consists of t copies of the diagonal δ^i with some copies being reflected. Considering t copies of each codeword instead of a single copy is equivalent to replacing the repetition codes REP(1), REP(2), REP(4) in the above analysis by REP(t), REP($2t$), REP($4t$), where $t = jk/g^2$ is the number of tiles. □

Decoding the cycle code

Here we consider the cycle code subject to random errors. We give a polynomial-time decoding algorithm that achieves the error threshold of 50%. Fix some integer $m \geq 3$ and consider the cycle code $\mathcal{C}_m$ defined in Section 3.3.2. Recall that $\mathcal{C}_m$ has length $n = m(m-1)/2$. We consider independent and identically distributed (i.i.d.) bit-flip errors such that each bit is flipped with probability $p \in [0, 1/2)$. Define an error bias $\epsilon > 0$ such that

$$2p(1-p) = \frac{1}{2} - \epsilon. \tag{3.4}$$

Lemma 3.1 (Cycle code decoder)**.** *Let $e \in \{0,1\}^n$ be a random iid error with a bias ϵ. There exists an algorithm that takes as input the syndrome of e and outputs a bit string $e' \in \{0,1\}^n$ such that*

$$\text{Prob}[e' = e] \geq 1 - 2m^2 \cdot \exp\left(-2\epsilon^2 m\right). \tag{3.5}$$

The algorithm has runtime $O(m^3)$.

Proof. Recall that the cycle code $\mathcal{C}_m$ is defined on the complete graph with m vertices such that each bit of $\mathcal{C}_m$ is located on some edge (i,j) of the graph. Let $e_{i,j}$ be the error bit associated with an edge (i,j). We begin by giving a subroutine that identifies a single error bit $e_{i,j}$. Without loss of generality, consider the edge $(1,2)$. This edge is contained in $m-2$ triangles that give rise to syndrome bits

$$\begin{aligned} s_3 &= e_{1,2} \oplus e_{2,3} \oplus e_{3,1}, \\ s_4 &= e_{1,2} \oplus e_{2,4} \oplus e_{4,1}, \\ &\ldots \\ s_m &= e_{1,2} \oplus e_{2,m} \oplus e_{m,1}. \end{aligned} \tag{3.6}$$

Since errors on different edges of each triangle are independent, the conditional probability distributions of syndromes s_j for a given error bit $e_{1,2}$ are

$$\begin{aligned} \text{Prob}[s_j = 1 | e_{1,2} = 0] &= \frac{1}{2} - \epsilon, \\ \text{Prob}[s_j = 0 | e_{1,2} = 0] &= \frac{1}{2} + \epsilon, \\ \text{Prob}[s_j = 1 | e_{1,2} = 1] &= \frac{1}{2} + \epsilon, \\ \text{Prob}[s_j = 0 | e_{1,2} = 1] &= \frac{1}{2} - \epsilon. \end{aligned}$$

Furthermore, since different triangles in Eq. (3.6) intersect only on the edge $(1,2)$, we have

$$\text{Prob}[s_3, \ldots, s_m | e_{1,2}] = \prod_{j=3}^{m} \text{Prob}[s_j | e_{1,2}]. \tag{3.7}$$

This is an iid distribution of $m-2$ bits which is ϵ-biased towards $e_{1,2}$. Hoeffding's inequality gives

$$\text{Prob}[s_3 + \ldots + s_m \geq m/2 | e_{1,2} = 0] \leq 4 \exp\left(-2\epsilon^2 m\right)$$

and

$$\text{Prob}[s_3 + \ldots + s_m \leq m/2 | e_{1,2} = 1] \leq 4 \exp\left(-2\epsilon^2 m\right).$$

The desired subroutine outputs $e_{1,2} = 0$ if $s_3 + \ldots + s_m \leq m/2$ and $e_{1,2} = 1$ otherwise. Clearly, the above calculations take time $O(m)$.

The full decoding algorithm applies the above subroutine independently to each edge of the graph learning error bits one by one. By the union bound, such algorithm misidentifies the error with probability at most $2m^2 \exp\left(-2\epsilon^2 m\right)$ since the complete graph K_m has $m(m-1)/2$ edges. The overall runtime of the algorithm is $O(m^3)$. □

Note that the decoding algorithm of Lemma 3.1 can be viewed as a single round of the standard Belief Propagation algorithm, which is commonly used to decode classical low-density parity check (LDPC) codes. Also recall that the cycle code $\mathcal{C}_m$ has length $n \sim m^2/2$. Thus the probability of a logical error in Eq. (3.5) decays exponentially with $\sqrt{n}$ (this scaling is unavoidable since the cycle code $\mathcal{C}_m$ has distance $O(m)$). As a consequence, the proposed decoder performs very poorly in the small bias regime. For example, reducing the error rate from 49% to 1% would require code length $n \approx 10^{17}$ (here we used Eq. (3.5) as a rough estimate of the logical error probability). In contrast, the logical error probability of the repetition code REP(n) decays exponentially with n.

3.3.3 Threshold of the surface code with pure Y noise

The surface code with pure Y noise is equivalent to a concatenation of two classical codes, as shown above, and both of these classical codes have thresholds of 50%. These results lead directly to the fact that the threshold of the surface code with pure Y noise is 50%. Indeed, let us employ the level-by-level decoding strategy such that the bottom-level repetition codes are decoded first. Assume that the pure Y noise has error rate $p < 1/2$. Then the j-th repetition code makes a logical error with probability $p_j \leq p < 1/2$. The effective error model for the top-level cycle code is a product of symmetric binary channels with error rates $p_1, \ldots, p_m \leq p$, where $m = g(g+1)/2$ is the length of the cycle code. One can easily verify that the decoder of Lemma 3.1 corrects such random error with probability given by Eqs. (3.4) and (3.5). Finally, Theorem 3.1 implies that each parity check of the repetition or the cycle code is a linear combination (modulo two) of the plaquette and vertex parity checks of Eq. (3.2). The coefficients in this linear combination can be found by solving a suitable system of linear equations in time $O(n^3)$. This enables an efficient conversion between the surface code syndrome and the syndromes of the bottom-level and the top-level code. To conclude, Theorem 3.1 and Lemma 3.1 have the following corollary.

Corollary 3.1 (Y-threshold)**.** *The error correction threshold for the surface code with pure Y noise is* 50%. *This error threshold can be achieved by a polynomial-time decoding algorithm.*

A numerical demonstration of the 50% threshold of the surface code with pure Y noise is given in Section 3.4.1.

In previous work [3.7], we used the BSV decoder [3.8] to estimate the error threshold of the surface code with a range of biases including with pure Y noise (equivalently pure Z noise on the modified surface code of Ref. [3.7]). The BSV decoder is tuned via an approximation parameter χ, which defines the scale of correlations between syndrome bits. With χ exponentially large in the number of physical qubits, the approximation becomes exact but decoding becomes inefficient. In order to keep the simulations tractable, we used $\chi = 48$ and found that the decoder performance saturated for large bias, $\eta \geq 300$, giving an estimate for the threshold with pure Y noise of 43.7(1)%, where error bars correspond to fitting variance not decoder approximation. We noted that, although clear thresholding behavior was seen, the BSV decoder had not completely converged with $\chi = 48$. Corollary 3.1 indicates that the saturation of the decoder performance was likely a side effect of having too small a value of χ.

3.3.4 Y-type logical operators

The structure of surface codes with pure Y noise, described in Section 3.3.2, also manifests itself in the structure and, consequently, the minimum weight and count of Y-type logical operators, i.e. logical operators consisting only of Y and identity single-qubit Paulis. In this section, we give explicit formulas for the minimum weight and count of Y-type logical operators. Highlighting the cases of co-prime and square codes, as well as comparing the formulas to those for X- and Z-type logical operators, we remark on how the minimum weight and count of Y-type logical operators contributes to the performance advantage with pure Y noise and Y-biased noise seen in Ref. [3.7], for surface codes in general, and in Sections 3.4 and 3.5, for co-prime codes in particular.

Logical operator minimum weight

We show that the minimum-weight Y-type logical operator is comparatively heavy. The X-distance, d_X, of a code is the weight of the minimum-weight X-type logical operator. Clearly, the minimum-weight X-type logical operator on a $j\times k$ code is a full column of X operators on horizontal edges, and hence $d_X = j$; similarly $d_Z = k$. It is also clear that the minimum-weight Y-type logical operator on a square $j\times j$ code is a full diagonal of Y operators, and hence $d_Y = 2j - 1$. From the proof of Theorem 3.1, it is apparent that, in the case of pure Y noise, a $j\times k$ surface code can be viewed as a tiling of jk/g^2 copies of a square $g\times g$ code, where $g = \gcd(j, k)$. Therefore, the Y-distance of a general $j\times k$ surface code is given by the following corollary.

Corollary 3.2 (Y-distance)**.** *For a $j\times k$ surface code, the weight of the minimum-weight Y-type logical operator, and hence the distance of the code to pure Y noise, is*

$$d_Y = \frac{(2g-1)jk}{g^2}$$

where $g = \gcd(j, k)$.

The distances to pure noise for various surface code families are summarised in Table 3.1. We note that, for all code families, Y-distance exceeds X- or Z-distance, which is consistent with the increase in error threshold of surface codes with biased noise seen in Ref. [3.7]. Furthermore, we note that the Y-distance of square codes is $d_Y = O(\sqrt{n})$ while that of a co-prime codes is $d_Y = O(n)$, where n is the number of physical qubits in a $j\times k$ surface code. This feature contributes to the significant improvement in logical failure rate of co-prime codes over square codes with pure Y noise and Y-biased noise, see Sections 3.4 and 3.5.

Code family	d_X	d_Y	d_Z
square	j	$2j-1$	j
co-prime	j	jk	k
$\gcd(j,k) = g$	j	$(2g-1)(jk/g^2)$	k

Table 3.1: Distances to pure noise for $j\times k$ surface code families. (d_P refers to the distance to pure P noise, where $P \in \{X, Y, Z\}$.)

Logical operator count

We show that Y-type logical operators are comparatively rare. The number c_X of X-type logical operators is equal to the number of ways the logical $\bar{X}$ operator can be deformed by X-type stabiliser generators. The number of X-type stabiliser generators (i.e. vertices) on a $j\times k$ surface code is $j(k-1)$, and hence $c_X = 2^{j(k-1)}$; similarly $c_Z = 2^{(j-1)k}$. From the proof of Theorem 3.1, it is apparent that the g basis codewords of the Y-code correspond to a single logical operator and a full set of $g-1$ linearly independent Y-type stabilisers of a $j\times k$ surface code, where $g = \gcd(j,k)$. Therefore, the number of Y-type logical operators of a general $j\times k$ surface code is given by the following corollary.

Corollary 3.3 (Y-count)**.** *For a $j\times k$ surface code, the number of Y-type logical operators is*

$$c_Y = 2^{g-1}$$

where $g = \gcd(j,k)$. The number of Y-type stabilisers is also c_Y.

The counts of pure noise logical operators for various surface code families are summarised in Table 3.2. We note that, for all code families, the number of logical operators for pure Y noise is much lower than the number for pure X or Z noise, which is consistent with the increase in error threshold of surface codes with biased noise seen in Ref. [3.7]. Furthermore, we note that the number of Y-type logical operators for square codes is $c_Y = O(2^{\sqrt{n}})$ while for co-prime codes it is $c_Y = O(1)$, where n is the number of physical qubits in a $j\times k$ surface code. This feature contributes to the significant improvement in logical failure rate of co-prime codes over square codes with pure Y noise and Y-biased noise, see Sections 3.4 and 3.5.

Code family	c_X	c_Y	c_Z
square	2^{j^2-j}	2^{j-1}	2^{j^2-j}
co-prime	$2^{j(k-1)}$	1	$2^{(j-1)k}$
$\gcd(j,k)=g$	$2^{j(k-1)}$	2^{g-1}	$2^{(j-1)k}$

Table 3.2: Counts of pure noise logical operators for $j\times k$ surface code families. (c_P refers to the number of P-type logical operators, where $P\in\{X,Y,Z\}$.)

3.4 Advantage of co-prime surface codes with pure Y noise

In Section 3.3, we presented our analytical results for surface codes with pure Y noise, highlighting features that contribute to the ultra-high thresholds seen in Ref. [3.7] with Y-biased noise. Our analytical results also indicate that co-prime codes should achieve lower logical failure rates than square codes with pure Y noise.

Here we present our numerical investigation into the performance of surface codes with different g with pure Y noise. In particular, we present results for co-prime, g=4 and square surface code families, confirming the 50% error threshold. We also demonstrate a significant reduction in the logical failure rate for co-prime codes compared to square codes. Specifically, quadratically fewer physical qubits may be used to achieve a target logical failure rate by using co-prime codes.

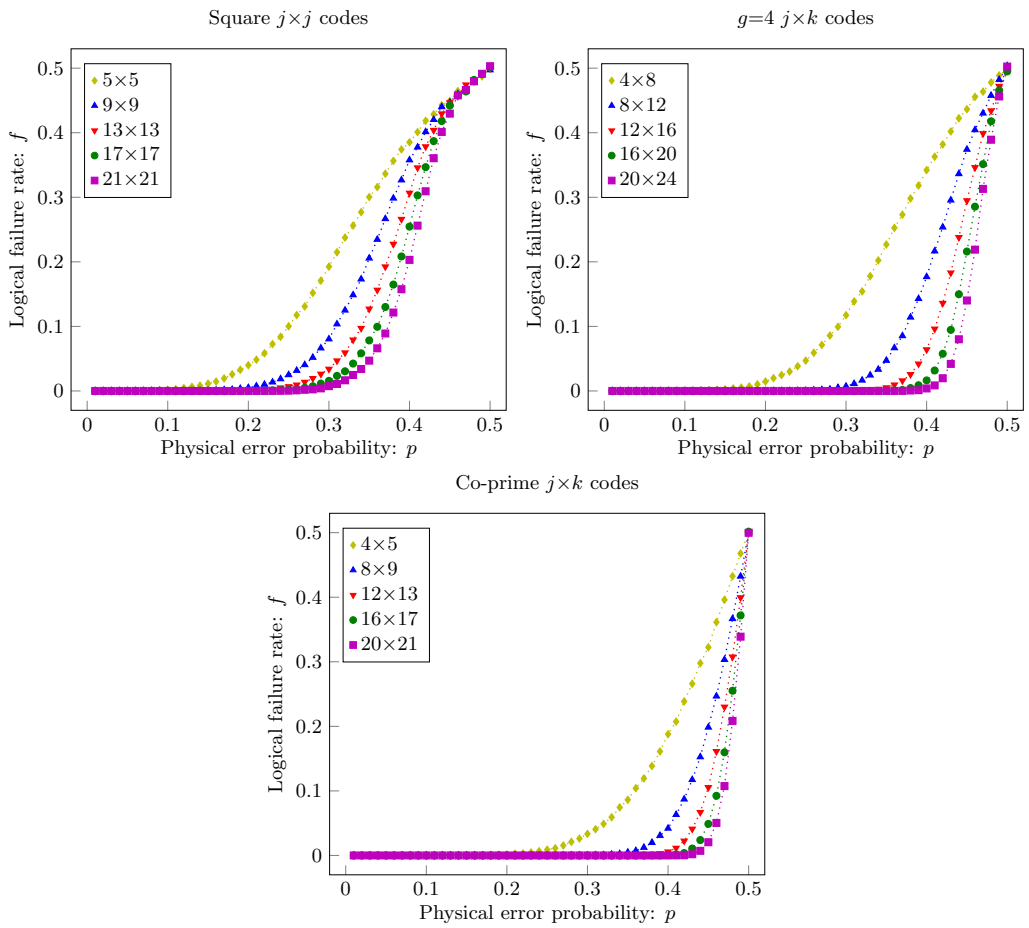


Figure 3.9: Logical failure rate f as a function of physical error probability p for $j\times k$ surface code families: square, g=4, and co-prime, where $g = \gcd(j, k)$, subject to pure Y noise. Data points are sample means over 60 000 runs using the Y-decoder described in Section 3.4.2. Dotted lines connect successive data points for a given code size.

We also define the exact optimal decoder for pure Y noise that we used in order to achieve optimal results and avoid the limitations of an approximate (see Section 3.3.3) or non-optimal (see Section 3.3.2) decoder.

3.4.1 Performance of surface codes with pure Y noise

We investigate the performance of surface codes with pure Y noise. Besides confirming the 50% threshold for the surface code, we demonstrate a significant reduction in logical failure rate for co-prime surface codes compared to square surface codes such that a target logical failure rate may be achieved with quadratically fewer physical qubits using co-prime codes in place of square codes. That is, we demonstrate that logical failure rate decays exponentially with Y-distance for square and co-prime codes but since, in accordance with Corollary 3.2, the Y-distance of these codes is $O(\sqrt{n})$ and $O(n)$ respectively, the logical failure rate decays quadratically faster with n for co-prime codes, where

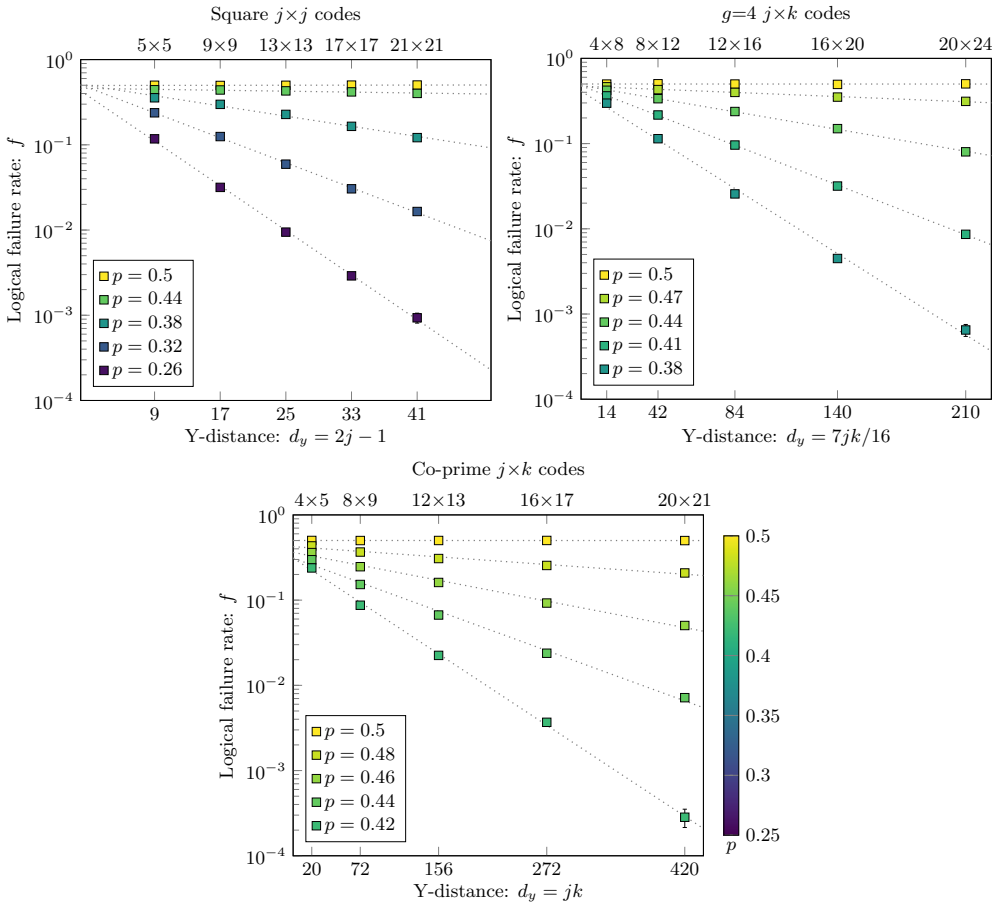


Figure 3.10: Exponential decay of the logical failure rate f with respect to code distance d_Y to pure Y noise in the regime of physical error probability p at and below the error threshold for $j\times k$ surface code families: square, g=4, and co-prime, where $g = \gcd(j, k)$, subject to pure Y noise. Data points are sample means over 60 000 runs using the Y-decoder described in Section 3.4.2. Dotted lines indicate least squares fit to data for a given p and error bars indicate 1 standard deviation.

n is the number of physical qubits.

In Fig. 3.9, we plot logical failure rate f as a function of physical failure rate p for $j\times k$ surface codes belonging to the following families: square, g=4, and co-prime codes, where $g = \gcd(j, k)$. For co-prime codes, we see clear evidence of an error threshold at $p_c = 50\%$, consistent with Corollary 3.1. For g=4 and square codes, the data is consistent with a threshold $p_c = 50\%$ but the evidence is less definitive. Within a code family, it is expected that smaller codes will perform worse than larger codes below threshold. However, comparing the performance of smaller co-prime codes to square codes, we see a significant improvement in logical failure rate across the full range of physical error probabilities. For example, the 20×21 co-prime code clearly outperforms the 21×21 square code. This can be seen as a qualitative demonstration of the effect of the features of surface codes with pure Y noise identified in Section 3.3. For g=4 codes, we see logical failure rates intermediate between square codes and co-prime codes, as expected.

In Fig. 3.10, we plot logical failure rate f as a function of code distance d_Y to pure Y noise at physical error probabilities p at and below the threshold $p_c = 50\%$ for surface codes belonging to the following families: square, g=4, and co-prime codes. For each code family, we see exponential decay of the logical failure rate $f \sim \exp(-\alpha d_Y)$, where α is a function of $(p_c - p)$, which is consistent with the threshold $p_c = 50\%$ predicted by Corollary 3.1. Considering $j\times k$ surface codes, according to Corollary 3.2, $d_Y = 2j - 1$ for square codes and $d_Y = jk$ for co-prime codes. That is, $d_Y = O(\sqrt{n})$ for square codes and $d_Y = O(n)$ for co-prime codes. As a result, based on the observed exponential decay, quadratically fewer physical qubits are required to achieve a target logical failure rate for a given physical error rate by using co-prime codes in place of square codes.

To investigate the performance of different families of surface codes with pure Y noise, we used the Y-decoder, defined in Section 3.4.2, to sample the logical failure rate across a full range of physical error probabilities for square, g=4 and co-prime codes. We used code sizes: {5×5, 9×9, 13×13, 17×17, 21×21} for square codes, {4×8, 8×12, 12×16, 16×20, 20×24} for g=4 codes, and {4×5, 8×9, 12×13, 16×17, 20×21} for co-prime codes, and 60 000 runs per code size and physical error probability. In our implementation of the Y-decoder, we used the Python language with SciPy and NumPy libraries [3.23,3.24] for fast linear algebra and the mathmp library [3.25] for arbitrary-precision floating-point arithmetic, enabling up to 50 decimal places of precision.

3.4.2 Exact optimal Y-decoder

Here we define the exact optimal decoder for pure Y noise that we used in our numerical simulations of Section 3.4.1. As mentioned in Section 3.3.2, it is possible to decode Y noise on the planar code by treating it as the concatenation of a cycle code and repetition codes and decoding level-by-level. However, while efficient, such a decoder is not necessarily optimal. Also, as mentioned in Section 3.3.3, the performance of the approximate maximum likelihood decoder [3.8] used in previous studies [3.7] was found to saturate with pure Y noise when tuned for efficiency. Here we explicitly define an exact maximum likelihood decoder for the surface code with pure Y noise that is efficient for $j\times k$ surface code families with small $\gcd(j, k)$, such as co-prime codes, and tractable for moderate-sized square codes.

Consider a surface code with n physical qubits and m independent vertex and plaquette stabiliser generators. In the case of pure Y noise, the only possible error configurations are Y-type Pauli operators, i.e. operators consisting only of Y and identity single-qubit Paulis. Let $\mathcal{P}_Y$ denote the

group of n-qubit Y-type Pauli operators, let $\mathcal{G}_Y$ denote the group of Y-type stabilisers, and define the centraliser of $\mathcal{G}_Y$ as $\mathcal{C}(\mathcal{G}_Y) = \{f \in \mathcal{P}_Y : fg = gf \ \forall \ g \in \mathcal{G}_Y\}$. If the result of measuring the vertex and plaquette stabiliser generators is given by syndrome $s \in \{0,1\}^m$ and $f_s \in \mathcal{P}_Y$ is some fixed Y-type Pauli operator with syndrome s then the set $f_s\mathcal{C}(\mathcal{G}_Y)$ of all Y-type Pauli operators with syndrome s is the disjoint union $f_s\mathcal{C}(\mathcal{G}_Y) = f_s\mathcal{G}_Y \cup f_s\bar{L}\mathcal{G}_Y$, where $\bar{L}$ is one of the single class of logical operators possible with pure Y noise.

For a given syndrome s and probability distribution π on the Pauli group, the maximum likelihood decoder for pure Y noise can be implemented by constructing a candidate Y-type recovery operator f_s consistent with s, and returning $\arg\max_f \pi(f\mathcal{G}_Y)$ where $f \in \{f_s, f_s\bar{L}\}$ and $\pi(f\mathcal{G}_Y) = \sum_{g\in\mathcal{G}_Y} \pi(fg)$.

On a $j\times k$ surface code, the size of the group of Y-type stabilisers is $|\mathcal{G}_Y| = c_Y = 2^{g-1}$ where $g = \gcd(j,k)$, see Corollary 3.3. Therefore, for surface codes with small g, such as co-prime codes, the Y-decoder is efficient, provided that a candidate Y-type recovery operator f_s, the group of Y-type stabilisers $\mathcal{G}_Y$ and logical operator $\bar{L}$ can be constructed efficiently. In the next two subsections, we describe these constructions.

Constructing Y-type stabilisers and logical operators

The construction of Y-type stabilisers and logical operators for a $j\times k$ code is illustrated in Fig. 3.11. A minimum-weight Y-type logical operator is constructed by applying Y operators along a path starting at the top-left corner of the lattice and descending diagonally to the right, reflecting at boundaries, until another corner is encountered from within the lattice. We construct Y-type stabilisers similarly, starting at each of the next $\gcd(j,k) - 1$ qubits of the top row and reflecting until the path cycles. Together these stabilisers generate the full group of 2^{g-1} Y-type stabilisers, and combine with the minimum-weight logical operator to give the 2^{g-1} Y-type logical operators of the $j\times k$ code.

Constructing candidate Y-type recovery operators

The construction of a candidate Y-type recovery operator, consistent with a given syndrome, depends on whether the code is co-prime, square or neither.

For co-prime codes, it is possible to construct an operator, consisting only of Y and identity single-qubit Paulis, that anticommutes with any single syndrome location. We refer to such operators as Y-type destabilisers. Given a complete syndrome, a candidate Y-type recovery operator is then simply constructed by taking the product of Y-type destabilisers for each syndrome location. One way to construct Y-type destabilisers for co-prime codes is illustrated in Fig. 3.12. For a given syndrome location, a partial recovery operator is constructed by applying seed Y operators along a path starting directly below the syndrome location and descending diagonally to the right until a boundary is encountered; further Y operators are applied along paths descending diagonally to the left of each of these seed Y operators, reflecting at boundaries, until the bottom boundary is encountered. The partial recovery operator then anticommutes with the original syndrome location and residual syndrome locations on the bottom boundary. A residual recovery operator is constructed for each residual syndrome location by applying Y operators along a line starting directly to the right of the syndrome location and ascending diagonally to the right, reflecting at boundaries, until a corner is encountered from within the lattice. The residual recovery operators then anticommute with the residual syndrome locations. The destabiliser for the original syndrome location is then simply the product of the partial

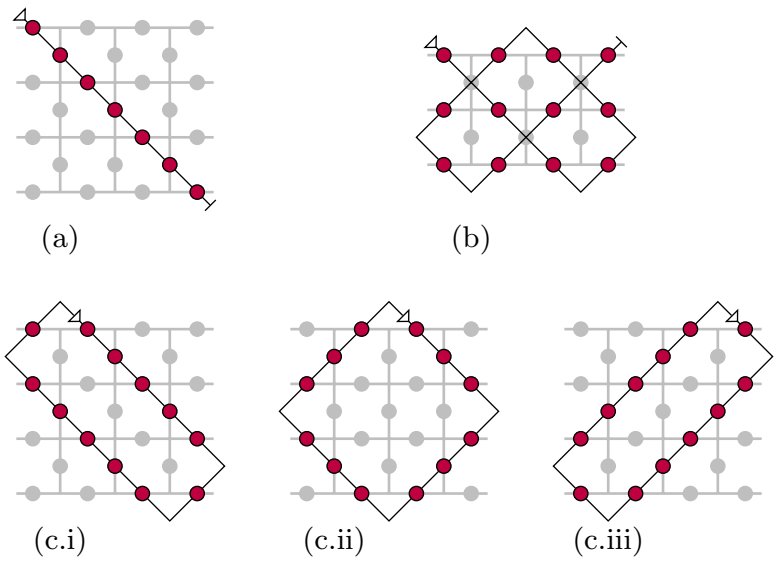


Figure 3.11: Examples of Y-type stabiliser and logical operator construction by applying Y operators along the indicated path until a corner is encountered or the path cycles. Minimum weight Y-type logical operators (a) and (b) for square 4×4 and co-prime 3×4 codes, respectively, are constructed by starting at the top-left qubit. Generators of the group of Y-type stabilisers (c) for the square 4×4 code are constructed by starting at each of the next $\gcd(j,k)-1=3$ qubits of the top row. (For co-prime codes, there are no Y-type stabilisers other than the identity.)

and residual recovery operators.

For square codes, Y-type destabilisers do not exist in general, and hence a different approach to constructing a candidate Y-type recovery operator must be adopted. Given a complete syndrome for a square code, a candidate Y-type recovery operator can be constructed by taking the product of partial recovery operators for each syndrome location, since the residual boundary syndrome locations cancel in the case of square codes, see Fig. 3.13.

For surface codes that are neither co-prime nor square, a candidate Y-type recovery operator is constructed by dividing the lattice into a co-prime region and square regions. Partial recovery operators are constructed for each region leaving residual syndrome locations only on plaquettes between regions. Residual syndrome locations can then be moved off the lattice using Y-type stabilisers on the square regions.

3.5 Advantage of co-prime surface codes with biased noise

In Section 3.4, we gave a demonstration that co-prime surface codes outperform square surface codes with pure Y noise in terms of logical failure rate. It is natural to ask if co-prime codes also outperform square codes with Y-biased noise, i.e., when X and Z errors may also occur. We demonstrate that a significant reduction in logical failure rate against biased noise can be achieved using a smaller co-prime $(j-1)\times j$ code compared to a $j\times j$ square code.

Our results are summarised in Fig. 3.1. With standard depolarising noise, i.e., $\eta = 0.5$, and

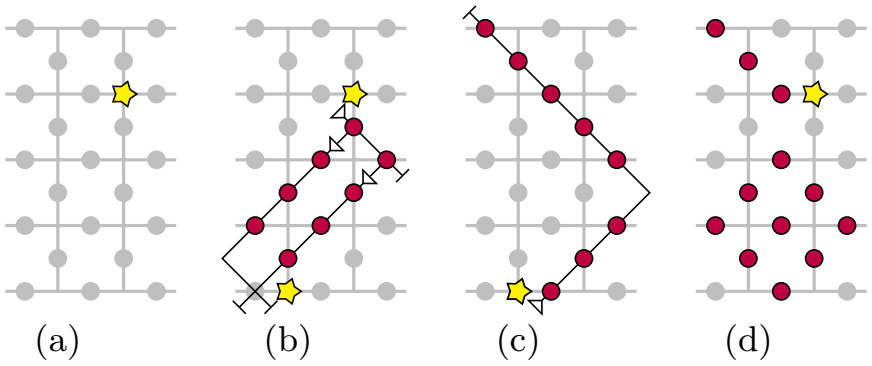


Figure 3.12: Example of Y-type destabiliser construction for a co-prime code. (a) Single syndrome location. (b) Partial recovery operator is constructed by applying Y operators, from below the syndrome location along a diagonal to any boundary, then from that diagonal along perpendicular diagonals, until the bottom boundary is encountered. (c) Residual recovery operators are constructed by applying Y operators, from right of each residual boundary syndrome location along a diagonal away, until a corner is encountered. (d) Destabiliser is product of partial and residual recovery operators.

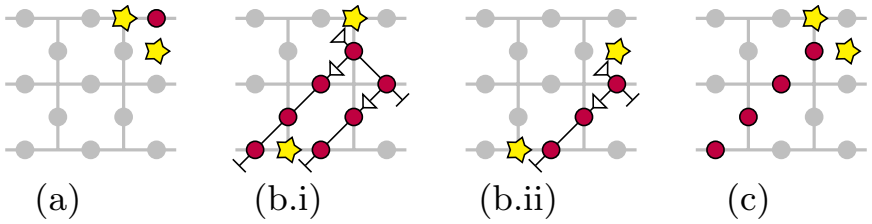


Figure 3.13: Example of candidate Y-type recovery operator construction for a square code using partial recovery operators. (a) Original error and complete syndrome. (b) Partial recovery operators with residual boundary syndrome locations. (c) Candidate recovery operator is the product of all partial recovery operators, since residual boundary syndrome locations cancel in the case of square codes.

with low bias, e.g. $\eta = 10$ (where Y errors are 10 times more likely than both X and Z), we see similar performance for the co-prime 7×8 and square 8×8 codes below the respective thresholds [3.7] of 18.9(3)% and 28.2(2)% for square codes with those biases. For bias as low as $\eta = 30$, we see a clear improvement in logical failure rate below threshold. In the limit of pure Y noise, we see the very large improvement, across the full range of physical error probabilities, that was already demonstrated in Section 3.4.1. Most interestingly, the improvement remains large through the intermediate bias regime, $\eta = 100$, over a wide range of physical error probabilities, indicating that the advantage of co-prime codes over square codes persists with modest noise biases.

The advantage of co-prime codes with biased noise can be explained in terms of the features of surface codes with Y noise identified in Section 3.3. The co-prime 7×8 code is only slightly more sensitive to X noise, having X-distance, $d_X = 7$, which is only slightly less than the X-distance, $d_X = 8$, of the square 8×8 code. The Z-distance, $d_Z = 8$, of both codes is identical. However, the co-prime code is much less sensitive to Y noise, having a much larger Y-distance, $d_Y = 56$, than the square code, $d_Y = 15$, and having only one Y-type logical operator, $c_Y = 1$, compared to many more

such operators, $c_Y = 2^7 = 128$, on the square code. Therefore, for sufficient bias, we expect co-prime $(j-1)\times j$ codes to outperform square $j\times j$ codes. Importantly, we also expect the relative advantage to increase with code size, as the slight increase in X-noise sensitivity becomes relatively smaller and the decrease in Y-noise sensitivity becomes relatively larger.

To compare the performance of co-prime and square codes with Y-biased noise, we sampled the logical failure rate across a full range of physical error probabilities for a co-prime 7×8 code and a square 8×8 code with noise biases $\eta \in \{0.5, 10, 30, 100, 300, \infty\}$. Sample means were taken over 60 000 runs per code, bias and physical error probability. To avoid any advantage being attributed to using a larger code, we chose the co-prime code to be smaller than the square code. Since the noise is biased, we could not use the Y-decoder for exact maximum likelihood decoding, so the natural choice was the BSV Decoder [3.8], which approximates maximum likelihood decoding. The BSV decoder is tuned via an approximation parameter χ, which defines the scale of correlations between syndrome bits. With pure Y noise, we observed that larger χ was required to achieve a performance close to that of the Y-decoder for co-prime codes than for square codes, which is consistent with the larger Y-distance of co-prime codes. For this reason we used relatively small codes with a large $\chi = 96$ to achieve good convergence and keep the computational requirements to a reasonable level.

3.6 Discussion

In this paper, we have described the structure of the surface code with pure Y noise and shown that this implies a 50% error threshold and a significant performance advantage in terms of logical failure rate with co-prime codes compared to square codes. Furthermore, we have provided numerics confirming our analytical results with pure Y noise and demonstrating the performance advantage of co-prime codes with Y-biased noise. It is important to note that our results apply equally to pure Z noise, i.e., dephasing noise, and the Z-biased noise prevalent in many quantum architectures, through the simple modification [3.7] of the surface code that exchanges the roles of Z and Y operators in stabiliser and logical operator definitions. We have, therefore, identified and characterised the features of surface codes that contribute to their ultra-high thresholds with Z-biased noise, seen in Ref. [3.7], and to the further improvement in logical failure rate with co-prime codes demonstrated in this paper.

In the limit of pure Y noise, we have shown that the surface code is equivalent to a concatenation of classical codes: a single top-level cycle code and a number of bottom-level repetition codes. We have shown that this implies the surface code with pure Y noise has a threshold of 50% and, for $j\times k$ surface codes with small $g = \gcd(j, k)$, the more effective repetition code dominates leading to a reduction in logical failure rate. In terms of logical operators, we have shown that Y-type logical operators are rarer and heavier than X- or Z-type equivalents, and co-prime surface codes, in particular, have only one Y-type logical operator and its weight is $O(n)$.

We have confirmed, numerically, the 50% error threshold of the surface code with pure Y noise, and demonstrated that co-prime codes with pure Y-noise significantly outperform similar-sized square codes in terms of logical failure rates such that a target logical failure rate may be achieved with quadratically fewer physical qubits using co-prime codes. Furthermore, we have demonstrated that this advantage persists with Y-biased noise. In particular, we found that a *smaller* co-prime code outperforms a square code, over a wide range of physical error probabilities, for biases as low as $\eta = 30$, where Y errors are 30 times more likely that X or Z errors. We argued that the relative

advantage of co-prime codes over square codes increases with code size.

Although our analytical results focus on features of the surface code with pure Y noise, it is interesting to put our observations of the performance of surface codes with biased noise in the context of other proposals to adapt quantum codes to biased noise [3.4,3.10–3.21]. Several proposals have been made for constructing asymmetric quantum codes for biased noise from classical codes [3.10–3.13] (see Ref. [3.12] for an extensive list of references) but of particular interest here are approaches that can be applied to topological codes. A significant increase in threshold with biased noise has been demonstrated by concatenating repetition codes at the bottom-level with another, possibly topological, code at the top-level [3.4, 3.14, 3.15]; interestingly, this mirrors the structure we have found to be inherent to the surface code. Performance improvements with biased noise have also been demonstrated by modifying the size and shape of stabilisers in Bacon-Shor codes [3.16–3.18] and surface / compass codes [3.19], by randomising the lattice of the toric code [3.20], or by concatenating a small Z-error detection code to the surface code [3.21]. These approaches are distinct from the use of co-prime codes (with the modification of Ref. [3.7]), which maintains the size and locality of surface code stabiliser generators, and so they could potentially be combined to yield further performance improvements.

Looking forward, the identified features of surface codes and the insights behind them suggest several interesting avenues of research. For the surface code, specifically, different geometries may be more robust to logical errors than co-prime codes in the high bias regime, where a few well-placed X and Z errors can combine with strings of Y errors to produce more common and lighter logical operators. Similarly, certain geometries of surface code used to encode multiple qubits [3.26] may or may not maintain the high performance of simple surface codes with biased noise. For topological codes, more generally, one can ask which codes exhibit an increase in performance with biased noise and what are the family traits of such codes; we have seen, for example, that the standard triangular 6.6.6 colour code does not exhibit an increase in performance. (Although this colour code is equivalent, in some sense, to a folded surface code [3.27], the mapping that relates the two does not preserve the biased noise model.)

Finally, although this paper focuses on features of surface codes with Y or Y-biased noise rather than the issue of efficient decoding, it may be possible to adapt the cycle code decoder or Y-decoder, defined herein, to implement a fast fault-tolerant decoder for biased noise. This would help to address the highly significant open question of whether the high performance of surface codes with biased noise can be preserved in the context of fault-tolerant quantum computing.

Acknowledgments

This work was supported by the Australian Research Council (ARC) via the Centre of Excellence for Engineered Quantum Systems (EQUS) project number CE170100009 and Discovery Project number DP170103073. SB acknowledges support from the IBM Research Frontiers Institute. Access to high-performance computing resources was provided by the National Computational Infrastructure (NCI), which is supported by the Australian Government, and by the Sydney Informatics Hub, which is funded by the University of Sydney.

3.A Color code thresholds with biased noise

We demonstrate that the threshold of the triangular 6.6.6 colour code [3.9] decreases when the noise is biased. This is in stark contrast to the surface code, which exhibits a significant increase in threshold with biased noise [3.7]. Our results are summarised in Fig. 3.14, in which, we contrast our new results for the colour code with those for the surface code reproduced from Ref. [3.7]. We note that the results for the surface code where originally presented in terms of Z-biased noise on a modified surface code but to simplify comparison with the colour code we restate the results in terms of Y-biased noise on the standard surface code. From statistical physics arguments, the optimal error threshold for the standard surface code with pure Z noise is estimated to be 10.93(2)% [3.3, 3.28], and with depolarising noise it is estimated to be 18.9(3)% [3.29]. The colour code has similar error thresholds [3.29, 3.30] to the surface code with pure Z noise and depolarising noise. Our results for the colour code, using an approximate maximum likelihood decoder, reveal a decrease in threshold with Y-biased noise; 18.7(1)% with standard ($\eta = 0.5$) depolarising noise, 13.3(1)% with bias $\eta = 3$, 11.4(1)% with bias $\eta = 10$, 10.6(1)% with bias $\eta = 100$, and 10.6(2)% in the limit of pure Y noise. In contrast, previous results [3.7] for the surface code revealed a significant increase in threshold with Y-biased noise; 18.7(1)% with standard ($\eta = 0.5$) depolarising noise, 22.3(1)% with bias $\eta = 3$, 28.2(2)% with bias $\eta = 10$, 40.3(8)% with bias $\eta = 100$, and 43.7(1)% (analytically shown to be 50% assuming exact optimal decoding, see Section 3.3.3) in the limit of pure Y noise. Our decoder implementation and numerics are described below. The features of surface codes that contribute to their exceptional performance with biased noise are discussed in the body of the paper.

Decoder.— In order to take account of correlations between X- and Z-type stabiliser syndromes, we implemented a tensor network approximate maximum likelihood decoder for triangular 6.6.6 colour codes following the same principles as the BSV decoder [3.8] used in Ref. [3.7] for surface codes.

Consider a colour code with n physical qubits and m independent stabiliser generators. Let $\mathcal{P}$ denote the group of n-qubit Pauli operators, let $\mathcal{G}$ denote the stabiliser group, and recall that the centraliser of $\mathcal{G}$ is given by $\mathcal{C}(\mathcal{G}) = \{f \in \mathcal{P} : fg = gf \ \forall \ g \in \mathcal{G}\}$. If the result of measuring the stabiliser generators is given by syndrome $s \in \{0,1\}^m$ and $f_s \in \mathcal{P}$ is some fixed Pauli operator with syndrome s then the set $f_s\mathcal{C}(\mathcal{G})$ of all Pauli operators with syndrome s is the disjoint union $f_s\mathcal{C}(\mathcal{G}) = f_s\mathcal{G} \cup f_s\bar{X}\mathcal{G} \cup f_s\bar{Y}\mathcal{G} \cup f_s\bar{Z}\mathcal{G}$, where $\bar{X}$, $\bar{Y}$ and $\bar{Z}$ are the logical operators on the encoded qubit.

For a given syndrome s and probability distribution π on the Pauli group, the maximum likelihood decoder can be implemented by constructing a candidate recovery operator f_s consistent with s, and returning $\arg\max_f \pi(f\mathcal{G})$ where $f \in \{f_s, f_s\bar{X}, f_s\bar{Y}, f_s\bar{Z}\}$ and $\pi(f\mathcal{G}) = \sum_{g\in\mathcal{G}} \pi(fg)$.

By analogy with the BSV decoder [3.8] for the surface code, we define a tensor network whose exact contraction yields the coset probability $\pi(f\mathcal{G})$ for the colour code. Fig. 3.15 (a-b) illustrates a distance 5 colour code, whereas (c) illustrates a tensor network with the same layout of qubits and stabilisers. Bonds have dimension 4. Stabiliser tensors are defined such that each element has value 1 if all indices are identical, and value 0 otherwise. Qubit tensors are defined such that each element has the single-qubit probability π of the product of the restriction of f to that qubit with the Paulis associated with bond indices where indices map to Paulis as $0 \mapsto I$, $1 \mapsto X$, $2 \mapsto Y$, $3 \mapsto Z$. In this way, all possible combinations of stabilisers are applied to f and the exact contraction of such a tensor network yields the coset probability $\pi(f\mathcal{G})$.

The exact contraction of the tensor network is inefficient with a runtime exponential in the number

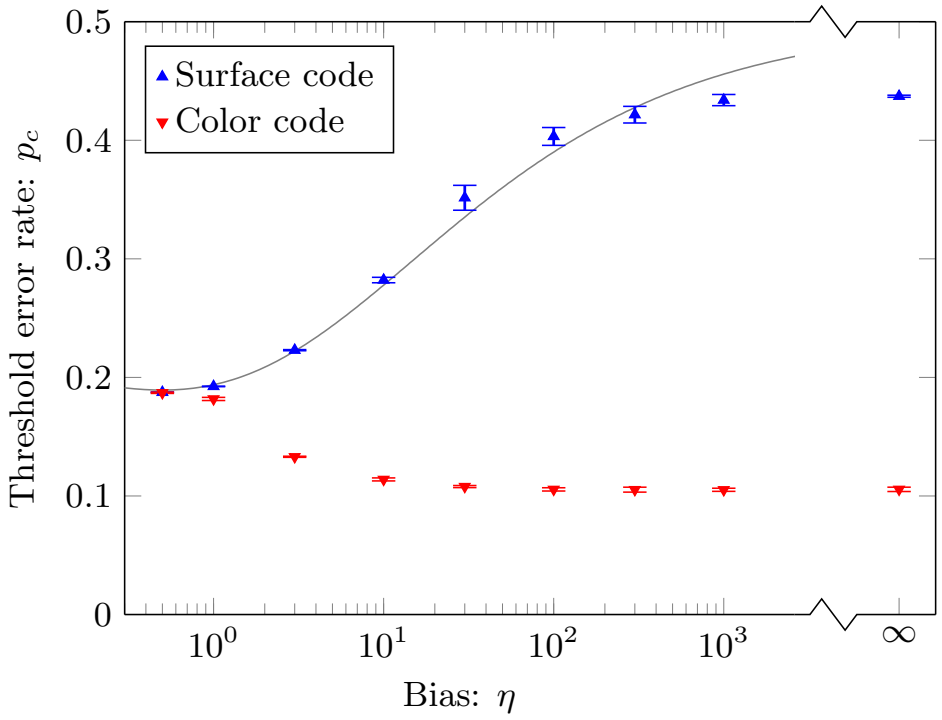


Figure 3.14: Threshold error rate p_c as a function of bias η. Red inverted triangles show threshold estimates for the triangular 6.6.6 colour code. For comparison, blue triangles show threshold estimates for the surface code (reproduced from Ref. [3.7]), whose saturation at high bias is due to decoder approximation. Error bars indicate 1 standard deviation relative to the fitting procedure. The points at the largest bias value correspond to infinite bias, i.e. only Y errors. The gray line is the zero-rate hashing bound for the associated Pauli error channel.

of qubits n. However, by merging neighbouring qubit tensors in pairs, the tensor network can be transformed into a square lattice, see Fig. 3.15 (d), so that techniques, used in the BSV decoder [3.8], can be applied to efficiently approximate the coset probability. The approximation is controlled by a parameter χ which defines the maximum bond dimension retained as the tensor network is contracted. We refer the reader to [3.8] for full details of the approximate contraction algorithm. We found that the performance of the decoder converged well for $\chi = 36$ across all noise biases, see Numerics below.

Numerics.— We followed the general approach taken in Ref. [3.7]; we give a brief summary here and refer the reader to Ref. [3.7] for full details. We used triangular 6.6.6 colour codes of distances $d = 7, 11, 15, 19$. We estimated the threshold for biases $\eta = 0.5, 1, 3, 10, 30, 100, 300, 1000, \infty$, where $\eta = p_Y/(p_X + p_Z)$ and $p_X = p_Z$, such that $\eta = 0.5$ corresponds to standard depolarising noise and $\eta = \infty$ corresponds to pure Y noise (see Section 3.2). We approximated maximum likelihood decoding using the decoder, described above, with approximation parameter $\chi = 36$. The decoder converged well (generally better than in Ref. [3.7]) across the full range of biases with weakest convergence in the low bias regime, see Fig. 3.16. We ran 30 000 simulations per code distance and physical error probability. As in Ref. [3.7], we used the critical exponent method of Ref. [3.31] to obtain threshold

estimates with jackknife resampling over the code distances to determine error bounds.

Bibliography

3.1 S. B. Bravyi and A. Y. Kitaev, (1998), quant-ph/9811052 .

3.2 B. M. Terhal, Rev. Mod. Phys. **87**, 307 (2015), arXiv:1302.3428 .

3.3 E. Dennis, A. Kitaev, A. Landahl, and J. Preskill, J. Math. Phys. **43**, 4452 (2002), quant-ph/0110143 .

3.4 P. Aliferis, F. Brito, D. P. DiVincenzo, J. Preskill, M. Steffen, and B. M. Terhal, New J. Phys. **11**, 013061 (2009), arXiv:0806.0383 .

3.5 M. D. Shulman, O. E. Dial, S. P. Harvey, H. Bluhm, V. Umansky, and A. Yacoby, Science **336**, 202 (2012), arXiv:1202.1828 .

3.6 D. Nigg, M. Müller, E. A. Martinez, P. Schindler, M. Hennrich, T. Monz, M. A. Martin-Delgado, and R. Blatt, Science **345**, 302 (2014), arXiv:1403.5426 .

3.7 D. K. Tuckett, S. D. Bartlett, and S. T. Flammia, Phys. Rev. Lett. **120**, 050505 (2018), arXiv:1708.08474 .

3.8 S. Bravyi, M. Suchara, and A. Vargo, Phys. Rev. A **90**, 032326 (2014), arXiv:1405.4883 .

3.9 H. Bombin and M. A. Martin-Delgado, Phys. Rev. Lett. **97**, 180501 (2006), quant-ph/0605138 .

3.10 L. Ioffe and M. Mézard, Phys. Rev. A **75**, 032345 (2007), quant-ph/0606107 .

3.11 P. K. Sarvepalli, A. Klappenecker, and M. Rötteler, Proc. R. Soc. A **465**, 1645 (2009).

3.12 G. G. La Guardia, Int. J. Theor. Phys. **53**, 2312 (2014).

3.13 A. Robertson, C. Granade, S. D. Bartlett, and S. T. Flammia, Phys. Rev. Applied **8**, 064004 (2017), arXiv:1703.08179 .

3.14 P. Aliferis and J. Preskill, Phys. Rev. A **78**, 052331 (2008), arXiv:0710.1301 .

3.15 A. M. Stephens, W. J. Munro, and K. Nemoto, Phys. Rev. A **88**, 060301 (2013), arXiv:1308.4776 .

3.16 A. M. Stephens, Z. W. E. Evans, S. J. Devitt, and L. C. L. Hollenberg, Phys. Rev. A **77**, 062335 (2008), arXiv:0708.3969 .

3.17 J. Napp and J. Preskill, Quantum Inf. Comput. **13**, 0490 (2013), arXiv:1209.0794 .

3.18 P. Brooks and J. Preskill, Phys. Rev. A **87**, 032310 (2013), arXiv:1211.1400 .

3.19 M. Li, D. Miller, M. Newman, Y. Wu, and K. R. Brown, (2018), arXiv:1809.01193 .

3.20 B. Röthlisberger, J. R. Wootton, R. M. Heath, J. K. Pachos, and D. Loss, Phys. Rev. A **85**, 022313 (2012), arXiv:1112.1613 .

3.21 X. Xu, Q. Zhao, X. Yuan, and S. C. Benjamin, (2018), arXiv:1812.01505 .

3.22 P. Webster, S. D. Bartlett, and D. Poulin, Phys. Rev. A **92**, 062309 (2015), arXiv:1509.05032 .

3.23 E. Jones, T. Oliphant, P. Peterson, *et al.*, "SciPy: Open source scientific tools for Python," (2001), `https://www.scipy.org/`.

3.24 T. E. Oliphant, *A guide to NumPy*, Vol. 1 (2006) `https://www.numpy.org/`.

3.25 F. Johansson *et al.*, *mpmath: a Python library for arbitrary-precision floating-point arithmetic (version 1.0)* (2017), `http://mpmath.org/`.

3.26 N. Delfosse, P. Iyer, and D. Poulin, (2016), arXiv:1606.07116 .

3.27 A. Kubica, B. Yoshida, and F. Pastawski, New J. Phys. **17**, 083026 (2015), arXiv:1503.02065 .

3.28 F. Merz and J. T. Chalker, Phys. Rev. B **65**, 054425 (2002), cond-mat/0106023 .

3.29 H. Bombin, R. S. Andrist, M. Ohzeki, H. G. Katzgraber, and M. A. Martin-Delgado, Phys. Rev. X **2**, 021004 (2012), arXiv:1202.1852 .

3.30 H. G. Katzgraber, H. Bombin, and M. A. Martin-Delgado, Phys. Rev. Lett. **103**, 090501 (2009), arXiv:0902.4845 .

3.31 C. Wang, J. Harrington, and J. Preskill, Annals of Physics **303**, 31 (2003), quant-ph/0207088 .

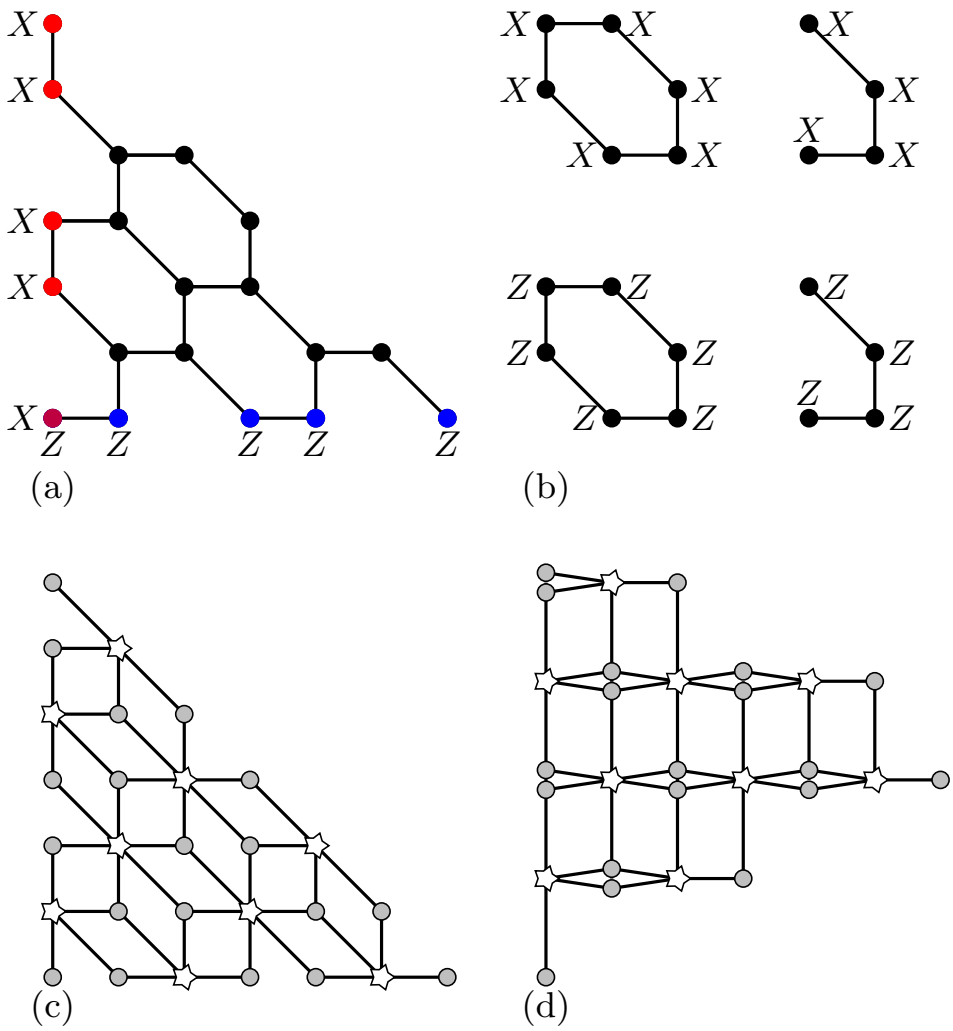


Figure 3.15: (a) Distance 5 triangular 6.6.6 colour code with logical operators given by a product of Z along the bottom edge and a product of X along the left edge. (b) Color code stabilisers. (c) Tensor network corresponding to coset probability of distance 5 colour code; gray discs represent qubit tensors; white stars represent stabiliser tensors; lines represent bonds. (d) Equivalent tensor network as a square lattice.

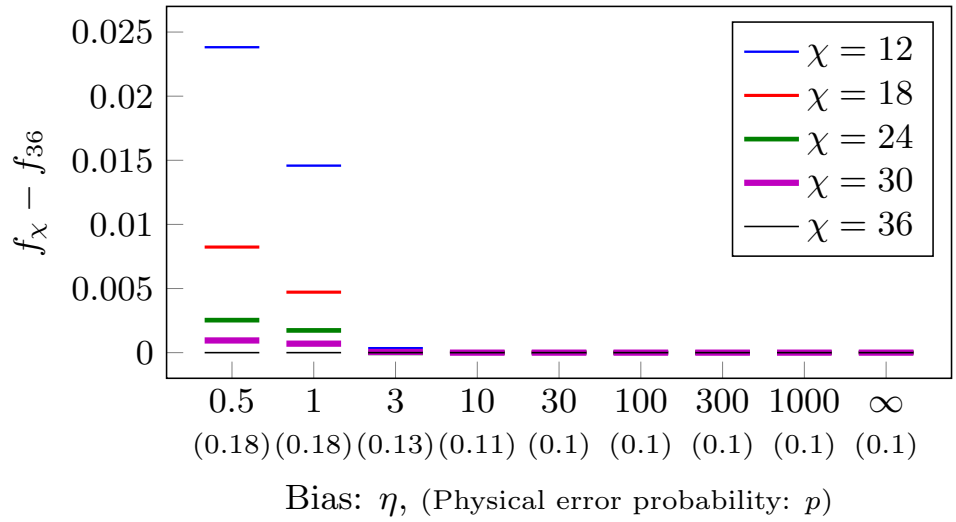


Figure 3.16: Decoder convergence, represented by shifted logical failure rate $f_\chi - f_{36}$, as a function of χ near the threshold physical error probability p for distance $d = 19$ triangular 6.6.6 colour codes. Each data point corresponds to 60 000 runs with identical errors generated across all χ for a given bias.

Chapter 4

Approximate symmetries of Hamiltonians

Christopher T. Chubb[1] and Steven T. Flammia[1,2]
[1] *Centre for Engineered Quantum Systems, University of Sydney, Sydney, NSW, Australia*
[2] *Center for Theoretical Physics, Massachusetts Institute of Technology, Cambridge, MA, USA*

Journal of Mathematical Physics **58**, 082202 (2017)
Doi: 10.1063/1.4998921
ArXiv: 1608.02600

Abstract

We explore the relationship between approximate symmetries of a gapped Hamiltonian and the structure of its ground space. We start by considering approximate symmetry operators, defined as unitary operators whose commutators with the Hamiltonian have norms that are sufficiently small. We show that approximate symmetry operators can be restricted to the ground space while approximately preserving certain mutual commutation relations. We generalise the Stone-von Neumann theorem to matrices that *approximately* satisfy the canonical (Heisenberg-Weyl-type) commutation relations, and use this to show that approximate symmetry operators can certify the degeneracy of the ground space even though they only approximately form a group. Importantly, the notions of "approximate" and "small" are all independent of the dimension of the ambient Hilbert space, and depend only on the degeneracy in the ground space. Our analysis additionally holds for any gapped band of sufficiently small width in the excited spectrum of the Hamiltonian, and we discuss applications of these ideas to topological quantum phases of matter and topological quantum error correcting codes. Finally, in our analysis we also provide an exponential improvement upon bounds concerning the existence of shared approximate eigenvectors of approximately commuting operators under an added normality constraint, which may be of independent interest.

4.1 Introduction

Given a quantum system described by a Hamiltonian H, a symmetry is simply an operator that commutes with H. The symmetry can be block diagonalised with respect to the energy eigenspaces, and

so the degeneracy within these blocks is constrained by the symmetry. In a system that possesses exact symmetries, a sufficiently weak perturbation will preserve the number of states of any band gapped away from the rest of the spectrum, but the symmetries will generally become only approximate.

In this work we consider a natural converse to this: suppose we know that a system has some approximate symmetries and a gapped band, such as the ground space band. Can we "unperturb" the symmetries into exact symmetries within the given band? Can we also use the approximate group structure of the approximate symmetries to count the degeneracy within the band? We answer these questions in the affirmative, giving quantitive bounds on when such a procedure can be performed, and thus when such approximate symmetries can be used as certificates of ground space degeneracy.

A related area of mathematical research with a long and rich history is the relationship between the properties of approximately and exactly commuting matrices. An exemplary problem which dates back as far as the 1950s [4.2–4.6] is whether a pair of approximately commuting matrices lie near an exactly commuting pair, i.e. whether there exists a dimension independent $\delta > 0$ for each $\epsilon > 0$ such that for all H, S with $\|H\|, \|S\| \leq 1$,

$$\|[H, S]\| \leq \delta \qquad \Longrightarrow \qquad \exists \tilde{H}, \tilde{S} : \left[\tilde{H}, \tilde{S}\right] = 0, \text{ where } \left\|H - \tilde{H}\right\|, \left\|S - \tilde{S}\right\| \leq \epsilon,$$

where here and throughout the norm $\|\cdot\|$ is the operator norm. Interpreting H as the Hamiltonian and S as a symmetry, this problem can be interpreted as whether approximate symmetries are necessarily near exact symmetries of a perturbed system. It has been shown that just such a theorem holds if all matrices are Hermitian [4.7–4.10]. A physical consequence of this is that a pair approximately commuting observables can be approximately simultaneously measured [4.10].

For unitary matrices the above is however known to be generally false [4.11]. This is due to a K-theoretic obstruction [4.12–4.14], though it is true if this obstruction vanishes [4.7, 4.8, 4.15], or under the assumption of a spectral gap [4.16]. Imposing a form of self-duality analogous to time-reversal symmetry the relevant K-theoretic obstruction reduces to the spin Chern number of a fermionic system [4.17], highlighting a link between the fields of topologically ordered quantum systems [4.18] and approximately commuting matrices.

Here we will consider Hamiltonians H with multiple non-commuting approximate symmetries, and establish a connection to the ground space degeneracy. Ground space degeneracy is a property of a quantum system that plays a special role in several important applications, such as quantum coding theory and the study of phases of matter. Quantum codes, especially those encoded into ground spaces of local Hamiltonians, are the leading candidates for thermally stable quantum memories [4.19, 4.20]; in these models approximate symmetries constitute approximate logical operators and the ground space degeneracy corresponds to the code size. In the context of condensed matter systems, the link between symmetries and degeneracies plays an important role both in classical symmetry-breaking phases [4.21], and also in exotic quantum phases, such as those exhibited by topologically ordered models [4.18]. Unfortunately, determining the ground space degeneracy of (finite) systems is generally #P-complete, even for gapped bands [4.22]. However, if we restrict to more structured examples such as 1D-local spin systems, then ground spaces can in fact be efficiently approximated [4.23, 4.24]. Our results show that when structure is present in the form of certain mutual commutation relations, one can obtain certifiable bounds on the degeneracy of a ground space by only knowing bounds on these relations. We go into more detail about these two applications in Section 4.5.

The form of non-commutation we will consider will involve *twisted* commutation relations.

Definition 4.1 (Twisted commutator)**.** *For $\alpha \in [0, 1)$, the* twisted commutator *is defined as*

$$[X, Y]_\alpha := XY - e^{2i\pi\alpha} YX.$$

We will refer to α as the twisting parameter*, and for some unitarily invariant norm $\|\cdot\|$ we will refer to $\|[\cdot, \cdot]_\alpha\|$ as the* twisted commutation value*. When considering a pair of operators in tandem such that each has a small twisted commutation value we will refer to it as a* twisted pair*.*

We note that the $\alpha = 0$ and $\alpha = 1/2$ cases correspond to the commutator and anti-commutator respectively.

Commuting operators exist in all dimensions, finite or infinite. Twisted commuting operators on finite-dimensional spaces, however, only exist in certain dimensions depending on the twisting parameter, e.g. no $\alpha \neq 0$ twisted commutator can non-trivially vanish in a one-dimensional space. For general operators, twisted commuting operators were studied in some detail in Ref. [4.25]. If we restrict to unitary operators however, the Stone-von Neumann Theorem[1] [4.26, 4.27] classifies the dimensions in which twisted commutation can occur.

Theorem 4.1 (Finite-dimensional Stone-von Neumann theorem)**.** *Given $\alpha = p/q$ with p, q coprime, then unitary operators X and Y which exactly twisted commute as*

$$[X, Y]_\alpha = 0$$

only exist in dimensions which are multiples of q.

In this paper we will generalise this connection into the regime of *approximately* twisted-commuting operators. Properties of both approximate commuting, and approximately twisted-commuting operators are reviewed in Ref. [4.28]. The rigidity of algebraic structures to small perturbations in the commutation relations that define them has been studied in several other settings, such as the soft torus [4.29, 4.30] and approximate representations of groups [4.31–4.33].

Suppose we have a physical system with a self-adjoint Hamiltonian H, acting on a possibly infinite-dimensional Hilbert space. Let Π be the orthogonal projector onto the finite-dimensional ground space, and $\bar{\Pi} := I - \Pi$. For simplicity, take the ground state energy of H to be zero, such that $\Pi H = 0$. As well as this, we will assume that the excited states are gapped away from the ground space, such that they all have an energy at least Δ, i.e. $H \geq \Delta\bar{\Pi}$. For such a system there exist two notions of symmetry we will discuss.

Definition 4.2 (Symmetry)**.** *We define a* ground symmetry *as an operator U that commutes with the ground space projector*

$$[U, \Pi] = 0,$$

and acts unitarily on the ground space $\Pi U^\dagger U \Pi = \Pi U U^\dagger \Pi = \Pi$. Moreover, we refer to a unitary as an ϵ-approximate symmetry *if it approximately commutes with the Hamiltonian with respect to a given unitarily invariant norm*

$$\|[U, H]\| \leq \epsilon.$$

[1] As usually stated, the Stone-von Neumann theorem is much more general than Theorem 4.1. We will only be concerned with twisted commutation in finite-dimensional spaces, and unconcerned with uniqueness, so this form will suffice for our purposes.

Here we use $\|\!|\cdot|\!\|$ to denote any unitarily invariant norm.

The error thresholds we are going to consider will depend on the spectral gap Δ of the system in question. One way to improve the scaling with the gap would be to consider symmetries defined not by commutation with the Hamiltonian, but by commutation with functions of the Hamiltonian. For example we could consider commutation with an (unnormalised) Gibbs state

$$\left\|\!\left|\left[U, e^{-\beta H}\right]\right|\!\right\| \leq \epsilon.$$

Such a symmetry can be seen to be an ϵ-approximate symmetry of $H' = I - e^{-\beta H}$, which shares a ground space with H and has a gap of $1 - e^{-\beta\Delta}$. If we have some control over the temperature, such as in Monte Carlo simulations, then this gives a tradeoff we can use to improve the gap scaling. If for example we set $\beta = \ln(2)/\Delta$, then we get a fixed gap of $1/2$. A similar analysis could be performed with any function of H which leaves the relevant band gapped.

4.1.1 Results

The main goal of this paper will be to establish a connection between twisted commuting symmetries and the ground space dimension, even when the relevant commutation relations are only approximate. A key feature of our bounds is that they can be expressed entirely in terms of the Hamiltonian, and do not require objects such as the ground space projector, which can often be prohibitively difficult to calculate, represent, or perform calculations with. Without access to the ground space projector, whether or not a unitary is a ground symmetry cannot be directly verified.

In Section 4.2 we will explore the relationship between approximate and ground symmetries, showing that an approximate symmetry is always near a ground symmetry. Extending this to the case of multiple symmetries, we will see that approximate symmetries can be restricted to the ground space with low distortion, implying the existence of unitaries on the ground space with certain twisted commutation relations. In showing these results, we will make repeated use of the following function and note some simple bounds on it,

$$f : [0,1] \to [0,1]\,, \qquad f(x) := 1 - \sqrt{1-x}\,, \qquad \tfrac{x}{2} \leq f(x) \leq x\,.$$

Then our first main result is the following.

Theorem 4.2 (Restriction to the ground space)**.** *For two ϵ-approximate symmetries U and V which approximately twisted commute*

$$\|\!|[U,V]_\alpha|\!\| \leq \delta,$$

then if $\xi := \epsilon/\Delta < 1$ there exist unitaries u and v acting on the ground space which also approximately twisted commute as

$$\|\!|[u,v]_\alpha|\!\| \leq \delta + 2\xi^2 + 4f(\xi^2)\,.$$

Rather importantly, we note that the above theorem holds independent of the ground space dimension. This will allow us to use approximate symmetries alone as witnesses of ground space degeneracy, circumventing the need for direct access to the ground space, which is often inaccessible in non-exactly solvable models.

Note that for simplicity we will henceforth take the band in consideration to be an exactly degenerate ground space. We will see however that our proof will rely not on the bound $H \geq \Delta\bar{\Pi}$, but on its relaxation $H^2 \geq \Delta^2\bar{\Pi}$, meaning that the band could be anywhere in the spectrum, so long as it is gapped on both sides by at least Δ. Furthermore we can take $w := \|H\Pi\| \geq 0$ when our band has a potentially non-zero width. By considering the new Hamiltonian $H' := H - H\Pi$, we get that our restricted result holds for more general bands once the necessary changes have been made.

Corollary 4.1 (Restriction to a general band)**.** *If there are two ϵ-approximate symmetries U and V which approximately twisted commute*

$$\|[U, V]_\alpha\| \leq \delta,$$

then if $\xi' := (\epsilon + w)/\Delta < 1$ there exists unitaries u and v acting on band of gap Δ and width w which also approximately twisted commute

$$\|[u, v]_\alpha\| \leq \delta + 2\xi'^2 + 4f(\xi'^2)\,.$$

Now that we have restricted our symmetries down to the ground space, by studying the relationship between dimensionality and approximate twisted commutation, we can hope to use these twisted symmetries as witnesses of ground space degeneracy. As above, we will henceforth adhere to the convention of upper case letters denoting operators which act on the system as a whole, and lower case operators which only act on the ground space.

In Section 4.3 we start by giving a proof of Theorem 4.1, and consider generalising this argument to the case of *approximately* twisted commuting operators. We consider a twisted pair of unitaries, and construct states which can be used to lower bound the number of eigenvalues these operators possess. By doing so we will show that if these operators have a sufficiently small twisted commutation value in the operator norm, then a lower bound on their degeneracy can be inferred.

Theorem 4.3. *If u and v are unitaries such that for some $d \in \mathbb{N}$*

$$\left\|[u, v]_{1/d}\right\| < \frac{2}{d-1}\left[1 - \cos \pi/d\right],$$

then the dimension of each operator is at least d.

While we do not have a closed form bound on the twisted commutation value required to certify other dimensions ($d \neq 1/\alpha$), in Section 4.B we discuss an algorithm to determine which degeneracies are certified by twisted pairs of given parameters. Using this we will plot the dimension that can be certified as a function of both the twisting parameter and the corresponding twisted commutation value.

In Section 4.A we strengthen existing results on shared approximate eigenvectors for approximately commuting operators when a normality condition is introduced, exponentially improving the dimension dependence of the bounds relative to known results [4.34]. Using this, in Section 4.3.3 we consider extending this procedure to the case of two pairs of twisted commuting unitaries. Here we will once again construct a set of ground states, showing that for sufficient parameters that they are linearly independent. Using this we can obtain a similar dimensionality lower bound.

Theorem 4.4. *If u_1, u_2, v_1 and v_2 are unitaries such that they satisfy the commutation relations*

$$\|[u_1, u_2]\| \leq \gamma \qquad \|[u_1, v_2]\| \leq \delta \qquad \|[u_2, v_1]\| \leq \delta$$

and twisted commutation relations

$$\left\|[u_1, v_1]_{1/d_1}\right\| \leq \delta \qquad \left\|[u_2, v_2]_{1/d_2}\right\| \leq \delta$$

with $d_1 \leq d_2$ and

$$\sqrt{\gamma} d_1 d_2 + (d_1 + d_2)\delta < \frac{\sin^2(\pi/2d_1)}{(d_1 d_2 - 1)^2},$$

then the dimension of each operator is at least $d_1 d_2$.

In Section 4.4 we provide a more comprehensive analysis for the case of a single twisted pair. Leveraging results from spectral perturbation theory, we find an explicit closed form for the minimum twisted commutation value for a class of norms known as the (p, k)-Schatten-Ky Fan norms. These are defined as the p-norm of the largest k singular values, or more formally as

$$\|X\|_{(p,k)} := \sup_A \left\{ (\operatorname{Tr} |AX|^p)^{1/p} \;\middle|\; \|A\| \leq 1,\, \operatorname{rank}(A) \leq k \right\}.$$

For a g-dimensional operator, the special case $k = g$ reduces to the Schatten p-norm, and the case $p = 1$ reduces to the Ky Fan k-norm. In particular, the $p = \infty$ and $(p, k) = (2, g)$ special cases reduce to the operator and Frobenius norms respectively.

Theorem 4.5 (Minimum twisted commutation value)**.** *Suppose that u and v are g-dimensional unitaries, then for any $p \geq 2$ the twisted commutator is lower bounded*

$$\left\|[u, v]_\alpha\right\|_{(p,k)} \geq 2k^{1/p} \sin\left(\pi \left| \frac{\lfloor g\alpha \rceil - g\alpha}{g} \right|\right),$$

where $\|\cdot\|_{(p,k)}$ is the (p, k)-Schatten-Ky Fan norm. Moreover this bound is tight, in that sense that there exist families of g-dimensional unitaries which saturate the above bounds and only depend on $\lfloor g\alpha \rceil$, the nearest integer to $g\alpha$.

For a given twisted pair, all dimensions for which the twisted commutation value falls below this minimum can therefore be ruled out as valid dimensions. As this bound is not monotonic as a function of g, it not only provides a lower bound, but a full classification of which dimensions are disallowed.

After giving proofs of the main results outlined above in Sections 4.2, 4.3, and 4.4, we turn to broader discussion and applications of these ideas. Section 4.5 is devoted to discussion of future directions for this work that add the additional constraint that the Hamiltonian is *local*, and we discuss the relationship to the notions of topological order and topological quantum codes. In particular we show how recent numerical methods for studying quantum many-body systems [4.35] could leverage the bounds presented here to provide certificates of the topological degeneracy of certain quantum systems.

4.2 Restriction to the ground space

In this section we will make precise the notion that approximate symmetries can be utilised as proxies of ground symmetries. We first establish a relationship between approximate symmetries and the ground symmetries that they imply. Then we consider operators with approximate twisted commutation relations, and we show that these can also be restricted faithfully to the ground space with low distortion.

Constructing a ground symmetry from an approximate symmetry will come in two steps. First we will pinch the symmetry U with respect to Π, giving an operator P for which $[P, \Pi] = 0$. While this will render P no longer unitary, we will see that its action upon the ground space will still be *approximately* unitary. Using this we will construct a nearby operator $\tilde{U}$ that retains commutation with the ground space projector, and acts unitarily on the ground space, thus constituting a ground symmetry.

We will start by showing that the off-diagonal blocks of an approximate symmetry are small, and then follow by showing that its action on the ground space is approximately unitary.

Lemma 4.2 (Small off-diagonal blocks)**.** *If U is an ϵ-approximate symmetry, then off-diagonal blocks of U with respect to Π have bounded norms, in particular $\left\|\bar{\Pi}U\Pi + \Pi U\bar{\Pi}\right\| \leq \epsilon/\Delta$. For Hamiltonians of the form $H = \Delta\bar{\Pi}$, this inequality is tight.*

Proof. We start by noting that $|A|^2 \geq |B|^2$ implies $|AX|^2 \geq |BX|^2$ for any X, where $|M| := \sqrt{M^\dagger M}$. Taking X to be finite-rank, we have from Weyl's inequalities [4.1] that the singular values of AX majorise those of BX. Unitarily invariant norms[2] act as symmetric gauge functions on finite-rank operators [4.36–4.38], which implies from Refs. [4.36, Prop.IV.1.1, Thm.IV.2.2] that $\|AX\| \geq \|BX\|$—a similar argument for the adjoint also gives $\|XA\| \geq \|XB\|$. Because H has a gapped band with projector Π, we have that $H^2 \geq \Delta^2\bar{\Pi}$. Using this, we can bound the off-diagonal blocks in terms of the commutator

$$\begin{aligned}
\Delta\left\|\bar{\Pi}U\Pi + \Pi U\bar{\Pi}\right\| &= \left\|\left(\bar{\Pi}\Delta + \Pi\right)\left(\bar{\Pi}U\Pi + \Pi U\bar{\Pi}\right)\left(\Delta\bar{\Pi} + \Pi\right)\right\| \\
&\leq \left\|\left(\bar{\Pi}H + \Pi\right)\left(\bar{\Pi}U\Pi + \Pi U\bar{\Pi}\right)\left(H\bar{\Pi} + \Pi\right)\right\| \\
&= \left\|\bar{\Pi}HU\Pi + \Pi UH\bar{\Pi}\right\|,
\end{aligned}$$

where the inequality follows from the aforementioned monotonicity property. Now using the unitary invariance of the norm (since $\bar{\Pi} - \Pi$ is unitary), we find that

$$\begin{aligned}
\left\|\bar{\Pi}HU\Pi + \Pi UH\bar{\Pi}\right\| &= \left\|\bar{\Pi}HU\Pi - \Pi UH\bar{\Pi}\right\| \\
&= \left\|\bar{\Pi}\left[H, U\right]\Pi + \Pi\left[H, U\right]\bar{\Pi}\right\| \\
&\leq \|[H, U]\| \leq \epsilon,
\end{aligned}$$

where the second equality makes use of $H\Pi = 0$ and the first inequality is the pinching inequality.

With regard to tightness, if we take $H = \Delta\bar{\Pi}$ then we can see $[H, U]$ has no on-diagonal blocks, and therefore $\Delta\left(\bar{\Pi}U\Pi - \Pi U\bar{\Pi}\right) = [H, U]$. Taking norms of both sides of this equation give $\Delta\left\|\bar{\Pi}U\Pi + \Pi U\bar{\Pi}\right\| = \|[H, U]\|$, meaning that $\left\|\bar{\Pi}U\Pi + \Pi U\bar{\Pi}\right\| \leq \epsilon/\Delta$ is tight.

□

[2] Following Ref. [4.36, 4.37] we adopt the normalisation $\|A\| = \|A\|$ for all rank-1 operators A.

Lemma 4.3 (Approximate unitarity on the ground space)**.** *For an ϵ-approximate symmetry U with $\xi := \epsilon/\Delta \leq 1$, the action on the ground space is approximately unitary*

$$|||\Pi - |\Pi U \Pi|||| \leq f(\xi^2),$$

where $f(x) := 1 - \sqrt{1-x}$. In the operator norm, this expression is tight.

Proof. First we can bound $|\Pi U \Pi|^2$ near Π by using the unitarity of U itself as

$$\begin{aligned}
\Pi - |\Pi U \Pi|^2 &= \Pi - \Pi U^\dagger \Pi U \Pi \\
&= \Pi U^\dagger U \Pi - \Pi U^\dagger \Pi U \Pi \\
&= \Pi U^\dagger \bar{\Pi} U \Pi \\
&= \left|\bar{\Pi} U \Pi\right|^2 .
\end{aligned}$$

Together with Lemma 4.2, the sub-multiplicativity of unitarily invariant norms on finite-rank operators let us conclude that $|||\Pi - |\Pi U \Pi|^2|| \leq \xi^2$. Next we need to use this bound on $|||\Pi - |\Pi U \Pi|^2|||$, and create a bound on $|||\Pi - |\Pi U \Pi|||$.

Consider a function $f(x) = \sum_{n=1}^{\infty} a_n x^n$, where $a_n > 0$ and $f(1) < \infty$. For any finite-rank operator $0 \leq X \leq 1$, we can use the triangle inequality and submultiplicativity of $|||\cdot|||$ to derive a Jensen-like inequality

$$|||f(X)||| = \left|\left|\left| \sum_{n=1}^{\infty} a_n X^n \right|\right|\right| \leq \sum_{n=1}^{\infty} a_n |||X^n||| \leq \sum_{n=1}^{\infty} a_n |||X|||^n = f\left(|||X|||\right).$$

If we let $a_n = \Gamma(n-1/2)/2\sqrt{\pi} n!$, then we get $f(x) = 1 - \sqrt{1-x}$ for $x \in [0,1]$. If we let $X = \Pi - |\Pi U \Pi|$, then applying the above gives

$$\begin{aligned}
|||\Pi - |\Pi U \Pi|||| &= \left|\left|\left| f\left(\Pi - |\Pi U \Pi|^2\right) \right|\right|\right| \\
&\leq f\left(\left|\left|\left| \Pi - |\Pi U \Pi|^2 \right|\right|\right|\right) \\
&\leq f(\xi^2).
\end{aligned}$$

We note that $x/2 \leq f(x) \leq x$, which means that this bound improves upon the bound trivially given by the contractivity of $\Pi U \Pi$,

$$|||\Pi - |\Pi U \Pi|||| \leq \left|\left|\left| \Pi - |\Pi U \Pi|^2 \right|\right|\right| \leq \xi^2.$$

For the purposes of tightness, consider a two-dimensional Hilbert space, and a Hamiltonian H and unitary U given by

$$H = \begin{pmatrix} 0 & \\ & \Delta \end{pmatrix} \qquad \text{and} \qquad U = \begin{pmatrix} \cos\phi & \sin\phi \\ -\sin\phi & \cos\phi \end{pmatrix}.$$

In the operator norm $\|[U, H]\| = \Delta \sin\phi$ and $\|\Pi - |\Pi U \Pi|\| = 1 - \cos\phi$, which saturates the above bound. □

Using these bounds we can now construct a ground symmetry $\tilde{U}$ by pinching U with respect to Π, and then restoring unitarity on the ground space.

Lemma 4.4 (Approximate symmetries are nearly ground symmetries)**.** *For an ϵ-approximate symmetry U with $\xi := \epsilon/\Delta \leq 1$, there exists a ground symmetry $\tilde{U}$ which is close to U*

$$\left\|U - \tilde{U}\right\| \leq \xi + f(\xi^2),$$

and closer still in the ground space

$$\left\|\Pi\left(U - \tilde{U}\right)\Pi\right\| \leq f(\xi^2)\,.$$

The first inequality is tight to leading order in ξ, and the second is tight in the operator norm.

Proof. We start by considering the polar of decompositon $\Pi U \Pi = W\,|\Pi U \Pi|$. As the ground space $\mathrm{im}(\Pi)$ is an invariant subspace of $\Pi U \Pi$, we can take[3] W to also leave the ground space invariant, $[W, \Pi] = 0$. Given this, we define our ground symmetry to be $\tilde{U} := \Pi W \Pi + \bar{\Pi} U \bar{\Pi}$.

We will now consider bounding the distance between U and $\tilde{U}$ block-wise. The off-diagonal blocks are bounded by Lemma 4.2 as

$$\left\|\Pi\left(\tilde{U} - U\right)\bar{\Pi} + \bar{\Pi}\left(\tilde{U} - U\right)\Pi\right\| = \left\|\bar{\Pi} U \Pi + \bar{\Pi} U \Pi\right\| \leq \xi.$$

The bound on the ground space however follows from Lemma 4.3

$$\left\|\Pi\left(\tilde{U} - U\right)\Pi\right\| = \left\|W\Pi - W\,|\Pi U \Pi|\right\| = \left\|\Pi - |\Pi U \Pi|\right\| \leq f(\xi^2).$$

Finally the fact that U was unchanged on the excited space trivially implies

$$\left\|\bar{\Pi}\left(\tilde{U} - U\right)\bar{\Pi}\right\| = 0.$$

Putting everything together, this gives the desired bound

$$\begin{aligned}
\left\|\tilde{U} - U\right\| &= \left\|\bar{\Pi}\left(\tilde{U} - U\right)\Pi + \Pi\left(\tilde{U} - U\right)\bar{\Pi} + \Pi\left(\tilde{U} - U\right)\Pi + \bar{\Pi}\left(\tilde{U} - U\right)\bar{\Pi}\right\| \\
&\leq \left\|\bar{\Pi}\left(\tilde{U} - U\right)\Pi + \Pi\left(\tilde{U} - U\right)\bar{\Pi}\right\| + \left\|\Pi\left(\tilde{U} - U\right)\Pi\right\| + \left\|\bar{\Pi}\left(\tilde{U} - U\right)\bar{\Pi}\right\| \\
&\leq \xi + f(\xi^2).
\end{aligned}$$

As for tightness, Lemma 4.2 gives that $\left\|\bar{\Pi} U \Pi + \Pi U \bar{\Pi}\right\| = \|[H, U]\|/\Delta$ for Hamiltonians of the form $H = \Delta\bar{\Pi}$. If we assume that $\|[U, H]\| = \epsilon$, then applying the pinching inequality gives

$$\left\|U - \tilde{U}\right\| \geq \left\|\bar{\Pi}\left(U - \tilde{U}\right)\Pi + \Pi\left(U - \tilde{U}\right)\bar{\Pi}\right\| = \left\|\bar{\Pi} U \Pi + \Pi U \bar{\Pi}\right\| = \xi,$$

which proves our bound on $\left\|U - \tilde{U}\right\|$ is tight to leading order in ξ. The tightness of the norm distance in the ground space follows directly from the tightness of Lemma 4.3. □

[3] Such a W could be found by performing the polar decomposition restricted to the ground space, and padding the unitary out to act as the identity on the rest of the space.

We will now consider how the existence of nearby ground symmetries allows twisted commutation relations of approximate symmetries to be pulled down into the ground space.

Theorem 4.2 (Restriction to the ground space)**.** *For two ϵ-approximate symmetries U and V which approximately twisted commute*

$$\left\|\left[U,V\right]_\alpha\right\| \leq \delta,$$

then if $\xi := \epsilon/\Delta \leq 1$ there exists unitaries u and v acting on the ground space which also approximately twisted commute as

$$\left\|\left[u,v\right]_\alpha\right\| \leq \delta + 2\xi^2 + 4f(\xi^2)\,.$$

Proof. Consider a $\tilde{U}$ and $\tilde{V}$ given by applying Lemma 4.4 to U and V respectively, such that

$$\left\|\Pi U\Pi - \Pi\tilde{U}\Pi\right\|,\ \left\|\Pi V\Pi - \Pi\tilde{V}\Pi\right\| \leq f(\xi^2).$$

Next we consider the twisted commutator of U and V, and that of $\tilde{U}$ and $\tilde{V}$, both projected into the ground space. By expanding out the twisted commutators we have

$$\begin{aligned}
\Pi\left[U,V\right]_\alpha\Pi - \Pi\left[\tilde{U},\tilde{V}\right]_\alpha\Pi &= \left[\Pi U\Pi,\Pi V\Pi\right]_\alpha - \left[\Pi\tilde{U}\Pi,\Pi\tilde{V}\Pi\right]_\alpha + \Pi U\bar{\Pi}\cdot\bar{\Pi}V\Pi - e^{2\pi i\alpha}\Pi V\bar{\Pi}\cdot\bar{\Pi}U\Pi,\\
&= \left(\Pi U\Pi - \Pi\tilde{U}\Pi\right)\cdot\Pi V\Pi - e^{2\pi i\alpha}\Pi\tilde{V}\Pi\cdot\left(\Pi U\Pi - \Pi\tilde{U}\Pi\right)\\
&\quad + \Pi\tilde{U}\Pi\cdot\left(\Pi V\Pi - \Pi\tilde{V}\Pi\right) - e^{2\pi i\alpha}\left(\Pi V\Pi - \Pi\tilde{V}\Pi\right)\cdot\Pi U\Pi\\
&\quad + \Pi U\bar{\Pi}\cdot\bar{\Pi}V\Pi - e^{2\pi i\alpha}\Pi V\bar{\Pi}\cdot\bar{\Pi}U\Pi.
\end{aligned}$$

Using the triangle inequality, the contractivity of $\Pi U\Pi$ and $\Pi V\Pi$, and the bound on the off-diagonal blocks from Lemma 4.2, we can bound this as required:

$$\begin{aligned}
\left\|\Pi\left[U,V\right]_\alpha\Pi - \Pi\left[\tilde{U},\tilde{V}\right]_\alpha\Pi\right\| &\leq \left\|\left(\Pi U\Pi - \Pi\tilde{U}\Pi\right)\cdot\Pi V\Pi\right\| + \left\|\Pi\tilde{V}\Pi\cdot\left(\Pi U\Pi - \Pi\tilde{U}\Pi\right)\right\|\\
&\quad + \left\|\Pi\tilde{U}\Pi\cdot\left(\Pi V\Pi - \Pi\tilde{V}\Pi\right)\right\| + \left\|\left(\Pi V\Pi - \Pi\tilde{V}\Pi\right)\cdot\Pi U\Pi\right\|\\
&\quad + \left\|\Pi U\bar{\Pi}\cdot\bar{\Pi}V\Pi\right\| + \left\|\Pi V\bar{\Pi}\cdot\bar{\Pi}U\Pi\right\|\\
&\leq 2\left\|\Pi U\Pi - \Pi\tilde{U}\Pi\right\| + 2\left\|\Pi V\Pi - \Pi\tilde{V}\Pi\right\|\\
&\quad + \left\|\Pi U\bar{\Pi}\right\|\cdot\left\|\bar{\Pi}V\Pi\right\| + \left\|\Pi V\bar{\Pi}\right\|\cdot\left\|\bar{\Pi}U\Pi\right\|\\
&\leq 4f(\xi^2) + 2\xi^2.
\end{aligned}$$

Next we let u and v be the restriction of $\tilde{U}$ and $\tilde{V}$ to the ground space respectively. As each are ground symmetries, u and v are both unitaries. If we consider the embedding of operators on the ground space back into the larger Hilbert space, then we can use the above to bound the twisted

commutator of our ground space unitaries

$$\begin{aligned}
\|[u, v]_\alpha\| &= \|[u \oplus 0, v \oplus 0]_\alpha\| \\
&= \left\|[\Pi\tilde{U}\Pi, \Pi\tilde{V}\Pi]_\alpha\right\| \\
&= \left\|\Pi[\tilde{U}, \tilde{V}]_\alpha\Pi\right\| \\
&\leq \|\Pi[U, V]_\alpha\Pi\| + 2\xi^2 + 4f(\xi^2) \\
&\leq \|[U, V]_\alpha\| + 2\xi^2 + 4f(\xi^2) \\
&\leq \delta + 2\xi^2 + 4f(\xi^2).
\end{aligned}$$

Note that if we had a set of more than two unitaries, this additive growth in the twisted commutation value would hold equally for every pair separately. □

4.3 Degeneracy lower bounds

In this section we show how twisted pairs of unitary operators can be used to give lower bounds on the degeneracy of the ground space. We start by considering an exact twisted pair and the Stone-von Neumann theorem. We will then show how this argument can be generalised to approximate twisted pairs, and how a lower bound on the degeneracy follows from an upper bound on the twisted commutator value. Finally we will see how this can also be extended to more general collections of twisted commuting operators through the example case of two twisted pairs that are approximately mutually commuting.

4.3.1 Stone-von Neumann Theorem

Consider a u and v which exactly twisted commute, so that $uv = e^{2i\pi\alpha}vu$. Let $(\lambda, |\psi\rangle)$ be an eigenpair of u. Using the twisted commutation relation, we see that $|\psi'\rangle := v|\psi\rangle$ forms a $\lambda e^{2i\pi\alpha}$-eigenvector. It follows that v forms an isomorphism between the λ and $\lambda e^{2i\pi\alpha}$-eigenspaces of u, which allows us to conclude that their dimensions must be the same. Carrying this argument forward, we can see that any eigenspaces whose eigenvalues differ by any power of $e^{2i\pi\alpha}$ must also be isomorphic.

Suppose we take $\alpha \in \mathbb{Q}$, with $\alpha = p/q$ with p, q coprime. As we can see in Figure 4.1, a simple divisibility argument implies that the eigenspaces come in isomorphic multiples of q, which therefore implies that the overall dimension of u and v is a multiple of q also.

We now generalise this connection between the twisted commutator and the spectrum of one of the operators to allow for only approximate twisted commutation.

4.3.2 One twisted pair

Let us first extend the above argument to the case of a single approximate twisted pair. For simplicity, we consider the case where $\alpha = p/q$ with $p = 1$ and $q = d$, so the corresponding phase in the twisted commutator is $\eta := e^{2i\pi/d}$. This is not much of a restriction since if $p > 1$ we can replace v with $v^{\bar{p}}$ where $\bar{p}$ is the modular multiplicative inverse of p such that $\bar{p}p = 1 \bmod q$ and then apply the results of the $p = 1$ case. Under this substitution the twisted commutator will grow by at most a factor of

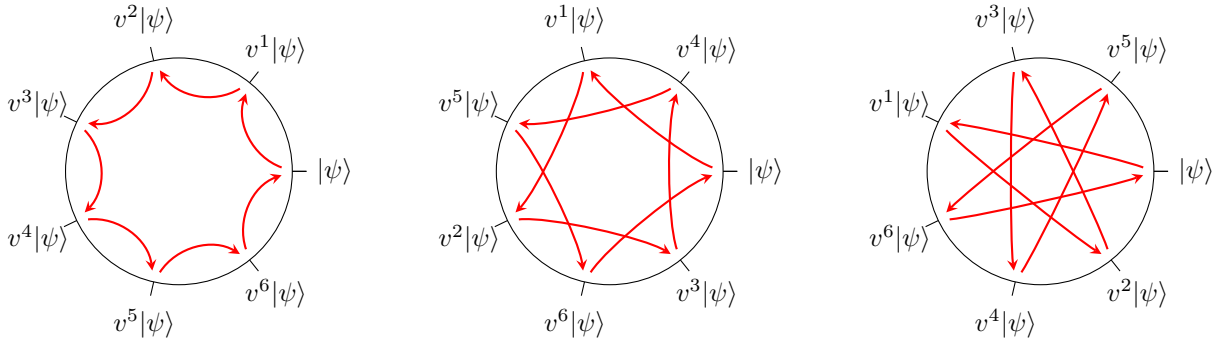


Figure 4.1: The action of powers of v on an eigenvector $|\psi\rangle$ of u. On the left $[u, v]_{1/7} = 0$, in the centre $[u, v]_{2/7} = 0$, and on the right $[u, v]_{3/7} = 0$. Here the position of the circle represents the corresponding eigenvalue of u.

$\lfloor q/2 \rfloor$. However, in Appendix 4.B we will show an alternative method that in fact works for arbitrary $\alpha \in \mathbb{R}$ and gives tighter bounds than this simple reduction. We also consider without loss of generality the case where u has at least one $+1$ eigenvalue, which can always be achieved by redefining u by multiplying by a complex unit phase factor.

Suppose we have two unitaries u and v such that

$$\left\| [u, v]_{1/d} \right\| = \|uv - \eta vu\| \leq \delta.$$

Our results will show that these operators must, for sufficiently small δ, be at least d-dimensional. To do this we will explicitly show that u has at least d distinct eigenvalues.

Let $|\psi\rangle$ be a $+1$ eigenvector of u, i.e. $u|\psi\rangle = |\psi\rangle$. Consider the orbit of $|\psi\rangle$ under v, i.e. the states $|j\rangle := v^j|\psi\rangle$ for $j = -\lfloor \frac{d-1}{2} \rfloor, \ldots, \lceil \frac{d-1}{2} \rceil$. These vectors are precisely the vectors depicted in Figure 4.1. We first show that these are approximate eigenstates of u.

Lemma 4.5 (Change in expectation value: One pair)**.** *The expectation value of u with respect to $|j\rangle$ is approximately η^j, specifically*

$$\left| \langle j|u|j\rangle - \eta^j \right| \leq |j|\,\delta.$$

Proof. This follows from the twisted commutator of u and v being small. By expanding the commutator and applying the triangle inequality we can see that $\|uv - \eta vu\| \leq \delta$ implies $\left\|uv^j - \eta^j v^j u\right\| \leq |j|\,\delta$. From this we can see that the expectation value of $|j\rangle$ lies close to η^j:

$$\begin{aligned}
|j|\,\delta &\geq \left\|uv^j - \eta^j v^j u\right\| \\
&= \left\|v^{-j}uv^j - \eta^j u\right\| \\
&\geq \left|\langle\psi|\left[v^{-j}uv^j - \eta^j u\right]|\psi\rangle\right| \\
&\geq \left|\langle\psi|v^{-j}uv^j|\psi\rangle - \eta^j \langle\psi|u|\psi\rangle\right| \\
&= \left|\langle j|u|j\rangle - \eta^j\right|.
\end{aligned}$$

□

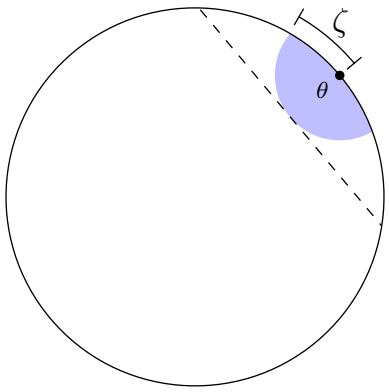


Figure 4.2: Lemma 4.6 gives that if there exists an expectation value in the blue region, there must exist an eigenvalue within the minor segment indicated by the dotted line.

So we can see that the $\{|j\rangle\}$ form a set of vectors with expectation values distributed approximately evenly around the unit circle, much like the states in the $\delta = 0$ case as seen in Figure 4.1. To relate these states to the dimensions of u and v, we will now show that there must exist an eigenvalue of u near the expectation value of each state.

Lemma 4.6 (Existence of eigenvalues)**.** *If there exists a state $|x\rangle$ such that*

$$\left|\langle x|u|x\rangle - e^{i\theta}\right| \leq \zeta$$

then u possesses a nearby eigenvalue $e^{i\phi}$ such that

$$|\phi - \theta| \leq \cos^{-1}(1-\zeta).$$

Proof. The bound on the expectation value with respect to u implies

$$\operatorname{Re}\left\langle x\middle|e^{-i\theta}u\middle|x\right\rangle \geq 1-\zeta.$$

As this expectation value is a convex combination of the eigenvalues of u, all of which lie on the unit circle, there must exists an eigenvalue of $e^{-i\theta}u$ with real value at least $1-\zeta$ (see Figure 4.2). This in turn implies that u possesses an eigenvalue $e^{i\phi}$ such that

$$\operatorname{Re} e^{i(\phi-\theta)} = \cos(\phi-\theta) \geq 1-\zeta.$$

□

Combining the two above lemmas, we can place a lower bound on the number of distinct eigenvalues of u.

Theorem 4.4. *If u and v are unitaries such that*

$$\left\|[u,v]_{1/d}\right\| < \frac{2}{d-1}\left[1-\cos\pi/d\right],$$

then the dimension of each operator is at least d.

Proof. From Lemma 4.5 we know that $\left|\langle j|u|j\rangle - e^{2i\pi j/d}\right| \leq |j|\,\delta$. Applying Lemma 4.6 we therefore get that u must have a corresponding eigenvalue $e^{i\phi_j}$ where

$$|\phi_j - 2j\pi/d| \leq \cos^{-1}\left(1 - |j|\,\delta\right).$$

As such we can see that each eigenvalue is within some error of a dth root of unity.

Next we want to find a bound for δ which ensures that these eigenvalues must be distinct, by bounding the regions in which these eigenvalues must exist away from each other. To do this we need $|\phi_j - \phi_k| > 0$ for all $j \neq k$. Taking the worst case over $j \neq k$:

$$\begin{aligned}
|\phi_j - \phi_k| &= \left|\frac{2\pi}{d}(j-k) + \left(\phi_j - \frac{2j\pi}{d}\right) - \left(\phi_k - \frac{2k\pi}{d}\right)\right| \\
&\geq \frac{2\pi}{d}\,|j-k| - \left|\phi_j - \frac{2j\pi}{d}\right| - \left|\phi_k - \frac{2k\pi}{d}\right| \\
&\geq \frac{2\pi}{d} - \cos^{-1}\left(1 - \left\lceil\frac{d-1}{2}\right\rceil\delta\right) - \cos^{-1}\left(1 - \left\lfloor\frac{d-1}{2}\right\rfloor\delta\right).
\end{aligned}$$

Here the last line follows from the fact that j and k cannot both saturate the worst-case distance of $\lceil\frac{d-1}{2}\rceil$. Therefore, the worst case can be chose without loss of generality to be $j = \lceil\frac{d-1}{2}\rceil$ and $k = -\lfloor\frac{d-1}{2}\rfloor$. Using the concavity of $\cos^{-1}(z)$ over $z \in [0,1]$, we can loosen this to

$$|\phi_j - \phi_k| \geq \frac{2\pi}{d} - 2\cos^{-1}\left(1 - \frac{d-1}{2}\,\delta\right).$$

Clearly this step is trivial for odd d.

Thus we get that a sufficient condition for all of the eigenvalues to be distinct is that the right-hand side of this inequality is strictly positive, and therefore we have the equivalent condition

$$\cos^{-1}\left(1 - \frac{d-1}{2}\,\delta\right) < \frac{\pi}{d}.$$

Rearranging, we find the specified bound on δ of

$$\delta < \frac{2}{d-1}\Big[1 - \cos(\pi/d)\Big].$$

□

Above we have only considered the case $d = 1/\alpha$, similar analysis could be performed for bounds required to certify dimensions $d' \neq 1/\alpha$. In Section 4.B we describe an algorithm for calculating which dimensions can be certified for an arbitrary pair of parameters α and δ — running this algorithm gives Figure 4.3.

4.3.3 Two twisted pairs

Next we are going to argue that the above analysis can be extended to more general collections of twisted commuting symmetries. By way of example, we are going to consider the case of two twisted pairs

$$\left\|[u_1, v_1]_{1/d_1}\right\| \leq \delta, \qquad \left\|[u_2, v_2]_{1/d_2}\right\| \leq \delta,$$

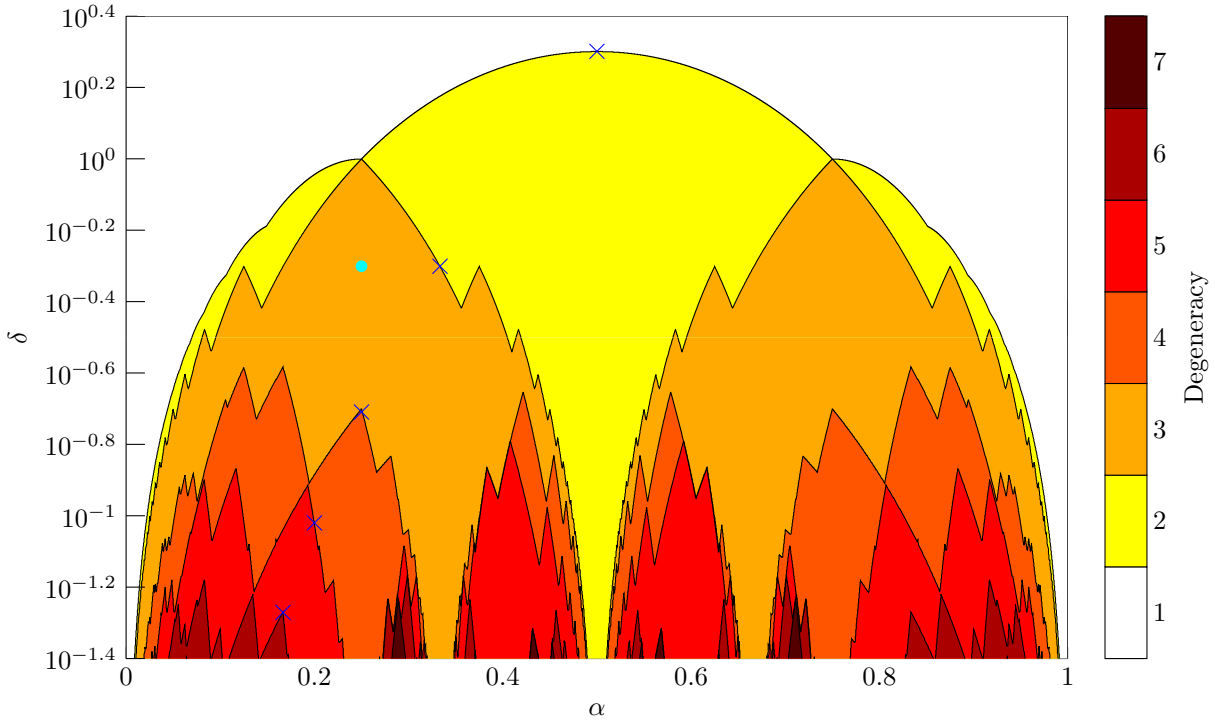


Figure 4.3: The dimension that can be certified, as a function of the twisted commutator value and twisting parameter, i.e. the minimum possible dimension of operators u and v for which $\|[u,v]_\alpha\| \leq \delta$ as a function of α and δ. The blue crosses indicate the bounds corresponding to degeneracy d and $\alpha = 1/d$, as considered in Theorem 4.3. The algorithm for calculating this figure is demonstrated in Section 4.B, by considering how the certification is calculated at the turquoise dot ($\alpha = 1/4$ and $\delta = 1/2$).

each of which approximately commute

$$\|[u_1, u_2]\| \leq \gamma, \qquad \|[u_1, v_2]\| \leq \delta, \qquad \|[u_2, v_1]\| \leq \delta.$$

The equivalent of Stone-von Neumann theorem laid our in Section 4.3.1 gives that for $\gamma = \delta = 0$, the dimension of such operators must be a multiple of $d_1 d_2$. We are going to give bounds on γ and δ below which we can prove the dimension to be at least $d_1 d_2$.

Previously we bounded the dimension from below by bounding the number of distinct eigenvalues. This is possible because these eigenvalues imply the existence of an orthonormal set of associated eigenvectors. As u_1 and u_2 do not commute, they will not necessarily possess an orthonormal set of shared eigenvectors. Instead we will have to address these vectors more directly, constructing *approximate shared eigenvectors* and proving their linear independence. First we will see that the approximate commutation of u_1 and u_2 can be used to demonstrate the existence of such a vector.

The existence of approximate shared eigenvectors of approximately commuting matrices was first proven in generality by Bernstein in Ref. [4.34]. Whilst Bernstein considers potentially non-normal matrices, in our case both u_1 and u_2 are unitary. In Section 4.A we leverage this additional structure to exponentially tighten the bounds on the approximate shared eigenvectors. One of the relevant bounds considered in Section 4.A gives the following immediate corollary.

Lemma 4.7 (Approximate eigenvector)**.** *There exists a vector $|\psi\rangle$ such that, after multiplying u_1 and u_2 by appropriate phase factor, it is an approximate shared $+1$-eigenvector of both, namely that*

$$\|u_1|\psi\rangle - |\psi\rangle\|, \|u_2|\psi\rangle - |\psi\rangle\| \leq \sqrt{\gamma} d_1 d_2/2.$$

Proof. Given an assumption that the dimension is at most $d_1 d_2$, this is a direct application of Theorem 4.1, which we consider in detail in Section 4.A. □

As in the case of a single pair, we will then consider the orbit of this vector under the action of products of v_1 and v_2. Let $|i,j\rangle := v_1^i v_2^j |\psi\rangle$ for $i = -\left\lfloor \frac{d_1-1}{2} \right\rfloor, \ldots, \left\lceil \frac{d_1-1}{2} \right\rceil$ and $j = -\left\lfloor \frac{d_2-1}{2} \right\rfloor, \ldots, \left\lceil \frac{d_2-1}{2} \right\rceil$. For convenience once again let $\eta_i := e^{2i\pi/d_i}$.

Lemma 4.8 (Change in expectation value: two pair)**.** *The states $|i,j\rangle$ are shared approximate eigenstates of u_1 and u_2. Specifically their approximate eigenvalues are the corresponding powers of η_1 and η_2*

$$\left|\langle i,j|u_1|i,j\rangle - \eta_1^i\right|, \left|\langle i,j|u_2|i,j\rangle - \eta_2^j\right| \leq \sqrt{\gamma} d_1 d_2/2 + (|i| + |j|)\,\delta.$$

Proof. From Lemma 4.7 we have

$$|\langle\psi|u_1|\psi\rangle - 1| \leq \sqrt{\gamma} d_1 d_2/2.$$

Applying an argument similar to that in Lemma 4.5 we can bound the change in eigenvalue under the action of v_c as

$$\left|\langle i,0|u_1|i,0\rangle - \eta_1^i \langle\psi|u_1|\psi\rangle\right| \leq |i|\,\delta.$$

Applying the same argument for v_2 gives

$$|\langle i,j|u_1|i,j\rangle - \langle i,0|u_1|i,0\rangle| \leq |j|\,\delta.$$

The triangle inequality allows us to merge these three inequalities, giving the stated bound. A similar argument can be performed for u_2. □

In the single pair case, we used the expectation values to imply the existence of nearby eigenvalues. Due to the lack of a shared eigenbasis of u_1 and u_2, we cannot do the same in the two pair case.

The reason that a set of distinct eigenvalues lower bounds the dimension is that, for normal operators such as unitaries, the eigenvalues imply the existence of an orthonormal eigenbasis. Instead of proving the existence of such vectors indirectly through the eigenvalues, we could instead prove our vectors $\{|i,j\rangle\}$ to be linearly independent — this is the approach we will take.

To this end, we will start by showing two approximate eigenvectors of a unitary with inconsistent expectation values are approximately orthogonal.

Lemma 4.9 (Low overlap)**.** *If two normalised vectors $|x\rangle$ and $|y\rangle$ have expectation values with some unitary w such that*

$$\left|\langle x|w|x\rangle - e^{i\theta_x}\right| \leq \zeta \qquad \text{and} \qquad \left|\langle y|w|y\rangle - e^{i\theta_y}\right| \leq \zeta$$

then the two vectors have a bounded overlap

$$|\langle x|y\rangle| \leq \sqrt{2\zeta}\left|\csc\left(\frac{\theta_y - \theta_x}{4}\right)\right|.$$

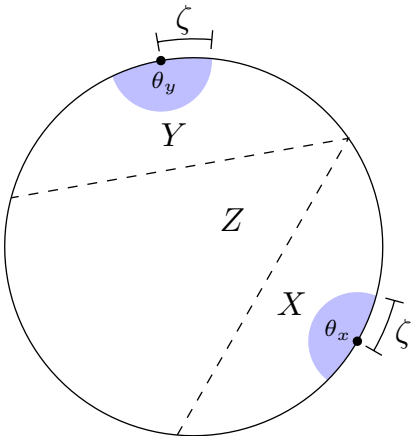


Figure 4.4: A disc representing the expectation values of vectors with respect to w, as well as the three regions X, Y, Z into which the disc is divided. The expectation values with respect to $|x\rangle$ and $|y\rangle$ lie in each of the blue regions.

Proof. Firstly, let $w' := e^{-i\theta_x} w$ and $\theta := \theta_y - \theta_x$. Next consider splitting the unit circle into three arcs X, Y, and Z. We let X and Y be centred on θ_x and θ_y respectively, and define them to be the largest possible regions such that they remain disjoint. We define Z to be the remaining arc, as shown in Figure 4.4. Note that by convexity any linear combination of eigenvectors whose eigenvalues lie in X will have an expectation value in the segment subtended by X, and similar for Y.

Now split $|x\rangle$ into two components

$$|x\rangle = \sqrt{1-\lambda_x}|x_X\rangle + \sqrt{\lambda_x}|x_{YZ}\rangle,$$

where $|x_X\rangle$ is in the span of eigenvectors with values in X, and $|x_{YZ}\rangle$ similar for $Y \cup Z$. By definition of X, we have

$$\operatorname{Re}\langle x_{YZ}|w'|x_{YZ}\rangle \leq \cos(\theta/2) \leq \operatorname{Re}\langle x_X|w'|x_X\rangle \leq 1.$$

Next we use the bound on the expectation value.

$$\begin{aligned}
\zeta &\geq |\langle x|w'|x\rangle - 1| \\
&\geq 1 - \operatorname{Re}\langle x|w'|x\rangle \\
&= 1 - (1-\lambda_x)\operatorname{Re}\langle x_X|w'|x_X\rangle - \lambda_x \operatorname{Re}\langle x_{YZ}|w'|x_{YZ}\rangle \\
&\geq 1 - (1-\lambda_x) - \lambda_x \cos(\theta/2) \\
&= 2\lambda_x \sin^2(\theta/4)
\end{aligned}$$

Thus we conclude that $\lambda_x \leq (\zeta/2)\csc^2(\theta/4)$. Similarly if we were to have decomposed $|y\rangle$ into parts contained in Y and XZ as $|y\rangle = \sqrt{1-\lambda_y}|y_Y\rangle + \sqrt{\lambda_y}|y_{XZ}\rangle$ then $\lambda_y \leq (\zeta/2)\csc^2(\theta/4)$.

Further decomposing

$$|x_{YZ}\rangle = \cos\varphi_x|x_Y\rangle + \sin\varphi_x|x_Z\rangle \qquad |y_{XZ}\rangle = \cos\varphi_y|y_X\rangle + \sin\varphi_y|y_Z\rangle\,,$$

then the inner product has the form

$$\begin{aligned}|\langle x|y\rangle| &= \left|\sqrt{1-\lambda_x}\sqrt{\lambda_y}\cos\varphi_y\,\langle x_X|y_X\rangle + \sqrt{\lambda_x}\sqrt{1-\lambda_y}\cos\varphi_x\,\langle x_Y|y_Y\rangle + \sqrt{\lambda_x}\sqrt{\lambda_y}\sin\varphi_x\sin\varphi_y\,\langle x_Z|y_Z\rangle\right| \\ &\leq \sqrt{1-\lambda_x}\sqrt{\lambda_y}\cos\varphi_y + \sqrt{\lambda_x}\sqrt{1-\lambda_y}\cos\varphi_x + \sqrt{\lambda_x}\sqrt{\lambda_y}\sin\varphi_x\sin\varphi_y.\end{aligned}$$

Using the identity $|A\cos\phi + B\sin\phi|^2 \leq |A|^2 + |B|^2$, we can maximise over φ_x to get

$$|\langle x|y\rangle| \leq \sqrt{1-\lambda_x}\sqrt{\lambda_y}\cos\varphi_y + \sqrt{\lambda_x}\sqrt{1-\lambda_y\cos^2\varphi_y}.$$

Using $\cos\varphi_y \leq 1$, we can simplify this bound to

$$|\langle x|y\rangle| \leq \sqrt{\lambda_y} + \sqrt{\lambda_x}.$$

Applying the ζ-dependent bounds on the λ values, we get the stated bounds. □

Now that we have a way of bounding the overlap between our vectors, we need to determine how low this overlap needs to be before linear independence can be ensured.

Lemma 4.10 (Overlap threshold)**.** *Take a set of normalised vectors $S = \{|v_i\rangle\}$ for $1 \leq i \leq n$. If the pairwise overlap between any two vectors is bounded $|\langle v_i|v_j\rangle| < 1/(n-1)$ for $i \neq j$, then S is linearly independent.*

Proof. Let G be the Gram matrix associated with S. As each of the vectors is normalised $G_{ii} = 1$ for all i. As all of the non-diagonal entries are strictly modulus-bounded by $1/(n-1)$, this matrix is strictly diagonally dominant, i.e.

$$|G_{ii}| > \sum_{j\neq i} |G_{ij}| \quad \text{for all } i.$$

From the Geshgorin circle theorem, such matrices are non-singular and full rank, allowing us to conclude that S is linearly independent.

Note that this analysis is tight, i.e. if $\langle v_i|v_j\rangle = -1/(n-1)$ for all $i \neq j$ then G is singular and $\sum_i |v_i\rangle = 0$. By considering the eigenvectors of such a Gram matrix, a set of vectors satisfying this can be backed out. □

Given this bound, we can finally find the condition for our vectors to be linearly independent and therefore lower bound the dimension of the space in which they reside.

Theorem 4.5. *If u_1, u_2, v_1 and v_2 are unitaries such that they satisfy the commutation relations*

$$\|[u_1, u_2]\| \leq \gamma \qquad \|[u_1, v_2]\| \leq \delta \qquad \|[u_2, v_1]\| \leq \delta$$

and twisted commutation relations

$$\left\|[u_1, v_1]_{1/d_1}\right\| \leq \delta \qquad \left\|[u_2, v_2]_{1/d_2}\right\| \leq \delta$$

with $d_1 \leq d_2$ and

$$\sqrt{\gamma}d_1d_2 + (d_1+d_2)\delta < \frac{\sin^2(\pi/2d_1)}{(d_1d_2-1)^2},$$

then the dimension of each operator is at least d_1d_2.

Proof. From Lemma 4.8 we have that our vectors have expectation values bounded near powers of η_1 and η_2

$$\left|\langle i,j|u_1|i,j\rangle - \eta_1^i\right|,\ \left|\langle i,j|u_2|i,j\rangle - \eta_2^j\right| \leq \sqrt{\gamma}d_1d_2/2 + (|i|+|j|)\,\delta.$$

Take a pair of vectors $|i,j\rangle$ and $|i',j'\rangle$ such that $i \neq i'$. Applying Lemma 4.9 with $w = u_1$ we get that their overlap is bounded as

$$\left|\langle i,j|i',j'\rangle\right|^2 \leq \left[\sqrt{\gamma}d_1d_2 + 2\max\{|i|+|j|\,,|i'|+|j'|\}\delta\right]\cdot \csc^2\left(\frac{\pi(i-i')}{2d_1}\right).$$

Combining this with a similar argument for u_2, and assuming $d_1 \leq d_2$, we get that for $(i,j) \neq (i',j')$

$$\left|\langle i,j|i',j'\rangle\right|^2 \leq \left[\sqrt{\gamma}d_1d_2 + (d_1+d_2)\delta\right]\csc^2\left(\frac{\pi}{2d_1}\right)$$

Thus we can see that

$$\left[\sqrt{\gamma}d_1d_2 + (d_1+d_2)\delta\right]\csc^2\left(\frac{\pi}{2d_1}\right) < \frac{1}{(d_1d_2-1)^2}$$

implies $|\langle i,j|i',j'\rangle| < 1/(d_1d_2-1)$ for all $(i,j) \neq (i',j')$. By Lemma 4.10 this means that the collection of vectors $\{|i,j\rangle\}_{i,j}$ are linearly independent, constructively proving the dimensionality of the operators in question to be at least d_1d_2. Rearranging this gives the specified bound. □

4.4 Minimum twisted commutation value

In the previous section we considered finding lower bounds on the dimensions of approximately twisting commuting operators. In the exact case, the Stone-von Neumann theorem (c.f. Theorem 4.1) tell us that unitaries x and y for which

$$[x,y]_{1/d} = 0$$

are not only *at least* d-dimensional, but are *a multiple of* d-dimensional. We might therefore hope for a more comprehensive understanding of twisted commutation that provides more information than simply a lower bound on the dimension. In this section we will consider the twisted commutator in the Schatten-Ky Fan norms $|||\cdot||| := \|\cdot\|_{(p,k)}$ with $p \geq 2$, and find the minimum possible twisted commutator value as a function of dimension.

Definition 4.3 (Minimum twisted commutator value)**.** *Let* $\Lambda_{g,\alpha}^{(p,k)}$ *be the* minimum twisted commutator value*, with respect to the Schatten-Ky Fan* (p,k)*-norm, over all pairs of unitary matrices of dimension* g

$$\Lambda_{g,\alpha}^{(p,k)} := \min_{u,v\in U(g)} \left\|[u,v]_\alpha\right\|_{(p,k)}.$$

In this language, the Stone-von Neumann theorem gives that $\Lambda_{g,\alpha}^{(p,k)} = 0$ if and only if $g\alpha \in \mathbb{Z}$. If we had an understanding of the values of $\Lambda_{g,\alpha}^{(p,k)}$ where $g\alpha \notin \mathbb{Z}$, then we could use twisted commutation value as a way of certifying dimension. In particular, if one thinks of α as fixed, and one knows the value $\left\|[u,v]_\alpha\right\|_{(p,k)}$ to be less than $\Lambda_{g,\alpha}^{(p,k)}$ for certain dimensions g, then these certain dimensions are ruled out as possible dimensions of u and v. In this section we will explicitly evaluate $\Lambda_{g,\alpha}^{(p,k)}$ for $p \geq 2$.

To lower bound $\Lambda_{g,\alpha}^{(p,k)}$, we will utilise techniques from spectral perturbation theory to bound a related quantity known as the *spectral distance*. By considering a family of operators which twisted commute, we will furthermore show this bound to be tight.

Definition 4.4 (Spectral distance)**.** *The* spectral (p,k)-distance $d_{(p,k)}(a,b)$ *between two matrices a and b is the (p,k)-norm of the vector containing the differences between eigenvalues of the two matrices, minimised over all possible orderings. If we let $\lambda(x)$ denote the vector of eigenvalues of a $g \times g$ matrix x then algebraically*

$$d_{(p,k)}(a,b) := \min_{\sigma \in S_g} \|\sigma\left[\lambda(a)\right] - \lambda(b)\|_{(p,k)} = \min_{\sigma \in S_g} \left(\sum_{j=1}^{k} \left|\lambda_{\sigma(j)}(a) - \lambda_j(b)\right|^p \right)^{1/p},$$

where the minimisation is over all elements σ of the permutation group S_g on g symbols.

4.4.1 Frobenius spectral bound

Before attacking the spectral distance, we are first going to restrict ourselves to the case of the Frobenius norm ($p = 2$, $k = g$), where we shall denote the norm by $\|\cdot\|_F$, the corresponding spectral distance by $d_F(\cdot,\cdot)$, and the twisted commutator minimum by $\Lambda_{g,\alpha}^{(F)}$. In this special case, the spectral distance between two normal matrices is bounded by their norm difference.

Lemma 4.11 (Wielandt-Hoffman inequality [4.39])**.** *For normal matrices a and b, $d_F(a,b) \leq \|a-b\|_F$.*

Once again let $\eta := e^{2i\pi\alpha}$. Applying Wielandt-Hoffman to $\Lambda_{g,\alpha}^{(F)}$ we see that the corresponding spectral distance provides a lower bound,

$$\Lambda_{g,\alpha}^{(F)} = \min_{u,v \in U(d)} \left\| v^\dagger u v - \eta u \right\|_F \geq \min_{u,v \in U(g)} d_F(v^\dagger u v, \eta u) = \min_{u \in U(g)} d_F(u, \eta u)\,.$$

Though $\left\|v^\dagger u v - \eta u\right\|_F$ depended on both u and v, $d_F(u,\eta u)$ depends only on the spectrum of u, making for a much simpler optimisation. This inequality will turn out to be tight for matrices minimising the twisted commutator value.

Denote the eigenvalues of u by $\{e^{i\theta_j}\}$, then the spectral distance in question is given by

$$d_F^2(u,\eta u) := \min_{\sigma \in S_g} \sum_{j=1}^{g} \left| e^{i\theta_{\sigma(j)}} - e^{i(\theta_j + 2\pi\alpha)} \right|^2 = \min_{\sigma \in S_g} \sum_{j=1}^{g} 4\sin^2\left(\frac{\theta_{\sigma(j)} - \theta_j - 2\pi\alpha}{2}\right).$$

Define $f(\sigma;\theta_1,\ldots,\theta_g)$ to be the argument of the above optimisation

$$f(\sigma;\theta_1,\ldots,\theta_g) := \sum_{j=1}^{g} 4\sin^2\left(\frac{\theta_{\sigma(j)} - \theta_j - 2\pi\alpha}{2}\right) \tag{4.1}$$

such that $d_F^2(u,\eta u) = \min_\sigma f(\sigma;\theta_1,\ldots,\theta_g)$. The optimisation of $d_F^2(u,\eta u)$ can therefore be reduced to an optimisation of $f(\sigma;\theta_1,\ldots,\theta_g)$.

We can now break the optimisation of f down into two parts. First we will show that for any assignment of permutation and angles, there exists a cyclic permutation, and adjusted angles, for which the value of f is the same. This will allow us to consider a minimising permutation which has only a single cycle without loss of generality. Secondly we shall see that, for such a cyclic permutation, the set of angles which minimise f are those that are equally distributed around the unit circle. Given these, we will find an explicit minimum for f, and thus for $d_F(u, \eta u)$.

Lemma 4.12 (Reduction to cyclic permutations)**.** *For a given multi-cycle permutation σ and set of angles $\{\theta_j\}$, there exists a cyclic permutation σ' and set of adjusted angles $\{\theta'_j\}$ such that*

$$f(\sigma; \theta_1, \ldots, \theta_g) = f(\sigma'; \theta'_1, \ldots, \theta'_g).$$

Proof. Firstly, our indices can be reordered such that the cycles of σ are contiguous, i.e. in cycle notation

$$\sigma = (1 \ \ldots \ k_1 - 1)\,(k_1 \ \ldots \ k_2 - 1)\ldots(k_n \ \ldots \ g),$$

for some $1 < k_1 \cdots < k_n \leq g$. (Note that the result is trivially true if $g = 1$, so we restrict to $g > 1$.) As f only depends on the difference between angles whose indices are within the same cycle of σ, if we shift all the angles within the same cycle by the same amount, the value of f will not change. For example if we take the change of angle

$$\theta'_j := \begin{cases} \theta_j - \theta_1 & 1 \leq j < k_1 \\ \theta_j - \theta_{k_1} & k_1 \leq j < k_2 \\ \quad \vdots & \\ \theta_j - \theta_{k_n} & k_n \leq j \leq g. \end{cases}$$

then $f(\sigma; \theta_1, \ldots, \theta_g) = f(\sigma; \theta'_1, \ldots, \theta'_g)$. Notice that $\theta'_1 = \theta'_{k_1} = \cdots = \theta'_{k_n} = 0$ by construction.

We now wish to merge the permutation σ into a single cyclic permutation

$$\sigma' := (1 \ \ldots \ g). \tag{4.2}$$

To do this, the only entries of the permutation which need to be changed are those at the end of each cycle.

$$\begin{aligned} \sigma(k_1 - 1) = 1 \quad &\rightarrow \quad \sigma'(k_1 - 1) = k_1 \\ \sigma(k_2 - 1) = k_1 \quad &\rightarrow \quad \sigma'(k_2 - 1) = k_2 \\ \vdots \quad &\qquad \vdots \\ \sigma(g) = k_n \quad &\rightarrow \quad \sigma'(g) = 1\,. \end{aligned}$$

By definition of the adjusted angles however, the only indices that change are those for which the angles have already been made identical in the previous step, i.e. $\theta'_{\sigma(j)} = \theta'_{\sigma'(j)}$ for all j. As f only depends on σ through how it acts on the angles, this means that this doesn't alter the value of f, therefore $f(\sigma; \theta'_1, \ldots, \theta'_g) = f(\sigma'; \theta'_1, \ldots, \theta'_g)$. □

Now that we have addressed the nature of the optimal permutation, namely showing that it can be taken to be cyclic, we turn out attention to the optimal angles.

Lemma 4.13. *For a given single-cycle permutation σ, the sets of angles which optimise f, as defined in Equation (4.1), correspond to those evenly distributed around the unit circle, and the difference between adjacent angles θ_j and $\theta_{\sigma(j)}$ is $2\pi\lfloor d\alpha \rceil / g$, where $\lfloor\cdot\rceil$ denotes integer rounding. Moreover the corresponding minimal value of f is*

$$\min_{\{\theta_j\}_j} f(\sigma;\theta_1,\ldots,\theta_g) = 2\sqrt{g}\sin\left(\pi\left|\frac{\lfloor g\alpha\rceil - g\alpha}{g}\right|\right).$$

Proof. Denote both of the terms[4] in f which depend non-trivially on θ_j by $f_j(\theta_j)$. Using the double angle formula and the auxiliary angle method, we can reduce the θ_j dependence to a single sinusoidal term.

$$\begin{aligned} f_j(\theta_j) &= 4\sin^2\left(\frac{\theta_j - \theta_{\sigma(j)} - 2\pi\alpha}{2}\right) + 4\sin^2\left(\frac{\theta_{\sigma^{-1}(j)} - \theta_j - 2\pi\alpha}{2}\right) \\ &= 4 - 4\cos\left(2\pi\alpha + \frac{\theta_{\sigma(j)} - \theta_{\sigma^{-1}(j)}}{2}\right)\cos\left(\theta_j - \frac{\theta_{\sigma(j)} + \theta_{\sigma^{-1}(j)}}{2}\right). \end{aligned}$$

We can therefore see that the optimal θ_j, leaving all other angles fixed, satisfies

$$\theta_j = \left(\theta_{\sigma(j)} + \theta_{\sigma^{-1}(j)}\right)/2 \mod \pi.$$

This implies that $\theta_{\sigma(j)} - \theta_j = \theta_j - \theta_{\sigma^{-1}(j)} \mod 2\pi$, i.e. θ_j lies in at the 'midpoint' of its neighbours, as described by σ. By inducting the above argument we find that $\theta_{\sigma(j)} - \theta_j = \theta_{\sigma(k)} - \theta_k \mod 2\pi$ for all j,k meaning that all adjacent angles are equally spaced around the unit circle. This means that if we have g angles, and label our indices such that $\sigma(j) = j+1 \mod g$, then for some fixed integer m, the optimal angles are of the form

$$\theta_j = \theta_1 + 2\pi m(j-1)/g. \tag{4.3}$$

The only free parameter left now is m, the spacing between adjacent points. Plugging these angles into the definition of f we find

$$f(\sigma;\theta_1,\ldots,\theta_g) = 2\sqrt{g}\left|\sin\left(\pi\left[m/g - \alpha\right]\right)\right|.$$

This is in turn minimised for $m = \lfloor g\alpha\rceil$, giving the stated spacing and minima. □

As this minimum of f is independent of the permutation σ, we get an overall minimum for f for free.

Corollary 4.14. *The minimum twisted commutator value (Definition 4.3) in the Frobenius norm $\Lambda^{(F)}_{g,\alpha}$ is lower bounded*

$$\Lambda^{(F)}_{g,\alpha} \geq 2\sqrt{g}\sin\left(\pi\left|\frac{\lfloor g\alpha\rceil - g\alpha}{g}\right|\right).$$

[4] In saying there are two such terms we have assumed $g \geq 3$. If $g = 1$ the lemma is trivial (f is constant), and if $g = 2$ then we have double counted in $f_j(\theta_j)$, but our analysis of its minimum remains valid.

Proof. This result can be seen by recalling that the definition of f in Equation (4.1) gives that

$$\min_{u\in U(g)} d_F(u,\eta u) = \min_{\sigma,\{\theta_j\}_j} f(\sigma,\theta_1,\ldots,\theta_g).$$

As Lemma 4.12 tells us that we can consider cyclic permutations without loss of generality, we can apply the minimum found in Lemma 4.13, giving

$$\min_{u\in U(g)} d_F(u,\eta u) = 2\sqrt{g}\sin\left(\pi\left|\frac{\lfloor g\alpha\rceil - g\alpha}{g}\right|\right).$$

Applying the Wielandt-Hoffman theorem (Lemma 4.11), we get that the above minimum spectral distance lower bounds the twisted commutator in the Frobenius norm, as required. □

4.4.2 Higher norms and tightness

With the above bound in hand, we now turn our attention to tightness. A canonical family of operators which exhibit twisted commutation is that of the *generalised Pauli operators*, also known as Sylvester's *clock and shift matrices*

$$C := \sum_j \omega^{j-1}|j\rangle\langle j|, \qquad S := \sum_j |j\oplus 1\rangle\langle j|$$

where $\omega = e^{2i\pi/g}$ is a primitive gth root of unity, and $\oplus$ denotes addition modulo g. As S simply cyclically permutes the eigenbasis of C, we can see that $S^\dagger CS = \omega C$, or $[C,S]_{1/g} = 0$. By taking appropriate powers these operators can also yield pairs which twisted commute with a phase that is any power of ω, specifically we see $\left[C,S^k\right]_{k/g} = 0$. Suppose we take such a pair and evaluate the twisted commutator at an arbitrary phase $\eta = e^{2i\pi\alpha}$. We then find,

$$\begin{aligned}\left\|\left[C,S^k\right]_\alpha\right\|_F &= \left\|CS^k - \eta S^k C\right\|_F \\ &= \left\|(1-\omega^k\eta)CS^k\right\|_F \\ &= \sqrt{g}\left|1-\omega^k\eta\right| \\ &= 2\sqrt{g}\left|\sin\left(\pi\left(\alpha + k/g\right)\right)\right|.\end{aligned}$$

If we now take $k = -\lfloor g\alpha\rceil$, then we saturate Corollary 4.14, proving tightness of the bound on $\Lambda^{(F)}_{g,\alpha}$, allowing us to conclude

$$\Lambda^{(F)}_{g,\alpha} = 2\sqrt{g}\sin\left(\pi\left|\frac{\lfloor g\alpha\rceil - g\alpha}{g}\right|\right).$$

For the above optimisations we restricted ourself to the $p = 2$ case of the Frobenius norm. The nature of the minimisers found allows us to pull this analysis up into minima for the $p > 2$ Schatten norms as well.

Theorem 4.6 (Minimum twisted commutation value)**.** *Suppose that u and v are g-dimensional unitaries, then for any $p \geq 2$ the twisted commutator is lower bounded*

$$\left\| [u, v]_\alpha \right\|_{(p,k)} \geq 2k^{1/p} \sin\left(\pi \left| \frac{\lfloor g\alpha \rceil - g\alpha}{g} \right|\right),$$

where $\|\cdot\|_{(p,k)}$ is the (p,k)-Schatten-Ky Fan norm. Moreover this bound is tight, in that sense that there exist families of g-dimensional unitaries which saturate the above bounds and only depend on $\lfloor g\alpha \rceil$, the nearest integer to $g\alpha$.

Proof. By the equivalence of Schatten-Ky Fan norms, the minimum Frobenius norm will also provide a lower bound for other (p,k)-norms as well. Specifically for $p \geq 2$ we have

$$\|M\|_{(p,k)} \geq k^{1/p} g^{-1/2} \|M\|_F .$$

Applying these to the definition of $\Lambda_{g,\alpha}^{(p,k)}$, this bound gives that $\Lambda_{g,\alpha}^{(p.k)} \geq k^{1/p} g^{-1/2} \Lambda_{g,\alpha}^{(F)}$ for $p \geq 2$. It turns out that this inequality is saturated by matrices M with flat spectra, i.e. those proportional to unitaries. It so happens that the clock and shift operators considered to demonstrate tightness have a twisted commutator with precisely this property, and therefore also saturate and demonstrate the tightness of the induced $p > 2$ bounds. We therefore conclude that

$$\Lambda_{g,\alpha}^{(p,k)} = k^{1/p} g^{-1/2} \Lambda_{g,\alpha}^{(F)} = 2k^{1/p} \sin\left(\pi \left| \frac{\lfloor g\alpha \rceil - g\alpha}{g} \right|\right).$$

□

Some plots of this bound are shown in Figure 4.5.

4.5 Applications and open questions

We now discuss several avenues for improvements, generalisations, refinements, and applications of these ideas.

4.5.1 Local Hamiltonians

In this paper the only assumption we made about our Hamiltonian H was the presence of a spectral gap. A natural additional structure to impose is that H be a many-body Hamiltonian: decompose our Hilbert space into a tensor product of many smaller Hilbert spaces, and let our Hamiltonian take the form

$$H = \sum_k h_k$$

where each term h_k acts non-trivially on a constant number of these tensor factor spaces. Additional to this we could also impose that the factors on which it acts are geometrically local as well. Under this special case it may be that either the bounds on degeneracy certification might be able to be improved, or we might be able to prove the existence of degeneracy witnesses with additional structure, e.g. such witnesses might act in a geometrically local fashion.

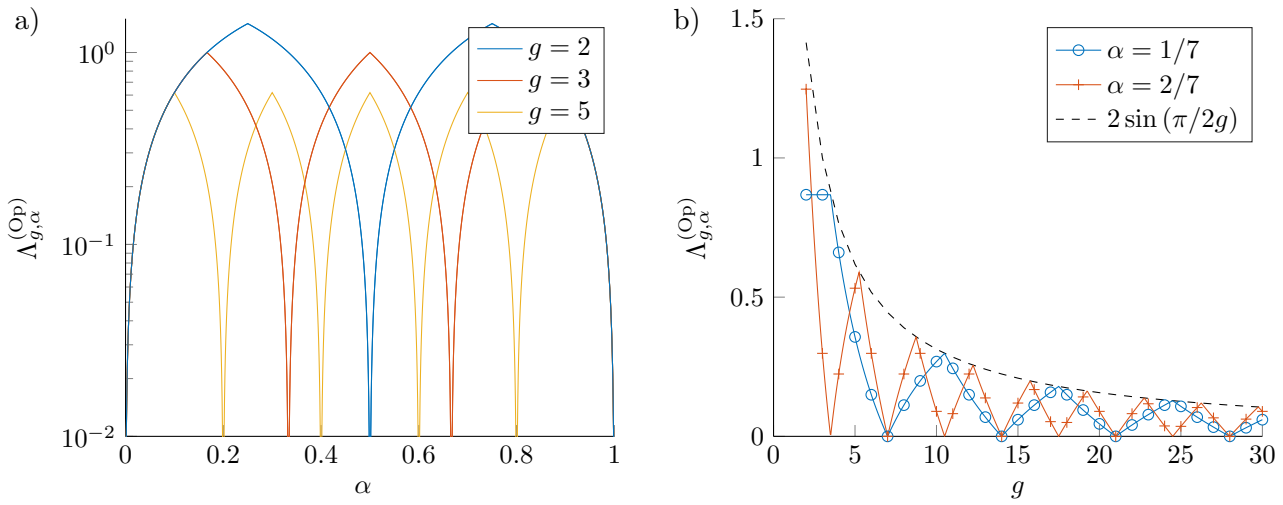


Figure 4.5: The twisted commutator value minimum (in the operator norm, $p = \infty$) $\Lambda^{(\mathrm{Op})}_{g,\alpha}$. a) The dependence on the twisting parameter α for a few fixed dimensions g. The presence of roots at multiples of $1/g$ are those predicted by Theorem 4.1. b) Now fixing the twisting parameter α, the dependence on the dimension g is shown. Note that g can only take integer values, indicated by the circles and pluses, with the continuous lines simply intended to guide the eye. The dotted black line indicates an α-independent upper bound on $\Lambda^{(\mathrm{Op})}_{g,\alpha}$ given by applying the bound $|x - \lfloor x \rceil| \leq 1/2$.

4.5.2 Topologically ordered systems

While the notions of approximate symmetry and degeneracy of a ground band are both robust to small perturbations, naïvely one can only consider perturbations of a strength no larger than the gap. For topologically ordered systems [4.18] however, we can afford much larger perturbations under certain locality assumptions.

Under the influence of local perturbations, the low-energy band structure, most notably the ground space degeneracy, is robust even if the overall strength of the perturbation is extensive [4.40, 4.41], for frustration-free Hamiltonians with local topological quantum order and a local gap. Moreover, any symmetries which witnesses this degeneracy can be quasi-adiabatically continued [4.42] into approximate symmetries which witness the degeneracy of the ground band in the perturbed system. It is in this sense that the existence of degeneracy witnesses can be considered robust to even rather strong perturbations, at least for the ground band.

The family of *abelian quantum double models* possess symmetries supported on quasi-1D regions which satisfy twisted commutation relations related to the braid and fusion rules of the underlying anyons [4.43]. More general models such as non-abelian/twisted quantum doubles [4.43–4.45], and Levin-Wen string net models [4.46] are all believed to possess symmetries which satisfy more general commutation-like relations based on more general notions of commutation. One possible example is the twist product [4.47] which only commutes the two operators on part of the system, braiding them together.

$$\Big(\sum_i A_i \otimes A'_i\Big) \infty \Big(\sum_j B_j \otimes B'_j\Big) := \sum_{ij} A_i B_j \otimes B'_j A'_i.$$

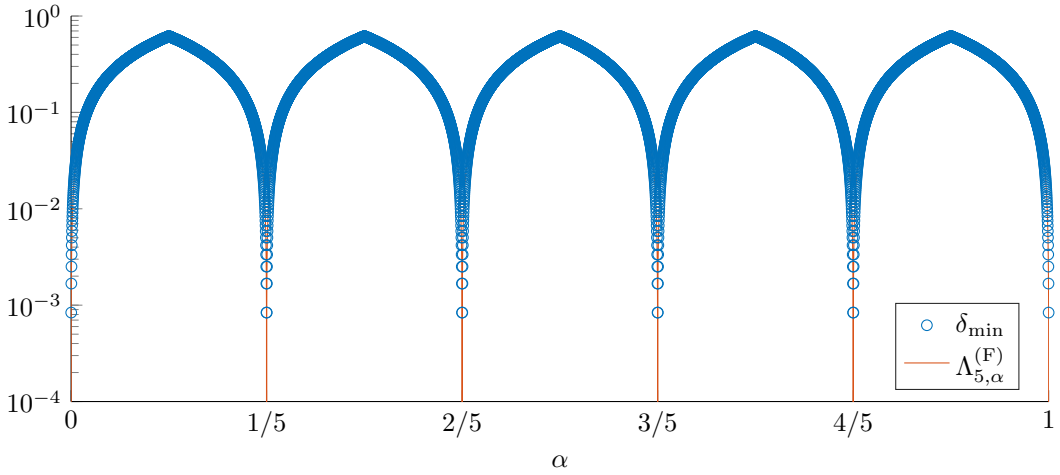


Figure 4.6: The twisted commutator value for ribbon operators on the $\mathbb{Z}_5$ quantum double model, as calculated using the algorithm of [4.35], as compared to the minimum possible twisted commutator value in 5-dimensions. Note that the difference between the two plots is no more than 3×10^{-13}.

An obvious extension of this work is to take various properties of these underlying systems implied by this commutation-like relations, and see if they too carry through into the regime of *approximate* relations.

In a recent paper, Bridgeman et. al. sought to classify the phases of 2D topologically ordered spin systems belonging to the same phase as abelian quantum doubles [4.35]. This was done by numerically optimising twisted pairs of symmetries. This optimisation was done over a tensor network [4.48, 4.49] ansatz of quasi-1D operators known as matrix product operators. For two operators L and R, supported on intersecting quasi-1D regions, the cost function takes the form

$$C(L, R; \alpha) \propto \epsilon_L^2 + \epsilon_R^2 + \delta^2$$

where $\epsilon_L := \|[L, H]\|_F$, $\epsilon_R := \|[R, H]\|_F$, and $\delta = \|[L, R]_\alpha\|_F$.

Minimising $C(L, R; \alpha)$ over L and R for a fixed α, they found that in the abelian quantum doubles the minimisers were unitary, and that both ϵ_L and ϵ_R vanish to within numerical accuracy, leaving only the twisted commutator value δ. By observing the values of α for which the minimum cost is low, they hoped to classify the topological phases of the underlying Hamiltonian. By Theorem 4.2 we know that, at least to within numerical accuracy, the ribbon operators found restrict down to ground symmetries with the same twisted commutation relations. In Figure 4.6 we compare, for the $\mathbb{Z}_5$ quantum double model, their numerically obtained values of this twisted commutator $\delta_{\min}$ with the minimal possible twisted commutator $\Lambda^{(F)}_{5,\alpha}$, showing close agreement and lending support to the efficacy of this numerical method.

4.5.3 Quantum codes

One class of systems for which twisted commuting symmetries play a special role are quantum codes, in which they can be interpreted as logical operators [4.19, 4.50, 4.51]. For a quantum code encoding

N codewords, the logical algebra must correspond to $\mathrm{Mat}_N(\mathbb{C})$, which necessarily contains a pair of operators X and Z such that $[X, Z]_{1/N} = 0$; indeed the algebra generated by any two such operators X and Z is itself $\mathrm{Mat}_N(\mathbb{C})$.

While the existence of logical operators which $\alpha = 1/N$ twisted commute can be ensured, we might only see and expect operators with twisted commutations characteristic of smaller ground spaces if we restrict the locality of these operators. Though the logical algebra is given by $\mathrm{Mat}_N(\mathbb{C})$, this space often naturally decomposes into a tensor product decomposition: the logical qudits. By geometrically restricting where on the system the operators can act, we can often restrict which factors the logical operators have nontrivial commutation relations with. This is the case for celebrated examples such as the toric code [4.43]. This can be seen above in Figure 4.6, where the logical operators are restricted to string-like regions that are only sensitive to one $\mathrm{Mat}_5(\mathbb{C})$ factor of the larger $\mathrm{Mat}_{25}(\mathbb{C})$ logical algebra; one of the two 5-level qudits. In the same way that Ref. [4.35] sought to use the existence of twisted commuting symmetries to classify topological phases, how this existence varies with respect to the geometry imposed on these operators might provide a tool to probe what portion of the logical algebra is accessible on certain regions.

In the language of quantum codes, our results can be interpreted as bounds below which approximate logical operators imply the existence of a certain number of code words. A possible avenue for future work is whether there exists bounds below which not only can the number of codestates be bounded, but reliable encoding, decoding, and error correction can all be performed with these approximate logical operators. Understanding when information stored in such states is approximately preserved, as opposed to exactly preserved [4.52], could have interesting applications in approximate quantum error correction.

Acknowledgements

We thank Jacob Bridgeman for fruitful comments, and for computing the numerical data for Figure 4.6 using the algorithm of Ref. [4.35]. For the use of American English in this paper, CTC apologises to the Commonwealth of Australia. This work was supported by the Australian Research Council via EQuS project number CE11001013 and STF was supported by an Australian Research Council Future Fellowship FT130101744.

4.A Approximate shared eigenvectors for approximately commuting matrices

In this section we will show that for two approximately commuting matrices, an approximate shared eigenvector exists. This problem has been considered before by Bernstein [4.34], who showed the following result.

Theorem 4.6 ([4.34]). *Take A and B to be complex matrices of dimension $n \geq 2$. If $\|B\| \leq 1$, and for some $\delta > 0$ we have*

$$\|[A, B]\| \leq \frac{\delta^n(1-\delta)}{1-\delta^{n-1}},$$

then for each eigenvalue λ of A, there exists a μ and normalised $|x\rangle$ such that

$$\|A|x\rangle - \lambda|x\rangle\|, \|B|x\rangle - \mu|x\rangle\| \leq \delta.$$

Notice above the required bound on the commutator scales like $\mathcal{O}(\delta^n)$ for small δ. Below we will improve this dimension scaling by adding the additional assumption that one of the matrices is normal, allowing us to bring this down to a $\mathcal{O}(\delta^2/n^2)$ dependence. First we will state the more general result, which only requires one of the matrices to be normal, followed by a more specialised result which applies when both matrices are normal.

The existence of an entire basis of shared approximate eigenvectors is closely related to approximate joint diagonalisation, a problem that has been widely considered and has found application in fields such as quantum chemistry [4.54], machine learning [4.55] and image processing [4.56]. This literature is too vast to review in this appendix, but see Ref. [4.53] for a discussion of the relationship between approximately commuting matrices and joint diagonalisation. Techniques similar to those used below have also been used in Ref. [4.57] to address the related problem of constructing nearby exactly commuting operators, in the case in which one matrix is Hermitian. Whilst this analysis gives better bounds than those presented below, it leverages a combinatorical construction [4.58] that explicitly uses the reality of the eigenvalues, and therefore cannot be directly applied to the case we will consider in which one matrix is normal, but not necessarily Hermitian.

Take A and B to be $n \times n$ matrices. Let A be normal, with an eigenvalue decomposition $A = \sum_i \lambda_i |i\rangle\langle i|$. Next take λ to be a specific eigenvalue of A. Let I_0 be the singleton set containing the index corresponding to λ, or all these indices if λ is degenerate. Define I_k to be all the indices whose eigenvalues are within some radius $r > 0$ in the complex plane (to be chosen later) of those in I_{k-1}, i.e.

$$I_k := \{i \,|\, \exists j \in I_{k-1} : |\lambda_i - \lambda_j| \leq r\}\,.$$

Clearly this sequence becomes fixed after at most n terms, and so let $I := I_n$ be this fixed point. Intuitively I can be thought of as the indices corresponding to eigenvalues which form a cluster around λ where every eigenvalue in the cluster is linked to at least one other by a disk of radius r in the complex plane.

By construction this set has two properties we require. First it is bounded away from any other index,

$$i \in I, j \notin I \implies |\lambda_i - \lambda_j| > r$$

Second, because all of the eigenvalues corresponding to elements in I have nearby neighbours in I, this means that the diameter of the disk containing all of the eigenvalues in I has a diameter bounded by at most nr,

$$i \in I \implies |\lambda_i - \lambda| \leq nr\,.$$

Next let V be the space spanned by the eigenvectors whose indices lies in I,

$$V := \text{Span}\,\{|i\rangle \,|\, i \in I\}\,.$$

Denoting the orthogonal complement of V by $\bar{V}$, then we can decompose both A and B into blocks on $V \oplus \bar{V}$ as

$$A = \begin{pmatrix} A_V & \\ & A_{\bar{V}} \end{pmatrix} \qquad \text{and} \qquad B = \begin{pmatrix} B_{VV} & B_{\bar{V}V} \\ B_{V\bar{V}} & B_{\bar{V}\bar{V}} \end{pmatrix}.$$

Lemma 4.15. *If $\|[A, B]\| \leq \epsilon$, with A normal and decomposed as above, then A_V is close to scalar,*

$$\|A_V - \lambda \mathbb{1}_V\| \leq nr,$$

and the off-diagonal blocks of B are bounded,

$$\|B_{\bar{V}V}\| \leq n\epsilon/2.$$

Proof. Given that A is normal, we can see that A_V is approximately scalar due to the bound between eigenvalues in I:

$$\|A_V - \lambda \mathbb{1}_V\| = \max_{i \in I} |\lambda_i - \lambda| \leq nr.$$

Next, using the fact that the operator norm dominates any component of a matrix, we can simply evaluate the relevant component of the commutator to bound elements of B:

$$\begin{aligned}
\epsilon &\geq \left\|[A, B]\right\| \\
&\geq |\langle i| [A, B] |j\rangle| \\
&= |\langle i| [AB - BA] |j\rangle| \\
&= |\lambda_i - \lambda_j| \cdot |\langle i|B|j\rangle| \,.
\end{aligned}$$

In the eigenbasis of A, the components of $B_{\bar{V}V}$ correspond to $\langle i|B|j\rangle$ for $i \notin I$, $j \in I$. By construction of I we have that $|\lambda_i - \lambda_j| > r$, and so

$$|\langle i|B|j\rangle| \leq \frac{\epsilon}{|\lambda_i - \lambda_j|} < \frac{\epsilon}{r}.$$

This implies therefore that $\|B_{\bar{V}V}\|_{\max} < \epsilon/r$, where $\|\cdot\|_{\max}$ denotes the elementwise max-norm. Using the fact that the operator norm exceeds the max-norm by at most the square root of the number of elements, we get

$$\|B_{\bar{V}V}\| \leq \|B_{\bar{V}V}\|_{\max} \times \sqrt{\dim V \times \dim \bar{V}}.$$

Given that $\dim V + \dim \bar{V} = n$, we have that $\dim V \times \dim \bar{V} \leq n^2/4$, and so

$$\|B_{\bar{V}V}\| < n\epsilon/2r \,.$$

□

Using these bounds, we can now put bounds on an approximate shared eigenvector. Imposing normality on both matrices, we can even impose the stricter requirement that both of the approximate eigenvalues are in fact exact eigenvalues.

Theorem 4.1 (Shared approximate eigenvector)**.** *Suppose that A and B are $n \times n$ matrices, such that A is normal and $\|[A, B]\| \leq \epsilon$. For any λ which is an eigenvalue of A, there exists a normalised $|u\rangle$ and μ such that*

$$\|A|u\rangle - \lambda|u\rangle\| \,, \|B|u\rangle - \mu|u\rangle\| \leq n\sqrt{\epsilon/2}.$$

If B is also normal, then for any λ which is an eigenvalue of A, there exists a ν which is also an eigenvalue of B and normalised $|w\rangle$ such that

$$\|A|w\rangle - \lambda|w\rangle\| \,, \|B|w\rangle - \nu|w\rangle\| \leq n\sqrt{\epsilon}.$$

Proof. Take $|u\rangle$ to be a right eigenvector of B_{VV} (contained within V), of eigenvalue μ. By Lemma 4.15, this then gives that the relevant errors with respect to A and B behave as:

$$\begin{aligned}\|A|u\rangle - \lambda|u\rangle\| &= \|A_V|u\rangle - \lambda|u\rangle\| \\ &= \|(A_V - \lambda\mathbb{1})|u\rangle\| \\ &\leq \|A_V - \lambda\mathbb{1}\| \\ &\leq nr\end{aligned} \qquad \begin{aligned}\|B|u\rangle - \mu|u\rangle\| &= \|B_{VV}|u\rangle - \mu|u\rangle + B_{\bar{V}V}|u\rangle\| \\ &= \|B_{\bar{V}V}|u\rangle\| \\ &\leq \|B_{\bar{V}V}\| \\ &< n\epsilon/2r\,.\end{aligned}$$

If we now let $r = \sqrt{\epsilon/2}$, we get the stated overall bound of $n\sqrt{\epsilon/2}$.

For the case of both matrices being normal, we can show that any approximate eigenvalue must lie near an exact eigenvalue. Taking $|w\rangle$ once again to be a right eigenvector of B_{VV} with eigenvalue ν' (for a different value of r to $|u\rangle$), we can see that

$$\big\|B|w\rangle - \nu'|w\rangle\big\| \leq n\epsilon/2r \qquad \Longrightarrow \qquad \langle w|(B-\nu')^\dagger(B-\nu)|w\rangle \leq n^2\epsilon^2/4r^2.$$

As $(B-\nu')^\dagger(B-\nu')$ is positive semi-definite, the existence of such a $|w\rangle$ implies $(B-\nu')^\dagger(B-\nu')$ possesses an eigenvalue at most $n^2\epsilon^2/4r^2$. By the normality of B, this implies in turn that B contains an eigenvalue ν such that $|\nu - \nu'| \leq n\epsilon/2r$. Using this we can see that the error with respect to B gains a factor of 2

$$\|B|w\rangle - \nu|w\rangle\| \leq \big\|B|w\rangle - \nu'|w\rangle\big\| + \big|\nu - \nu'\big| \leq n\epsilon/r.$$

Now taking $r = \sqrt{\epsilon}$, we find the stated bound of $n\sqrt{\epsilon}$. □

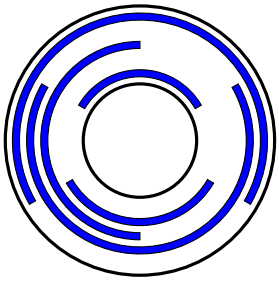
4.B An algorithm for the certifiable degeneracy of a twisted pair

In this appendix we sketch how, for a given pair of parameters α and δ, we can calculate the minimum possible dimension of unitaries u and v such that $\|[u,v]_\alpha\| \leq \delta$. Lemmas 4.5 and 4.6 give that for all $j \in \mathbb{Z}$, there exists an eigenvalue $e^{i\phi_j}$ of u such that

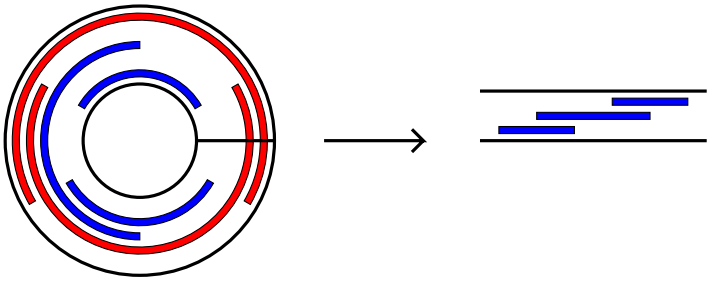
$$|\phi_j - 2\pi\alpha j| \leq \cos^{-1}(1 - |j|\,\delta).$$

The question now is to find the minimum number of eigenvalues such that at least one lies in each of the above arcs. This is known as the *transversal number*, and can be efficiently calculated by a greedy algorithm [4.59]. We now sketch this algorithm for the example parameters $\alpha = 1/4$ and $\delta = 1/2$ (indicated by the turquoise dot in Figure 4.3).

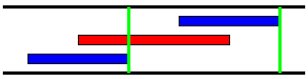
The first thing to note is that these arcs are trivial for $\delta\,|j| \geq 2$, in that they are the entire unit circle. For this reason we need only consider a finite number of arcs for $j = -\lfloor 2/\delta \rfloor, \ldots, \lfloor 2/\delta \rfloor$. In our case this corresponds $j = -3, \ldots, 3$. Below we have drawn these non-trivial arcs, omitting the trivial $j = 0$ arc.



Next we note that the $j = 0$ arc is simply a point, implying that u must contain a $+1$ eigenvalue. Given this, any arc containing $+1$ can be thrown away (indicated in red below), allowing us to unfold our arcs on a circle into intervals on a line.



We then take the intervals to be sorted by end-point. Considering each interval in order, we place an eigenvalue at the end of each interval as necessary, indicated as a green line below. Note that any interval which already contains an included eigenvalue when we arrive at it can be ignored, indicated by the red interval below.



Including the already found eigenvalue at $+1$, this gives us the minimum number of points necessary to satisfy each arc. Applying the algorithm for a large number of points, we can plot the certified degeneracy as in Figure 4.3.

Bibliography

4.1 H. Weyl, "Das asymptotische Verteilungsgesetz der Eigenwerte linearer partieller Differentialgleichungen (mit einer Anwendung auf die Theorie der Hohlraumstrahlung)," *Mathematische Annalen* **71**, 4, 441–469, (1911).

4.2 P. Rosenthal, "Are Almost Commuting Matrices Near Commuting Matrices?," *The American Mathematical Monthly* **76**, 925, (1969).

4.3 P. R. Halmos, "Some unsolved problems of unknown depth about operators on Hilbert space," *Proceedings of the Royal Society of Edinburgh, Section: A Mathematics* **76**, 1, 67–76, (1976).

4.4 S. Szarek, "On almost commuting Hermitian operators," *Rocky Mountain Journal of Mathematics* **20**, 583–591, (1990).

4.5 K. R. Davidson, "Almost commuting hermitian matrices," *Mathematica Scandinavica* **56**, 222, (1985).

4.6 I. D. Berg and K. R. Davidson, "Almost commuting matrices and a quantitative version of the Brown-Douglas-Fillmore theorem," *Acta Mathematica* **166**, 121–161, (1991).

4.7 H. Lin, "Almost commuting selfadjoint matrices and applications," in *Operator algebras and their applications*, 193–223, Fields Institute Communications, Vol. 13, (1997).

4.8 I. Kachkovskiy and Y. Safarov, "Distance to normal elements in C*-algebras of real rank zero," *Journal of the American Mathematical Society* **29**, 61–80, arXiv:1403.2021, (2015).

4.9 P. Friis and M. Rørdam, "Almost commuting self-adjoint matrices - a short proof of Huaxin Lin's theorem," *Journal für die reine und angewandte Mathematik* **1996**, 479, 121–131, (1996).

4.10 M. B. Hastings, "Making Almost Commuting Matrices Commute," *Communications in Mathematical Physics* **291**, 321–345, arXiv:0808.2474, (2009).

4.11 D. Voiculescu, "Asymptotically commuting finite rank unitary operators without commuting approximants," *Acta Scientiarum Mathematicarum* **45**, 429–431, (1983).

4.12 R. Exel and T. Loring, "Almost commuting unitary matrices," in *Proceedings of the American Mathematical Society*, 106, 913–913, (1989).

4.13 R. Exel and T. A. Loring, "Invariants of almost commuting unitaries," *Journal of Functional Analysis* **95**, 364–376, (1991).

4.14 M.-D. Choi, "Almost Commuting Matrices Need not be Nearly Commuting," *Proceedings of the American Mathematical Society* **102**, 529, (1988).

4.15 H. Lin, "Almost Commuting Unitaries and Classification of Purely Infinite Simple C*-Algebras," *Journal of Functional Analysis* **24**, 1–24, (1998).

4.16 T. J. Osborne, "Almost commuting unitaries with spectral gap are near commuting unitaries," arXiv:0809.0602, (2008).

4.17 T. A. Loring and A. P. W. Sørensen, "Almost commuting orthogonal matrices," *Journal of Mathematical Analysis and Applications* **420**, 1051–1068, arXiv:1311.4575, (2014).

4.18 X.-G. Wen, "Topological Order: From Long-Range Entangled Quantum Matter to a Unified Origin of Light and Electrons," *ISRN Condensed Matter Physics* **2013**, 1–20, arXiv:1210.1281, (2013).

4.19 J. Preskill, Quantum Error Correction. Cambridge University Press, Cambridge, (2013).

4.20 B. J. Brown, D. Loss, J. K. Pachos, C. N. Self, and J. R. Wootton, "Quantum memories at finite temperature," *Reviews of Modern Physics* **88**, 045005, arXiv:1411.6643, (2016).

4.21 L. Landau, E. Lifshitz, V. Berestetskii, and L. Pitaevskii, Course of Theoretical Physics. Pergamon Press, (1951).

4.22 B. Brown, S. T. Flammia, and N. Schuch, "Computational difficulty of computing the density of states," *Physical Review Letters* **107**, arXiv:1010.3060, (2011).

4.23 C. T. Chubb and S. T. Flammia, "Computing the Degenerate Ground Space of Gapped Spin Chains in Polynomial Time," 1–32, arXiv:1502.06967, (2015).

4.24 Y. Huang, "A polynomial-time algorithm for approximating the ground state of 1D gapped Hamiltonians," arXiv:1406.6355v3, (2014).

4.25 J. Yang and H.-k. Du, "A note on commutativity up to a factor of bounded operators," *Proceedings of the American Mathematical Society* **132**, 1713–1721, (2004).

4.26 B. C. Hall, Quantum Theory for Mathematicians, 267 of *Graduate Texts in Mathematics*. Springer New York, (2013).

4.27 J. Rosenberg, "A selective history of the Stone-von Neumann theorem," in *Contemporary Mathematics*, 331–353, (2004).

4.28 M. Said, "Almost Commuting Elements in Non-Commutative Symmetric Operator Spaces," (2014).

4.29 S. Eilers and R. Exel, "Finite dimensional representations of the soft torus," *Proceedings of the American Mathematical Society* **130**, 727–732, arXiv:math/9810165, (1998).

4.30 R. Exel, "The soft torus and applications to almost commuting matrices," *Pacific Journal of Mathematics* **160**, 207–217, (1993).

4.31 L. Babai and K. Friedl, "Approximate representation theory of finite groups," in *Proceedings 32nd Annual Symposium of Foundations of Computer Science*, 733–742, IEEE Comput. Soc. Press, (1991).

4.32 K. Friedl, "Near-representations of finite groups," (2003).

4.33 C. Moore and A. Russell, "Approximate Representations, Approximate Homomorphisms, and Low-Dimensional Embeddings of Groups," *SIAM Journal on Discrete Mathematics* **29**, 182–197, arXiv:1009.6230, (2015).

4.34 A. R. Bernstein, "Almost Eigenvectors for Almost Commuting Matrices," *SIAM Journal on Applied Mathematics* **21**, 232–235, (1971).

4.35 J. C. Bridgeman, S. T. Flammia, and D. Poulin, "Detecting Topological Order with Ribbon Operators," arXiv:1603.02275, (2016).

4.36 R. Bhatia, Matrix Analysis, 169 of *Graduate Texts in Mathematics*. Springer New York, (1997).

4.37 R. Schatten, Norm Ideals of Completely Continuous Operators. Springer Berlin, Heidelberg, (1960).

4.38 J. Von Neuman, "Some matrix inequalities and metrization of matric spaces," *Tomsk Univ. Rev.* **1**, 286–299, (1937).

4.39 A. J. Hoffman and H. W. Wielandt, "The variation of the spectrum of a normal matrix," *Duke Mathematical Journal* **20**, 37–39, (1953).

4.40 S. Bravyi, M. B. Hastings, and S. Michalakis, "Topological quantum order: Stability under local perturbations," *Journal of Mathematical Physics* **51**, 093512, arXiv:1001.0344, (2010).

4.41 S. Michalakis and J. Pytel, "Stability of Frustration-Free Hamiltonians," *Communications in Mathematical Physics* **322** (2), 277–302 arXiv:1109.1588, (2010).

4.42 M. B. Hastings and X.-G. Wen, "Quasiadiabatic continuation of quantum states: The stability of topological ground-state degeneracy and emergent gauge invariance," *Physical Review B* **72**, 045141, arXiv:cond-mat/0503554, (2005).

4.43 A. Kitaev, "Fault-tolerant quantum computation by anyons," *Annals of Physics* **303**, 2–30, arXiv:quant-ph/9707021, (2003).

4.44 Y. Hu, Y. Wan, and Y.-S. Wu, "Twisted Quantum Double Model of Topological Phases in Two–Dimension," *Physical Review B* **87**, 125114, arXiv:1211.3695, (2012).

4.45 R. Dijkgraaf and E. Witten, "Topological gauge theories and group cohomology," *Communications in Mathematical Physics* **129**, 393–429, (1990).

4.46 M. A. Levin and X.-G. Wen, "String-net condensation: A physical mechanism for topological phases," *Physical Review B* **02139**, arXiv:cond-mat/0404617, (2004).

4.47 J. Haah, "An Invariant of Topologically Ordered States Under Local Unitary Transformations," *Communications in Mathematical Physics* **342**, 771–801, arXiv:1407.2926, (2016).

4.48 J. C. Bridgeman and C. T. Chubb, "Hand-waving and Interpretive Dance: An Introductory Course on Tensor Networks," arXiv:1603.03039, (2016).

4.49 R. Orús, "A practical introduction to tensor networks: Matrix product states and projected entangled pair states," *Annals of Physics* **349**, 117–158, arXiv:1306.2164, (2014).

4.50 M. A. Nielsen and I. L. Chuang, Quantum Computation and Quantum Information 10th Anniversary Edition. Cambridge University Press, Cambridge, (2011).

4.51 E. Knill, R. Laflamme, and L. Viola, "Theory of Quantum Error Correction for General Noise," *Physical Review Letters* **84**, 2525–2528, arXiv:quant-ph/9604034, (2000).

4.52 R. Blume-Kohout, H. K. Ng, D. Poulin, and L. Viola, "Information-preserving structures: A general framework for quantum zero-error information," *Physical Review A* **82**, 062306, arXiv:1006.1358, (2010).

4.53 K. Glashoff and M. M. Bronstein, "Almost-commuting matrices are almost jointly diagonalizable," arXiv:1305.2135, (2013).

4.54 F. Gygi, J.-L. Fattebert, and E. Schwegler, "Computation of Maximally Localized Wannier Functions using a simultaneous diagonalization algorithm," *Computer Physics Communications* **155**, 1–6, (2003).

4.55 M. Congedo, B. Afsari, A. Barachant, and M. Moakher, "Approximate Joint Diagonalization and Geometric Mean of Symmetric Positive Definite Matrices," *Public Library of Science*, **10**, 4, (2015).

4.56 D. Eynard, A. Kovnatsky, and M. M. Bronstein, "Laplacian colormaps: a framework for structure-preserving color transformations," *Computer Graphics Forum* **33**, 215–224, arXiv:1311.0119, (2014).

4.57 C. Pearcy and A. Shields, "Almost commuting matrices," *Journal of Functional Analysis* **33**, 332–338, (1979).

4.58 S. H. Schanuel, "A combinatorial problem of Shields and Pearcy," *Proceedings of the American Mathematical Society* **65**, 185–185, (1977).

4.59 A. Gyárfás and J. Lehel, "Covering and coloring problems for relatives of intervals," *Discrete Mathematics* **55**, 167–180, (1985).

Part II

Resource-error trade-offs

Chapter 5

Moderate deviation analysis for classical communication over quantum channels

Christopher T. Chubb[1], Vincent Y.F. Tan[2,3], and Marco Tomamichel[1,4]

[1] *Centre for Engineered Quantum Systems, School of Physics, University of Sydney, Sydney, Australia*

[2] *Department of Electrical and Computer Engineering, National University of Singapore, Singapore, Singapore*

[3] *Department of Mathematics, National University of Singapore, Singapore, Singapore*

[4] *Centre for Quantum Software and Information, University of Technology Sydney, Sydney, Australia*

Communications in Mathematical Physics **355**, 3 (2017)

Doi: 10.1007/s00220-017-2971-1

ArXiv: 1701.03114

Abstract

We analyse families of codes for classical data transmission over quantum channels that have both a vanishing probability of error and a code rate approaching capacity as the code length increases. To characterise the fundamental tradeoff between decoding error, code rate and code length for such codes we introduce a quantum generalisation of the moderate deviation analysis proposed by Altŭg and Wagner as well as Polyanskiy and Verdú. We derive such a tradeoff for classical-quantum (as well as image-additive) channels in terms of the channel capacity and the channel dispersion, giving further evidence that the latter quantity characterises the necessary backoff from capacity when transmitting finite blocks of classical data. To derive these results we also study asymmetric binary quantum hypothesis testing in the moderate deviations regime. Due to the central importance of the latter task, we expect that our techniques will find further applications in the analysis of other quantum information processing tasks.

5.1 Introduction

The goal of information theory is to find the fundamental limits imposed on information processing and transmission by the laws of physics. One of the early breakthroughs in quantum information theory was the characterisation of the capacity of a classical-quantum (c-q) channel to transmit classical information by Holevo [5.1, 5.2] and Schumacher–Westmoreland [5.3]. The *classical capacity* of a quantum channel is defined as the maximal rate (in bits per channel use) at which we can transmit information such that the decoding error vanishes asymptotically as the length of the code increases. However, for many practical applications there are natural restrictions on the code length imposed, for example, by limitations on how much quantum information can be processed coherently. Therefore it is crucial to go beyond the asymptotic treatment and understand the intricate tradeoff between decoding error probability, code rate and code length.

For this purpose, we will study families of codes that have *both* a rate approaching the capacity and an error probability that vanishes asymptotically as the code length n increases. The following tradeoff relation gives a rough illustration of our main result: if the code rate approaches capacity as $\Theta(n^{-t})$ for some $t \in (0, 1/2)$, then the decoding error cannot be smaller than $\exp(-\Theta(n^{1-2t}))$. In fact, we will show that the constants implicit in the Θ notation are determined by a second channel parameter beyond the capacity, called the *channel dispersion*. We will also show that this relation is tight, i.e., there exist families of codes achieving equality asymptotically.

Our work thus complements previous work on the boundary cases corresponding to $t \in \{0, 1/2\}$. The error exponent (or reliability function) of c-q channels (see, e.g., Refs. [5.4–5.6]) corresponds to the case $t = 0$ where the rate is bounded away from capacity and the error probability vanishes exponentially in n. This is also called the *large deviations* regime. Moreover, the second-order asymptotics of c-q channels were evaluated by Tomamichel and Tan [5.7]. They correspond to the case $t = 1/2$ where the rate approaches capacity as $\Theta(n^{-1/2})$ and the error probability is non-vanishing. This is also called the *small deviations* regime.

In the present work, we consider the entire regime in between, which is dubbed the *moderate deviation* regime.[1] The different parameter regimes are illustrated in Fig. 5.1.

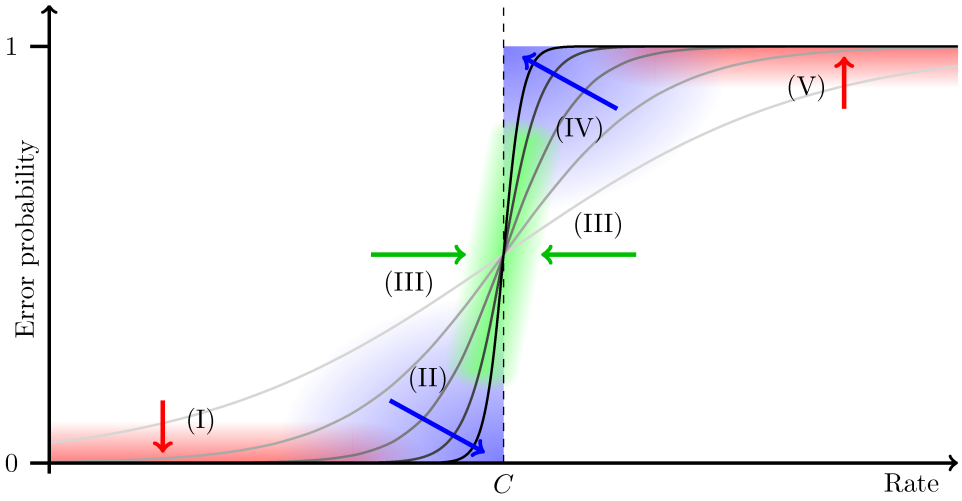
Main results. Before we present our main results, let us introduce the notion of a *moderate sequence* of real numbers, $\{x_n\}_n$ for $n \in \mathbb{N}$, whose defining properties are that $x_n \searrow 0$ and $\sqrt{n}\, x_n \to +\infty$ as $n \to \infty$.[2] Our two main results concern binary asymmetric quantum hypothesis testing and c-q channel coding.

1. The first result, presented in detail in Sect. 5.3, concerns binary quantum hypothesis testing between a pair of quantum states ρ and σ. We show that for any moderate sequence x_n, there exists a sequence of tests $\{Q_n\}_n$ such that the two kinds of errors satisfy

$$\operatorname{Tr} \rho^{\otimes n}(1 - Q_n) = e^{-nx_n^2} \text{ and } \operatorname{Tr} \sigma^{\otimes n} Q_n = \exp\Big(- n\Big(D(\rho\|\sigma) - \sqrt{2V(\rho\|\sigma)}\, x_n + o(x_n)\Big)\Big), \tag{5.1}$$

[1] In the technical analysis, we are considering moderate deviations from the mean of a sum of independent log-likelihood ratios, thus justifying the name emanating from statistics [5.8, Theorem 3.7.1].

[2] As mentioned above an archetypical moderate sequence is $x_n = \Theta(n^{-t})$ for some $t \in (0, \frac{1}{2})$. The boundary cases are not included — in fact $t = 0$ requires a large deviation analysis whereas $t = \frac{1}{2}$ requires a small deviation analysis.



	(I)	(II)	(III)	(IV)	(V)
regime	error exponent	moderate deviation (below capacity)	constant error (second-order)	moderate deviation (above capacity)	strong converse exponent
error prob.	$\exp(-\Theta(n))$	$\exp(-o(n))$ & $\omega(1)$	$\Theta(1)$	$1-\exp(-o(n))$ & $1-\omega(1)$	$1-\exp(-\Theta(n))$
code rate	$C-\Theta(1)$	$C-o(1)$ & $C-\omega(n^{-\frac{1}{2}})$	$C-\Theta(n^{-\frac{1}{2}})$	$C+o(1)$ & $C+\omega(n^{-\frac{1}{2}})$	$C+\Theta(1)$

Figure 5.1: The figure shows the optimal error probability as a function of the rate, for different block lengths. Darker lines correspond to longer block lengths, and the capacity is denoted by C. The table shows the asymptotics in each region, as the blocklength n goes to infinity. The functions of n implicit in the Θ, o, and ω notation are assumed to be positive-valued.

and another sequence of tests $\{Q'_n\}_n$ such that the errors satisfy

$$\operatorname{Tr}\rho^{\otimes n}(1-Q'_n) = 1-e^{-nx_n^2} \text{ and } \operatorname{Tr}\sigma^{\otimes n}Q'_n = \exp\Big(-n\Big(D(\rho\|\sigma)+\sqrt{2V(\rho\|\sigma)}\,x_n+o(x_n)\Big)\Big)\,, \tag{5.2}$$

where $D(\cdot\|\cdot)$ and $V(\cdot\|\cdot)$ denote the relative entropy [5.9] and relative entropy variance [5.10,5.11], respectively. (The reader is referred to the next section for formal definitions of all concepts discussed here.) Most importantly, we show that both of these tradeoffs are in fact optimal.

2. The main result, covered in Sect. 5.4, concerns coding over a memoryless classical-quantum channel $\mathcal{W}$. Let us denote by $M^*(\mathcal{W};n,\epsilon)$ the maximum $M\in\mathbb{N}$ such that there exists a code transmitting one out of M messages over n uses of the channel $\mathcal{W}$ such that the average probability of error does not exceed ϵ. For any sequence of tolerated error probabilities $\{\epsilon_n\}_n$ vanishing sub-exponentially with $\epsilon_n = e^{-nx_n^2}$, we find that

$$\frac{1}{n}\log M^*(\mathcal{W};n,\epsilon_n) = C(\mathcal{W}) - \sqrt{2V_{\min}(\mathcal{W})}\,x_n + o(x_n)\,, \tag{5.3}$$

$$\frac{1}{n}\log M^*(\mathcal{W};n,1-\epsilon_n) = C(\mathcal{W}) + \sqrt{2V_{\max}(\mathcal{W})}\,x_n + o(x_n)\,, \tag{5.4}$$

	asymmetric binary hypothesis testing	channel coding	quantum hypothesis testing	classical-quantum channel coding
large deviation (<)	[5.18]	[5.19, 5.20]	[5.5, 5.21]	unknown[3]
moderate deviation (<)	[5.22]	[5.12, 5.13]	*this work*	*this work*
small deviation	[5.23]	[5.23–5.25]	[5.10, 5.11]	[5.7]
moderate deviation (>)	*this work*	*this work*	*this work*	*this work*
large deviation (>)	[5.26, 5.27]	[5.28, 5.29]	[5.30, 5.31]	[5.32]

Table 5.1: Exposition of related work on finite resource analysis of hypothesis testing and channel coding problems. The rows correspond to different parameter regimes, labelled by the deviation from the critical rate (i.e., the relative entropy for hypothesis testing and the capacity for channel coding problems).

where $C(\cdot)$ denotes the channel capacity and $V_{\min}(\cdot)$ and $V_{\max}(\cdot)$ denote the minimal and maximal channel dispersion as defined in Ref. [5.7], respectively. This result holds very generally for channels with arbitrary input alphabet and without restriction on the channel dispersion, strengthening also the best known results for classical channels. Moreover, as in Ref. [5.7], this generality allows us to lift the above result to a statement about coding classical information over image-additive quantum channels and general channels as long as the encoders are restricted to prepare separable states.

Since quantum hypothesis testing underlies many other quantum information processing tasks such as entanglement-assisted classical communication as well as private and quantum communication, we expect that our techniques will have further applications in quantum information theory.

Related work. For classical channels, Altŭg and Wagner [5.12] first established the best decay rate of the average error probability for a class of discrete memoryless channels (DMCs) when the code rate approaches capacity at a rate slower than $\Theta(n^{-1/2})$. Shortly after the conference version of Ref. [5.12], Polyanskiy and Verdú [5.13] (see Refs. [5.14, 5.15] for applications of their techniques to the quantum setting) relaxed some of the conditions on the class of DMCs and also established the moderate deviations asymptotics for other important classical channels such as the additive white Gaussian noise channel. The other main contributions to the analysis of hypothesis testing, channel coding, quantum hypothesis testing, and c-q channel coding in the different parameter regimes are summarised in Table 5.1.

From a technical perspective the moderate deviations regime can be approached via a refined large deviations analysis (as was done in Ref. [5.12]) or via a variation of second-order analysis via the information spectrum method (as was proposed in Ref. [5.13]). In our work, we mostly follow the latter approach, interspersed with ideas from large deviation theory. In particular, we build on bounds from one-shot information theory by Wang and Renner [5.16] and use techniques developed for the second-order asymptotics in Ref. [5.7]. In concurrent work, Cheng and Hsieh [5.17] provide a moderate deviation analysis for c-q channels via a refined error exponent analysis. Their result holds for c-q channels with finite input alphabets and their techniques are complementary to ours.

[3] In contrast to classical channels a tight characterisation of the error exponent of c-q channels remains elusive to date even for high rates. See, e.g., Refs. [5.4–5.6] for partial progress.

5.2 Preliminaries

5.2.1 Notation and classical coding over quantum channels

Let $\mathcal{H}$ be a finite-dimensional Hilbert space and denote by $\mathcal{S} := \{\rho \in \mathcal{H} \mid \operatorname{Tr}\rho = 1, \rho \geq 0\}$ the quantum states on $\mathcal{H}$. We take $\exp(\cdot)$ and $\log(\cdot)$ to be in an arbitrary but compatible base (such that they are inverses), and denote the natural logarithm by $\ln(\cdot)$. For convenience, we will consider the dimension of this Hilbert space to be a fixed constant, and omit any dependence constants may have on this dimension. For $\rho, \sigma \in \mathcal{S}$ we write $\rho \ll \sigma$ if the support of ρ is contained in the support of σ. For any closed subset $\mathcal{S}_\circ \subseteq \mathcal{S}$, we will denote by $\mathcal{P}(\mathcal{S}_\circ)$ the space of probability distributions supported on $\mathcal{S}_\circ$. We equip $\mathcal{S}$ with the trace metric $\delta_{\mathrm{Tr}}(\rho, \rho') := \frac{1}{2}\|\rho - \rho'\|_1$ and $\mathcal{P}(\mathcal{S})$ with a weak-convergence metric[4] δ_{wc}, such that both are compact metric spaces with

$$f : \mathcal{S} \to \mathbb{R} \text{ continuous} \qquad \Longrightarrow \qquad \mathbb{P} \mapsto \int \mathrm{d}\mathbb{P}(\rho)\, f(\rho) \text{ continuous.} \tag{5.5}$$

We will use the *cumulative standard normal distribution* function Φ is defined as

$$\Phi(a) := \int_{-\infty}^{a} \frac{1}{\sqrt{2\pi}}\, e^{-\frac{x^2}{2}}\, \mathrm{d}x. \tag{5.6}$$

Following Ref. [5.7], we consider a general *classical-quantum channel* $\mathcal{W} : \mathcal{X} \to \mathcal{S}$ where $\mathcal{X}$ is any set (without further structure). We define the *image of the channel* as the set $\operatorname{im}\mathcal{W} \subset \mathcal{S}$ of all quantum states ρ such that $\rho = \mathcal{W}(x)$ for some $x \in \mathcal{X}$. For convenience we assume that our Hilbert space satisfies

$$\mathcal{H} = \operatorname*{Span}_{\rho \in \operatorname{im}\mathcal{W}} \operatorname{supp}(\rho) \tag{5.7}$$

such that $\sigma > 0$ is equivalent to $\rho \ll \sigma$ for all $\rho \in \operatorname{im}\mathcal{W}$.

For $M, n \in \mathbb{N}$, an (n, M)-*code* for a classical-quantum channel $\mathcal{W}$ is comprised of an encoder and a decoder. The *encoder* is a map $E : \{1, 2, \ldots, M\} \to \mathcal{X}^n$ and the *decoder* is a positive operator-valued measure $\{D_m\}_{m=1}^M$ on $\mathcal{H}^{\otimes n}$. Moreover, an (n, M, ϵ)-*code* is an (n, M)-*code* that satisfies

$$\frac{1}{M}\sum_{m=1}^{M} \operatorname{Tr}\left(\bigotimes_{i=1}^{n} \mathcal{W}\big(E_i(m)\big) D_m\right) \geq 1 - \epsilon\,, \tag{5.8}$$

i.e. the average probability of error does not exceed ϵ. The finite blocklength achievable region for a channel $\mathcal{W}$ is the set of triples (n, M, ϵ) for which there exists an (n, M, ϵ)-*code* on $\mathcal{W}$. We are particularly interested in the boundary

$$M^*(\mathcal{W}; n, \epsilon) := \max\left\{M \in \mathbb{N} : \exists \text{ a } (n, M, \epsilon)\text{-code for } \mathcal{W}\right\}. \tag{5.9}$$

Specifically we are going to be concerned with the behaviour of the *maximum rate*, which is defined as $R^*(\mathcal{W}; n, \epsilon) := \frac{1}{n}\log M^*(\mathcal{W}; n, \epsilon)$.

[4] An example of which is the induced Lévy–Prokhorov metric (see, e.g., Section 6 and Theorem 6.4 in Ref. [5.33]).

5.2.2 Channel parameters

An important parameter of a channel is the largest rate such that there exists a code of vanishing error probability in the large blocklength limit. This critical rate is known as the *capacity* of a channel $C(\mathcal{W})$, which is defined as

$$C(\mathcal{W}) := \inf_{\epsilon>0} \liminf_{n\to\infty} R^*(\mathcal{W}; n, \epsilon). \tag{5.10}$$

For classical-quantum channels there exists a *strong converse* bound, which states that the capacity described the asymptotic rate not just for vanishing error probability, but those for non-zero fixed error probabilities as well [5.34, 5.35]. Together with the original channel coding theorem [5.1, 5.36], this yields

$$\lim_{n\to\infty} R^*(\mathcal{W}; n, \epsilon) = C(\mathcal{W}) \quad \text{for all } \epsilon \in (0,1). \tag{5.11}$$

In essence the strong converse tells us that the capacity entirely dictates the asymptotic behaviour of the maximum rate at a fixed error probability. How quickly the rate approaches this asymptotic value for arbitrarily low and high error probabilities are described by the channel *min-dispersion* $V_{\min}(\mathcal{W})$ and *max-dispersion* $V_{\max}(\mathcal{W})$, which are defined respectively as

$$V_{\min}(\mathcal{W}) := \inf_{\epsilon>0} \limsup_{n\to\infty} \left(\frac{C(\mathcal{W}) - R^*(\mathcal{W}; n, \epsilon)}{\Phi^{-1}(\epsilon)/\sqrt{n}} \right)^2, \tag{5.12}$$

$$V_{\max}(\mathcal{W}) := \sup_{\epsilon<1} \limsup_{n\to\infty} \left(\frac{C(\mathcal{W}) - R^*(\mathcal{W}; n, \epsilon)}{\Phi^{-1}(\epsilon)/\sqrt{n}} \right)^2. \tag{5.13}$$

As with the strong converse, the min and max-dispersions also describe the dispersion at other fixed error probabilities [5.7]:

$$\lim_{n\to\infty} \left(\frac{C(\mathcal{W}) - R^*(\mathcal{W}; n, \epsilon)}{\Phi^{-1}(\epsilon)/\sqrt{n}} \right)^2 = \begin{cases} V_{\min}(\mathcal{W}) & \epsilon \in (0, 1/2) \\ V_{\max}(\mathcal{W}) & \epsilon \in (1/2, 1) \end{cases}. \tag{5.14}$$

5.2.3 Information quantities

Classically, for two distributions P and Q, the *relative entropy* $D(P\|Q)$ and *relative entropy variance* $V(P\|Q)$ are both defined as the mean and variance of the log-likelihood ratio $\log\big(P/Q\big)$ with respect to the distribution P. In the non-commutative case, for $\rho, \sigma \in \mathcal{S}$ with $\rho \ll \sigma$, these definitions are generalised as [5.9–5.11]

$$D(\rho\|\sigma) := \operatorname{Tr} \rho \left(\log \rho - \log \sigma\right), \tag{5.15}$$

$$V(\rho\|\sigma) := \operatorname{Tr} \rho \big(\log \rho - \log \sigma - D\left(\rho\|\sigma\right) \cdot \mathrm{id}\big)^2. \tag{5.16}$$

If $\rho \not\ll \sigma$ both quantities are set to $+\infty$.

Following Ref. [5.7], for a closed set $\mathcal{S}_\circ \in \mathcal{S}$, the *divergence radius*[5] $\chi(\mathcal{S}_\circ)$ is given by

$$\chi(\mathcal{S}_\circ) = \sup_{\mathbb{P}\in\mathcal{P}(\mathcal{S}_\circ)} \int \mathrm{d}\mathbb{P}(\rho)\, D\left(\rho \,\middle\|\, \int \mathrm{d}\mathbb{P}(\rho')\, \rho'\right). \tag{5.17}$$

[5] Whilst Equation (5.17) characterises the divergence radius, we will mostly rely on a more useful form presented in Definition 5.2.

where $\mathcal{P}(\mathcal{S}_\circ)$ denotes the space of distributions on $\mathcal{S}_\circ$. If we let $\Pi(\mathcal{S}_\circ)$ denote the distributions which achieve the above supremum, we also define the *minimal and maximal peripheral variance*, $v_{\min}(\mathcal{S}_\circ)$ and $v_{\max}(\mathcal{S}_\circ)$, as

$$v_{\min}(\mathcal{S}_\circ) := \inf_{\mathbb{P}\in\Pi(\mathcal{S}_\circ)} \int \mathrm{d}\mathbb{P}(\rho)\, V\left(\rho \,\middle\|\, \int \mathrm{d}\mathbb{P}(\rho')\,\rho'\right), \tag{5.18}$$

$$v_{\max}(\mathcal{S}_\circ) := \sup_{\mathbb{P}\in\Pi(\mathcal{S}_\circ)} \int \mathrm{d}\mathbb{P}(\rho)\, V\left(\rho \,\middle\|\, \int \mathrm{d}\mathbb{P}(\rho')\,\rho'\right). \tag{5.19}$$

For the image of a quantum channel, the above three information quantities correspond exactly to the three previously defined channel parameters [5.7]. Specifically, for $\mathcal{S}_\circ = \overline{\operatorname{im}\mathcal{W}}$, we have

$$C(\mathcal{W}) = \chi(\mathcal{S}_\circ), \qquad V_{\min}(\mathcal{W}) = v_{\min}(\mathcal{S}_\circ), \qquad V_{\max}(\mathcal{W}) = v_{\max}(\mathcal{S}_\circ). \tag{5.20}$$

5.2.4 Moderate deviation tail bounds

We now discuss the relevant tail bounds we will require in the moderate deviation regime. Let $\{X_{i,n}\}_{i\leq n}$ be independent zero-mean random variables, and define the average variance as

$$V_n := \frac{1}{n}\sum_{i=1}^{n} \operatorname{Var}[X_{i,n}]. \tag{5.21}$$

Recall that a sequence $\{t_n\}_n$ is moderate if $x_n \searrow 0$ and $\sqrt{n}x_n \to +\infty$ as $n\to\infty$. Given certain bounds on the moments and cumulants of these variables, which we will make explicit below, we will see that the probability that the average variable $\frac{1}{n}\sum_{i=1}^n X_{i,n}$ deviates from the mean by a moderate sequence $\{t_n\}_n$ decays asymptotically as

$$\ln \Pr\left[\frac{1}{n}\sum_{i=1}^{n} X_{i,n} \geq t_n\right] = -\big(1+o(1)\big)\frac{nt_n^2}{2V_n}. \tag{5.22}$$

Lemma 5.1 (Moderate deviation lower bound)**.** *If there exist constants $\nu > 0$ and τ such that $\nu \leq V_n$ and*

$$\frac{1}{n}\sum_{i=1}^{n} \mathbb{E}\left[|X_{i,n}|^3\right] \leq \tau \tag{5.23}$$

for all n, then for any $\eta > 0$ there exists a constant $N(\{t_i\},\nu,\tau,\eta)$ such that, for all $n \geq N$, the probability of a moderate deviation is lower bounded as

$$\ln \Pr\left[\frac{1}{n}\sum_{i=1}^{n} X_{i,n} \geq t_n\right] \geq -(1+\eta)\frac{nt_n^2}{2V_n}. \tag{5.24}$$

Lemma 5.2 (Moderate deviation upper bound)**.** *If there exists a constant γ such that*

$$\frac{1}{n}\sum_{i=1}^{n} \sup_{s\in[0,1/2]} \left|\frac{\mathrm{d}^3}{\mathrm{d}s^3} \ln \mathbb{E}\left[e^{sX_{i,n}}\right]\right| \leq \gamma, \tag{5.25}$$

for all n, then for any $\eta > 0$ there exists a constant $N(\{t_i\}, \gamma, \eta)$ such that, for all $n \geq N$, the probability of a moderate deviation is upper bounded as

$$\ln \Pr\left[\frac{1}{n}\sum_{i=1}^{n} X_{i,n} \geq t_n\right] \leq -\frac{nt_n^2}{2V_n + \eta}. \tag{5.26}$$

If V_n has a uniform lower bound, then as $\eta \searrow 0$ the above two bounds sandwich together, giving the two-sided asymptotic scaling of Eq. 5.22. In this case we can see that

$$\sigma\left[\frac{1}{n}\sum_{i=1}^{n} X_i\right] = \sqrt{V_n/n} = \Theta(1/\sqrt{n}) \qquad \text{and} \qquad \sqrt{\frac{1}{n}\sum_{i=1}^{n}\sigma^2\left[X_i\right]} = \sqrt{V_n} = \Theta(1), \tag{5.27}$$

where $\sigma\left[\cdot\right]$ denotes the standard deviation. If we interpret the standard deviation as setting the 'length-scale' on which a distribution decays, then the above two quantities—the deviation of the average, and average[6] of the deviation—set the length-scales of small and large deviation bounds respectively. Using this intuition, we can generalise moderate deviation bounds to give tight two-sided bounds for distributions with arbitrary normalisation, in which V_n is no longer bounded. To do this we will tail bound for deviations which are moderate, *in units of* $\sqrt{V_n}$.

Corollary 5.3 (Dimensionless moderate deviation bound)**.** *If there exists a γ such that*

$$\frac{1}{nV_n^{3/2}}\sum_{i=1}^{n}\sup_{s\in[0,1/2]}\left|\frac{\mathrm{d}^3}{\mathrm{d}s^3}\ln\mathbb{E}\left[e^{sX_{i,n}}\right]\right| \leq \gamma, \tag{5.28}$$

for all n, then there exists a constant $N(\{t_i\}, \gamma)$ such that, for all $n \geq N$, we have the two-sided bound

$$-(1+\eta)\frac{nt_n^2}{2} \leq \ln \Pr\left[\frac{1}{n}\sum_{i=1}^{n} X_{i,n} \geq t_n\sqrt{V_n}\right] \leq -(1-\eta)\frac{nt_n^2}{2}. \tag{5.29}$$

We present proofs of these lemmas in Section 5.A.

5.2.5 Reversing lemma

Intuitively one might expect that moderate deviation bounds can be 'reversed' e.g. that the bound on the probability given the deviation (see Lemmas 5.1 and 5.2) of the form

$$\lim_{n\to\infty}\frac{V_n}{nt_n^2}\ln\Pr\left[\frac{1}{n}\sum_{i=1}^{n} X_i \geq t_n\right] = -\frac{1}{2}, \tag{5.30}$$

is equivalent to a bound on the deviation given the probability

$$\lim_{n\to\infty}\frac{1}{t_n}\inf\left\{t\in\mathbb{R}\,\middle|\,\frac{V_n}{nt_n^2}\ln\Pr\left[\frac{1}{n}\sum_{i=1}^{n} X_i \geq t\right] \leq -\frac{1}{2}\right\} = \tag{5.31}$$

We will now see that such an ability to 'reverse' moderate deviation bounds is generic. We do this by considering two quantities A and B defined on the same domain, and considering the infimum value of each quantity for a fixed value of the other.

[6] More specifically the root-mean-square

Lemma 5.4 (Reversing Lemma)**.** *Let* $\{A_i\}_i$ *and* $\{B_i\}_i$ *be sequences of real functions with* $\inf_t A_i(t) \leq 0$ *and* $\inf_t B_i(t) \leq 0$ *for all* i*. If we define*

$$\hat{A}_n(b) := \inf_t \{A_n(t) | B_n(t) \leq b\} \qquad \text{and} \qquad \hat{B}_n(a) := \inf_t \{B_n(t) | A_n(t) \leq a\} \tag{5.32}$$

then

$$\lim_{n\to\infty} \frac{\hat{A}_n(b_n)}{b_n} = 1, \quad \forall \{b_n\} \text{ moderate} \qquad \iff \qquad \lim_{n\to\infty} \frac{\hat{B}_n(a_n)}{a_n} = 1, \quad \forall \{a_n\} \text{ moderate.} \tag{5.33}$$

Proof. See Section 5.B. □

5.3 Hypothesis testing

Whilst the divergence radius characterises the channel capacity, one-shot channel bounds are characterised by a quantity known as the *ϵ-hypothesis testing divergence* [5.16]. As the name suggests, as well as being relevant to one-shot channel coding bounds, the hypothesis testing divergence also has an operational interpretation in the context of hypothesis testing of quantum states. We will start by considering a moderate deviation analysis of this quantity.

5.3.1 Hypothesis testing divergence

Consider a hypothesis testing problem, in which ρ and σ correspond to the null and alternative hypotheses respectively. A test between these hypotheses will take the form of a POVM $\{Q, I - Q\}$, where $0 \leq Q \leq I$. For a given Q, the type-I and type-II error probabilities are given by

$$\alpha(Q; \rho, \sigma) := \operatorname{Tr}(I - Q)\rho, \qquad \beta(Q; \rho, \sigma) := \operatorname{Tr} Q\sigma. \tag{5.34}$$

If we define the smallest possible type-II error given a type-I error at most ϵ as

$$\beta_\epsilon(\rho\|\sigma) := \min_{0\leq Q\leq \mathbb{I}} \{\beta(Q; \rho, \sigma) \,|\, \alpha(Q; \rho, \sigma) \leq \epsilon\}, \tag{5.35}$$

then the ϵ-hypothesis testing divergence is defined as

$$D_{\mathrm{h}}^\epsilon(\rho\|\sigma) := -\log \frac{\beta_\epsilon(\rho\|\sigma)}{1-\epsilon}. \tag{5.36}$$

We note that the denominator of $1 - \epsilon$ follows the normalisation in [5.37] such that $D_{\mathrm{h}}^\epsilon(\rho\|\rho) = 0$ for all ρ.

An obvious extension of this hypothesis problem is to the case of n copies of each state, i.e. a hypothesis test between $\rho^{\otimes n}$ and $\sigma^{\otimes n}$, or more generally between two product states $\otimes_{i=1}^n \rho_i$ and $\otimes_{i=1}^n \sigma_i$. A second-order analysis of the ϵ-hypothesis testing divergence for a non-vanishing ϵ was given in [5.10, 5.11].

Theorem 5.1 (Moderate deviation of the hypothesis testing divergence)**.** *For any moderate sequence* $\{a_n\}_n$ *and states* $\{\rho_n\}_n$ *and* $\{\sigma_n\}_n$ *such that both* $\lambda_{\min}(\sigma_i)$ *and* $V(\rho_i\|\sigma_i)$ *are both uniformly bounded*

away from zero, the ϵ_n- and $(1-\epsilon_n)$-hypothesis testing divergences of non-uniform product states for $\epsilon_n = e^{-na_n^2}$ scale as

$$\frac{1}{n} D_{\mathrm{h}}^{\epsilon_n}\left(\bigotimes_{i=1}^n \rho_i \middle\| \bigotimes_{i=1}^n \sigma_i\right) = D_n - \sqrt{2V_n}\, a_n + o(a_n), \tag{5.37}$$

$$\frac{1}{n} D_{\mathrm{h}}^{1-\epsilon_n}\left(\bigotimes_{i=1}^n \rho_i \middle\| \bigotimes_{i=1}^n \sigma_i\right) = D_n + \sqrt{2V_n}\, a_n + o(a_n), \tag{5.38}$$

where $D_n := \frac{1}{n}\sum_{i=1}^n D(\rho_i\|\sigma_i)$ and $V_n := \frac{1}{n}\sum_{i=1}^n V(\rho_i\|\sigma_i)$. More specifically for any ρ and σ such that $\rho \ll \sigma$, the hypothesis testing divergences of uniform product states scale as

$$\frac{1}{n} D_{\mathrm{h}}^{\epsilon_n}\left(\rho^{\otimes n}\middle\|\sigma^{\otimes n}\right) = D(\rho\|\sigma) - \sqrt{2V(\rho\|\sigma)}\, a_n + o(a_n), \tag{5.39}$$

$$\frac{1}{n} D_{\mathrm{h}}^{1-\epsilon_n}\left(\rho^{\otimes n}\middle\|\sigma^{\otimes n}\right) = D(\rho\|\sigma) + \sqrt{2V(\rho\|\sigma)}\, a_n + o(a_n). \tag{5.40}$$

In Sect. 5.3.3 we will bound the regularised hypothesis testing divergences towards the relative entropy (the *inward bound*), and in Sect. 5.3.4 we will bound them away (the *outward bound*).

Remark 1. *For sequences ϵ_n bounded away from zero and one the second-order expansion in Refs. [5.10, 5.11] yields*

$$\frac{1}{n} D_{\mathrm{h}}^{\epsilon_n}\left(\rho^{\otimes n}\middle\|\sigma^{\otimes n}\right) = D(\rho\|\sigma) + \sqrt{\frac{V(\rho\|\sigma)}{n}}\,\Phi^{-1}(\epsilon_n) + O\left(\frac{\log n}{n}\right), \tag{5.41}$$

where Φ denotes the cumulative distribution function of the standard normal. As already pointed out in Ref. [5.13], for small ϵ_n we have $\Phi^{-1}(\epsilon_n) \approx \sqrt{-2\ln\epsilon_n}$. Ignoring all higher order terms, the substitution $\epsilon_n = e^{-na_n^2}$ into (5.41) then recovers the expression in (5.39). In this sense the two results thus agree at the boundary between small and moderate deviations.

Remark 2. *A similar argument can be sketched at the boundary between moderate and large deviations. The quantum Hoeffding bound [5.5, 5.21] states that if $\frac{1}{n} D_{\mathrm{h}}^{\epsilon_n}(\rho^{\otimes n}\|\sigma^{\otimes n}) \leq D(\rho\|\sigma) - r$ for some small $r > 0$ then ϵ_n drops exponentially in n with the exponent given by*

$$\sup_{0\leq\alpha<1} \frac{\alpha-1}{\alpha}\Big[D(\rho\|\sigma) - r - D_\alpha(\rho\|\sigma)\Big], \tag{5.42}$$

where $D_\alpha(\rho\|\sigma)$ is the Petz' quantum Rényi relative entropy [5.38]. For sufficiently small r, the expression in (5.42) attains its supremum close to $\alpha = 1$ and we can thus approximate $D_\alpha(\rho\|\sigma) \approx D(\rho\|\sigma) + \frac{\alpha-1}{2} V(\rho\|\sigma)$ by its Taylor expansion [5.39]. Evaluating this approximate expression yields

$$\epsilon_n = e^{-n\frac{r^2}{2V(\rho\|\sigma)}}. \tag{5.43}$$

up to leading order in r. Substituting $r = \sqrt{2V(\rho\|\sigma)}a_n$ into (5.43) then recovers (5.39). An essentially equivalent argument is also applicable to the strong converse exponent derived in Ref. [5.30].

5.3.2 Nussbaum–Szkoła distributions

To allow us to apply a moderate deviation analysis to the quantum hypothesis testing divergence, we leverage the results of Ref. [5.10] which allow us to reduce the hypothesis testing divergence of quantum states to a quantity known as the information spectrum divergence of certain classical distributions, known as the Nussbaum–Szkoła distributions.

Definition 5.1 (Nussbaum–Szkoła distributions [5.40])**.** *The* Nussbaum–Szkoła distributions *for a pair of states ρ and σ are given by*

$$P^{\rho,\sigma}(a,b) = r_a \, |\langle \phi_a | \psi_b \rangle|^2 \qquad \text{and} \qquad Q^{\rho,\sigma}(a,b) = s_b \, |\langle \phi_a | \psi_b \rangle|^2 \tag{5.44}$$

where the states are eigendecomposed as $\rho = \sum_a r_a \, |\phi_a\rangle\langle\phi_a|$ and $\sigma = \sum_b s_b \, |\psi_b\rangle\langle\psi_b|$.

The power of the Nussbaum–Szkoła distributions lies in their ability to reproduce both the divergence and variance of the underlying quantum states

$$D(\rho\|\sigma) = D(P^{\rho,\sigma}\|Q^{\rho,\sigma}), \qquad \text{and} \qquad V(\rho\|\sigma) = V(P^{\rho,\sigma}\|Q^{\rho,\sigma}). \tag{5.45}$$

As well as capturing these asymptotic quantities, the hypothesis testing relative entropy, which arises one-shot channel coding bounds, can also be captured by the Nussbaum–Szkoła distributions. Specifically this is done via the *information spectrum divergence*, which is defined for two classical distributions P and Q by a tail bound on the log-likelihood ratio as

$$D_{\mathrm{s}}^{\epsilon}(P\|Q) := \sup\left\{ R \;\middle|\; \Pr_{X\leftarrow P}\left[\log \frac{P(X)}{Q(X)} \leq R\right] \leq \epsilon \right\}. \tag{5.46}$$

Inserting the Nussbaum–Skzoła distributions, we find that the (classical) information spectrum divergence approximates the (quantum) hypothesis testing divergence.

Lemma 5.5 (Thm. 14, Ref. [5.10])**.** *There exists a universal constant K such that for any states ρ and σ with $\lambda_{\min}(\sigma) \geq \lambda$ and $\epsilon < 1/2$, we find that $D_{\mathrm{h}}^{\epsilon}(\rho\|\sigma)$ is bounded as*

$$D_{\mathrm{h}}^{\epsilon}(\rho\|\sigma) \leq D_{\mathrm{s}}^{2\epsilon}(P^{\rho,\sigma}\|Q^{\rho,\sigma}) + \log \frac{1-\epsilon}{\epsilon^3(1-2\epsilon)} + \log K \lceil \ln(1/\lambda) \rceil \tag{5.47}$$

$$D_{\mathrm{h}}^{\epsilon}(\rho\|\sigma) \geq D_{\mathrm{s}}^{\epsilon/2}(P^{\rho,\sigma}\|Q^{\rho,\sigma}) - \log \frac{1}{\epsilon(1-\epsilon)} - \log K \lceil \ln(1/\lambda) \rceil, \tag{5.48}$$

and $D_{\mathrm{h}}^{1-\epsilon}(\rho\|\sigma)$ is bounded as

$$D_{\mathrm{h}}^{1-\epsilon}(\rho\|\sigma) \leq D_{\mathrm{s}}^{1-\epsilon/2}(P^{\rho,\sigma}\|Q^{\rho,\sigma}) + \log \frac{1-\epsilon/2}{\epsilon^4} + \log K \lceil \ln(1/\lambda) \rceil \tag{5.49}$$

$$D_{\mathrm{h}}^{1-\epsilon}(\rho\|\sigma) \geq D_{\mathrm{s}}^{1-2\epsilon}(P^{\rho,\sigma}\|Q^{\rho,\sigma}) - \log \frac{1}{\epsilon^2} - \log K \lceil \ln(1/\lambda) \rceil. \tag{5.50}$$

As the information spectrum divergence is defined in terms of a tail bound, we will bound these quantities using the moderate deviation tail bounds of Sect. 5.2.4. To do this, we will start by showing that the log-likelihood ratio of Nussbaum–Skzoła distributions is sufficiently well behaved, specifically that its cumulant generating function has bounded derivatives.

Lemma 5.6 (Bounded cumulants)**.** *For $\lambda > 0$, there exists constants $C_k(\lambda)$ such that the cumulant generating function $h(t) := \ln \mathbb{E}\left[e^{tZ}\right]$ of the log-likelihood ratio $Z := \log P^{\rho,\sigma}/Q^{\rho,\sigma}$ for $\lambda_{\min}(\sigma) \geq \lambda$ is smooth and has uniformly bounded derivatives in a neighbourhood of the origin*

$$\sup_{|t|\leq 1/2} \left| \frac{\partial^k}{\partial t^k} h(t) \right| \leq C_k. \tag{5.51}$$

We present a proof of this lemma in Section 5.C.

5.3.3 Inward bound

Proposition 5.7 (Inward bound)**.** *For any constants $\lambda, \eta > 0$, there exists a constant $N(\{a_i\}, \lambda, \eta)$ such that, for $n \geq N$, the hypothesis testing divergence can be bounded for any states $\{\rho_i\}_i$ and $\{\sigma_i\}_i$ with $\lambda_{\min}(\sigma_i) \geq \lambda$ as*

$$\frac{1}{n} D_{\mathrm{h}}^{\epsilon_n}\left(\bigotimes_{i=1}^n \rho_i \middle\| \bigotimes_{i=1}^n \sigma_i\right) \geq D_n - \sqrt{2V_n} a_n - \eta a_n, \tag{5.52}$$

$$\frac{1}{n} D_{\mathrm{h}}^{1-\epsilon_n}\left(\bigotimes_{i=1}^n \rho_i \middle\| \bigotimes_{i=1}^n \sigma_i\right) \leq D_n + \sqrt{2V_n}\, a_n + \eta a_n. \tag{5.53}$$

where $D_n := \frac{1}{n}\sum_{i=1}^n D(\rho_i\|\sigma_i)$ and $V_n := \frac{1}{n}\sum_{i=1}^n V(\rho_i\|\sigma_i)$.

Proof. Firstly, let Z_i be the log-likelihood ratios

$$Z_i := \log \frac{P^{\rho_i,\sigma_i}(A_i, B_i)}{Q^{\rho_i,\sigma_i}(A_i, B_i)}, \qquad (A_i, B_i) \leftarrow P^{\rho_i,\sigma_i}. \tag{5.54}$$

In terms of these log-likelihood ratios, the lower and upper bound on the ϵ_n- and $(1-\epsilon_n)$-hypothesis testing divergences respectively from Lemma 5.5 become

$$D_{\mathrm{h}}^{\epsilon_n}\left(\bigotimes_{i=1}^n \rho_i \middle\| \bigotimes_{i=1}^n \sigma_i\right) \geq \sup\left\{R \;\middle|\; \Pr\left[\sum_{i=1}^n Z_i \leq R\right] \leq \epsilon_n/2\right\} - \log\frac{1}{\epsilon_n(1-\epsilon_n)} - \log Kn\lceil \ln(1/\lambda)\rceil, \tag{5.55}$$

$$D_{\mathrm{h}}^{1-\epsilon_n}\left(\bigotimes_{i=1}^n \rho_i \middle\| \bigotimes_{i=1}^n \sigma_i\right) \leq \sup\left\{R \;\middle|\; \Pr\left[\sum_{i=1}^n Z_i \leq R\right] \leq 1-\epsilon_n/2\right\} + \log\frac{1-\epsilon_n/2}{\epsilon_n^4} + \log Kn\lceil \ln(1/\lambda)\rceil. \tag{5.56}$$

Recalling that $\epsilon_n := e^{-na_n^2}$, we can see that in both cases the error terms scale like $\Theta(na_n^2)$ and $\Theta(\log n)$ respectively, which are both $o(na_n)$. As such, there must exist an $N_1(\{a_i\}, \lambda, \eta)$ such that, for $n \geq N_1$, these error terms are bounded by $\eta n a_n/2$ as

$$\frac{1}{n} D_{\mathrm{h}}^{\epsilon_n}\left(\bigotimes_{i=1}^n \rho_i \middle\| \bigotimes_{i=1}^n \sigma_i\right) \geq \frac{1}{n}\sup\left\{R \;\middle|\; \Pr\left[\sum_{i=1}^n Z_i \leq R\right] \leq \epsilon_n/2\right\} - \eta a_n/2, \tag{5.57}$$

$$\frac{1}{n} D_{\mathrm{h}}^{1-\epsilon_n}\left(\bigotimes_{i=1}^n \rho_i \middle\| \bigotimes_{i=1}^n \sigma_i\right) \leq \frac{1}{n}\sup\left\{R \;\middle|\; \Pr\left[\sum_{i=1}^n Z_i \leq R\right] \leq 1-\epsilon_n/2\right\} + \eta a_n/2. \tag{5.58}$$

Next we want to apply the tail bounds of Sect. 5.2.4. To this end, we will start by defining zero-mean variables $X_i := Z_i - D(\rho_i\|\sigma_i)$. In terms of these variables, the above bounds take the form

$$\frac{1}{n}D_{\mathrm{h}}^{\epsilon_n}\left(\bigotimes_{i=1}^n \rho_i \middle\| \bigotimes_{i=1}^n \sigma_i\right) \geq D_n - \inf\left\{t \;\middle|\; \Pr\left[\frac{1}{n}\sum_{i=1}^n(-X_i) \geq t\right] \leq \epsilon_n/2\right\} - \eta a_n/2, \tag{5.59}$$

$$\frac{1}{n}D_{\mathrm{h}}^{1-\epsilon_n}\left(\bigotimes_{i=1}^n \rho_i \middle\| \bigotimes_{i=1}^n \sigma_i\right) \leq D_n + \inf\left\{t \;\middle|\; \Pr\left[\frac{1}{n}\sum_{i=1}^n(+X_i) \geq t\right] \leq \epsilon_n/2\right\} + \eta a_n/2. \tag{5.60}$$

By Lemma 5.6 there exists constants $\bar{V}(\lambda)$ and $\gamma(\lambda)$, such that $V_i \leq \bar{V}$ and

$$\sup_{t\in[0,1/2]}\left|\frac{\mathrm{d}^3}{\mathrm{d}s^3}\ln\mathbb{E}\left[e^{s(\pm X_i)}\right]\right| \leq \gamma \tag{5.61}$$

for all i. If we let $t_n := \left(\sqrt{2V_n} + \eta/2\right)a_n$, then Lemma 5.2 gives an $N_2(\{a_i\},\lambda,\eta)$ such that, for $n \geq N_2$, the tail probability is bounded as

$$\ln\Pr\left[\frac{1}{n}\sum_{i=1}^n(\pm X_i) \geq t_n\right] \leq \frac{-nt_n^2}{2V_n + \eta^2/3} \tag{5.62}$$

$$\leq -\frac{\left(\sqrt{2V_n} + \eta/2\right)^2}{2V_n + \eta^2/5}na_n^2 \tag{5.63}$$

$$\leq -\frac{2V_n + \eta^2/4}{2V_n + \eta^2/5}na_n^2 \tag{5.64}$$

$$= -\left(1 + \frac{\eta^2}{40V_n + 4\eta^2}\right)na_n^2 \tag{5.65}$$

$$\leq -\left(1 + \frac{\eta^2}{40\bar{V} + 4\eta^2}\right)na_n^2. \tag{5.66}$$

As $\eta^2/(40\bar{V} + 4\eta)$ is a constant and $na_n^2 \to \infty$, there must exist a constant $N_3(\{a_i\},\lambda,\eta)$ such $n \geq N_3$ implies

$$\ln\Pr\left[\frac{1}{n}\sum_{i=1}^n(\pm X_i) \geq t_n\right] \leq -na_n^2 - 1 = \ln(\epsilon_n/2), \tag{5.67}$$

and therefore that

$$\inf\left\{t \;\middle|\; \Pr\left[\frac{1}{n}\sum_{i=1}^n(\pm X_i) \geq t\right] \leq \epsilon_n/2\right\} \leq t_n. \tag{5.68}$$

Putting everything together, we get that for any $n \geq N(\{a_i\},\lambda,\eta) := \max\{N_1, N_2, N_3\}$ we have

$$\frac{1}{n}D_{\mathrm{h}}^{\epsilon_n}\left(\bigotimes_{i=1}^n \rho_i \middle\| \bigotimes_{i=1}^n \sigma_i\right) \geq D_n - \sqrt{2V_n}\,a_n - \eta a_n, \tag{5.69}$$

$$\frac{1}{n}D_{\mathrm{h}}^{1-\epsilon_n}\left(\bigotimes_{i=1}^n \rho_i \middle\| \bigotimes_{i=1}^n \sigma_i\right) \leq D_n + \sqrt{2V_n}\,a_n + \eta a_n. \tag{5.70}$$

as required. □

5.3.4 Outward bound

Proposition 5.8 (Outward bound)**.** *For any constants $\lambda, \eta > 0$, there exists a constant $N(\{a_i\}, \lambda, \eta)$ such that, for $n \geq N$, the hypothesis testing divergence can be bounded for any states $\{\rho_i\}_i$ and $\{\sigma_i\}_i$ with $\lambda_{\min}(\sigma_i) \geq \lambda$ as*

$$\frac{1}{n} D_{\mathrm{h}}^{\epsilon_n}\left(\bigotimes_{i=1}^n \rho_i \middle\| \bigotimes_{i=1}^n \sigma_i\right) \leq D_n + \eta a_n, \tag{5.71}$$

$$\frac{1}{n} D_{\mathrm{h}}^{1-\epsilon_n}\left(\bigotimes_{i=1}^n \rho_i \middle\| \bigotimes_{i=1}^n \sigma_i\right) \geq D_n - \eta\, a_n. \tag{5.72}$$

where $D_n := \frac{1}{n}\sum_{i=1}^n D(\rho_i\|\sigma_i)$. Moreover, if we let $V_n := \frac{1}{n}\sum_{i=1}^n V(\rho_i\|\sigma_i)$, and there also exists a constant $\nu > 0$ such that $V_i \geq \nu$ for all i, then there exists an $N'(\{a_i\}, \lambda, \nu, \eta)$ such that, for $n \geq N'$, the hypothesis testing divergence is more tightly bounded as

$$\frac{1}{n} D_{\mathrm{h}}^{\epsilon_n}\left(\bigotimes_{i=1}^n \rho_i \middle\| \bigotimes_{i=1}^n \sigma_i\right) \leq D_n - \sqrt{2V_n} a_n + \eta a_n, \tag{5.73}$$

$$\frac{1}{n} D_{\mathrm{h}}^{1-\epsilon_n}\left(\bigotimes_{i=1}^n \rho_i \middle\| \bigotimes_{i=1}^n \sigma_i\right) \geq D_n + \sqrt{2V_n}\, a_n - \eta a_n. \tag{5.74}$$

Proof. Similar to Proposition 5.7, we will start by taking the upper and lower bounds on the ϵ_n- and $(1-\epsilon_n)$-hypothesis testing divergences respectively from Lemma 5.5. This gives that there exists an $N_1(\{a_i\}, \lambda, \eta)$ such that, for $n \geq N_1$, we have

$$\frac{1}{n} D_{\mathrm{h}}^{\epsilon_n}\left(\bigotimes_{i=1}^n \rho_i \middle\| \bigotimes_{i=1}^n \sigma_i\right) \leq D_n - \inf\left\{ t \;\middle|\; \Pr\left[\frac{1}{n}\sum_{i=1}^n (-X_i) \geq t\right] \leq 2\epsilon_n \right\} + \eta a_n/2, \tag{5.75}$$

$$\frac{1}{n} D_{\mathrm{h}}^{1-\epsilon_n}\left(\bigotimes_{i=1}^n \rho_i \middle\| \bigotimes_{i=1}^n \sigma_i\right) \geq D_n + \inf\left\{ t \;\middle|\; \Pr\left[\frac{1}{n}\sum_{i=1}^n (+X_i) \geq t\right] \leq 2\epsilon_n \right\} - \eta a_n/2. \tag{5.76}$$

where $X_i := Z_i - D(\rho_i\|\sigma_i)$.

Firstly, applying Chebyshev's inequality two standard deviations below the mean gives us that

$$\Pr\left[\frac{1}{n}\sum_{i=1}^n (\pm X_i) \geq -2\sqrt{V_n/n}\right] \geq 3/4 \geq 2\epsilon_n, \tag{5.77}$$

and so we conclude that

$$\inf\left\{ t \;\middle|\; \Pr\left[\frac{1}{n}\sum_{i=1}^n (\pm X_i) \geq t\right] \leq 2\epsilon_n \right\} \geq -2\sqrt{V_n/n}. \tag{5.78}$$

By Lemma 5.6, V_n must be bounded $V_n \leq \bar{V}(\lambda)$, and thus $\sqrt{V_n/n} = \mathcal{O}(1/\sqrt{n}) = o(a_n)$. As such, there must exist an $N_2(\{a_i\}, \lambda, \eta)$ such that $n \geq N_2$ implies $2\sqrt{V_n/n} \leq \eta a_n/2$. Inserting this tail bound, we

get that for any $n \geq N(\{a_i\}, \lambda, \eta) := \max\{N_1, N_2\}$ that

$$\frac{1}{n} D_{\mathrm{h}}^{\epsilon_n}\left(\bigotimes_{i=1}^{n} \rho_i \middle\| \bigotimes_{i=1}^{n} \sigma_i\right) \leq D_n - \eta a_n, \tag{5.79}$$

$$\frac{1}{n} D_{\mathrm{h}}^{1-\epsilon_n}\left(\bigotimes_{i=1}^{n} \rho_i \middle\| \bigotimes_{i=1}^{n} \sigma_i\right) \geq D_n + \eta a_n, \tag{5.80}$$

as required.

If there also exists an $\nu > 0$ such that $V_i \geq \nu$, then we can use a more refined moderate deviation bound. Specifically, Lemma 5.6 gives us a bound on the absolute third moment of X_i, which allows us to apply Lemma 5.1. If we let $t_n := (\sqrt{2V_n} - \eta/2)a_n$ and assume $\eta < \sqrt{8\nu}$ such that $\{t_n\}_n$ is moderate, then this gives us that there exists an $N_3(\{a_i\}, \lambda, \nu, \eta)$ such that, for any $n \geq N_3$, the tail probabilities are bounded

$$\ln \Pr\left[\frac{1}{n}\sum_{i=1}^{n}(\pm X_i) \geq t_n\right] \geq -\left(1 + \eta/\sqrt{2\bar{V}}\right)\frac{nt_n^2}{2V_n} \tag{5.81}$$

$$\geq -\frac{\left(1 + \eta/\sqrt{2\bar{V}}\right)\left(\sqrt{2V_n} - \eta/2\right)^2}{2V_n} n a_n^2 \tag{5.82}$$

$$\geq -\frac{\left(1 + \eta/\sqrt{2V_n}\right)\left(\sqrt{2V_n} - \eta/2\right)^2}{2V_n} n a_n^2 \tag{5.83}$$

$$\geq -\left(1 - \frac{5\eta^2}{8\bar{V}}\right) n a_n^2. \tag{5.84}$$

Once again, the second term in the parenthesis is a non-zero constant, and thus there must exist an $N_4(\{a_i\}, \lambda, \eta)$ such that

$$\log \Pr\left[\frac{1}{n}\sum_{i=1}^{n}(\pm X_i) \geq t_n\right] \geq -n a_n^2 + 1 = \ln 2\epsilon_n, \tag{5.85}$$

allowing us to conclude $\Pr\left[\frac{1}{n}\sum_{i=1}^{n}(\pm X_i) \geq t_n\right] \geq 2\epsilon_n$, and therefore

$$\inf\left\{t \;\middle|\; \Pr\left[\frac{1}{n}\sum_{i=1}^{n} X_i \geq t\right] \leq 2\epsilon_n\right\} \geq t_n. \tag{5.86}$$

Inserting this into the above bounds, we find that for any $n \geq N'(\{a_i\}, \lambda, \nu, \eta) := \max\{N_1, N_3, N_4\}$, we have the desired final bound

$$\frac{1}{n} D_{\mathrm{h}}^{\epsilon_n}\left(\bigotimes_{i=1}^{n} \rho_i \middle\| \bigotimes_{i=1}^{n} \sigma_i\right) \leq D_n - \sqrt{2V_n}\, a_n + \eta a_n, \tag{5.87}$$

$$\frac{1}{n} D_{\mathrm{h}}^{1-\epsilon_n}\left(\bigotimes_{i=1}^{n} \rho_i \middle\| \bigotimes_{i=1}^{n} \sigma_i\right) \geq D_n + \sqrt{2V_n}\, a_n - \eta a_n. \tag{5.88}$$

□

5.4 Channel Coding

We are now going to show how the above moderate deviation bounds can be applied to the capacity of a classical-quantum channel.

Theorem 5.2 (Moderate deviation of c-q channels)**.** *For any moderate sequence $\{a_n\}_n$ and memoryless c-q channel $\mathcal{W}$ with capacity $C(\mathcal{W})$ and min-dispersion $V_{\min}(\mathcal{W})$, being operated at error probability no larger than $\epsilon_n := e^{-na_n^2}$, the optimal rate deviates below the capacity as*

$$R^*\left(\mathcal{W}; n, \epsilon_n\right) = C(\mathcal{W}) - \sqrt{2V_{\min}(\mathcal{W})}\, a_n + o(a_n). \tag{5.89}$$

Conversely, if the channel has max-dispersion $V_{\max}$ and is operated at error probability no larger than $1-\epsilon_n$, then the optimal rate deviates above the capacity as

$$R^*\left(\mathcal{W}; n, 1-\epsilon_n\right) = C(\mathcal{W}) + \sqrt{2V_{\max}(\mathcal{W})}\, a_n + o(a_n). \tag{5.90}$$

If either the min- or max-dispersion is non-zero, an application of Lemma 5.4 gives an equivalent formulation in terms of the minimal error probability at a given rate.

Corollary 5.9. *For any moderate sequence $\{s_n\}_n$, the error probability for a code with min-dispersion $V_{\min} > 0$ deviating below capacity by s_n scales as*

$$\lim_{n\to\infty} \frac{1}{ns_n^2} \ln \epsilon^*(\mathcal{W}; n, C - s_n) = -\frac{1}{2V_{\min}}. \tag{5.91}$$

Similarly, for a code with max-dispersion $V_{\max} > 0$ deviating above capacity by s_n, the error probability scales

$$\lim_{n\to\infty} \frac{1}{ns_n^2} \ln\bigl(1 - \epsilon^*(\mathcal{W}; n, C + s_n)\bigr) = -\frac{1}{2V_{\max}}. \tag{5.92}$$

Remark 3. *Recall that our definition of c-q channels does not put any restriction on the input set. In particular, this set may be comprised of quantum states itself such that the c-q channel is just a representation of a quantum channel. Hence, as pointed out in Ref. [5.7], our results immediately also apply to classical communication over general image-additive channels [5.41] as well as classical communication over quantum channels with encoders restricted to prepare separable states. We refer the reader to Corollaries 6 and 7 of Ref. [5.7] for details.*

We will split the proof of Theorem 5.2 in two, in Sect. 5.4.1 we will prove a lower bound on the maximum rate ('achievability'), followed in Sect. 5.4.2 by a corresponding the upper bound ('optimality'). For the rest of this section, we will fix the channel $\mathcal{W}$, and omit any dependencies on $\mathcal{W}$ from here on for notational convenience.

5.4.1 Achievability

For achievability, we will use a lower bound on the ϵ-one-shot rate that is essentially due to Hayashi and Nagaoka [5.42] who analysed the coding problem using the information spectrum method.

Lemma 5.10 (Theorem 1 of Ref. [5.16]). *If we have a c-q channel which maps from a finite message space Y as $y \mapsto \rho^{(y)}$, then the maximum rate with error probability at most ϵ and $1-\epsilon$, $R^*(\epsilon)$ and $R^*(1-\epsilon)$ respectively, are lower bounded*

$$R^*(\epsilon) \geq \sup_{P_Y} D_{\mathrm{h}}^{\epsilon/2}\left(\pi_{YZ} \| \pi_Y \otimes \pi_Z\right) - \log \frac{8(2-\epsilon)}{\epsilon} \tag{5.93}$$

$$R^*(1-\epsilon) \geq \sup_{P_Y} D_{\mathrm{h}}^{1-2\epsilon}\left(\pi_{YZ} \| \pi_Y \otimes \pi_Z\right) - \log \frac{8(1-\epsilon)}{\epsilon} \tag{5.94}$$

where π_{YZ} is the joint state of the input and output, with inputs chosen according to the distribution P_Y

$$\pi_{YZ} := \sum_{y \in Y} P_Y(y) \left|y\right\rangle\!\left\langle y\right|_Y \otimes \rho_Z^{(y)}. \tag{5.95}$$

Proposition 5.11 (Channel coding: Achievability). *For any moderate sequence $\{a_n\}_n$ and error probability $\epsilon_n := e^{-na_n^2}$, the rate is at least*

$$R^*\left(n, \epsilon_n\right) \geq C - \sqrt{2V_{\min}} a_n + o(a_n). \tag{5.96}$$

Similarly, at error probability $1-\epsilon_n$, the rate is at least

$$R^*\left(n, 1-\epsilon_n\right) \geq C + \sqrt{2V_{\max}} a_n + o(a_n). \tag{5.97}$$

Proof. Let X be our, possibly infinite, message space. By Lemma 3 of Ref. [5.7], there exists a finite subset $Y \subseteq X$, and a distribution Q_Y thereon, such that $D(\rho\|\sigma) = C$ and $V(\rho\|\sigma) = V_{\min}$ for states

$$\rho := \sum_{y \in Y} Q_Y(y) \left|y\right\rangle\!\left\langle y\right| \otimes \rho^{(y)} \qquad \text{and} \qquad \sigma := \sum_{y \in Y} Q_Y(y) \left|y\right\rangle\!\left\langle y\right| \otimes \sum_{y' \in Y} Q_Y(y') \rho^{(y')}. \tag{5.98}$$

Clearly by restricting the message space we can only ever decrease the rate. By applying Lemma 5.10 to the restriction of the message space to Y, we can lower bound the maximum rate of the full code. Applying this reasoning to n memoryless applications of our channel we find

$$nR^*(n, \epsilon_n) \geq \sup_{P_{Y^n}} D_{\mathrm{h}}^{\epsilon_n/2}\left(\pi_{Y^n Z^n} \| \pi_{Y^n} \otimes \pi_{Z^n}\right) - \log \frac{8(2-\epsilon_n)}{\epsilon_n}. \tag{5.99}$$

Substituting in both the error probability, which is no larger than $\epsilon_n = e^{-na_n^2}$, and a product distribution $Q_{Y^n}(\vec{y}) := \prod_{i=1}^n Q_Y(y_i)$ then we get

$$R^*(n, \epsilon_n) \geq \frac{1}{n} D_{\mathrm{h}}^{\epsilon_n/2}\left(\rho^{\otimes n} \middle\| \sigma^{\otimes n}\right) + \mathcal{O}(a_n^2). \tag{5.100}$$

Applying Proposition 5.7, we get an overall bound on the rate of

$$R^*(n, \epsilon_n) \geq C - \sqrt{2V_{\min}}\, a_n + o(a_n). \tag{5.101}$$

If instead we were to take a distribution Q_Y such that $V(\rho\|\sigma) = V_{\max}$, then the same arguments would allow us to use Proposition 5.8 to analogously give

$$R^*(n, 1-\epsilon_n) \geq C + \sqrt{2V_{\max}}\, a_n + o(a_n). \tag{5.102}$$

□

5.4.2 Optimality

Similar to the second-order analysis of Ref. [5.7], we are going to do this by relating the capacity and one-shot maximum rates to geometric quantities known as the divergence radius and divergence centre.

Definition 5.2 (Divergence radius and centre)**.** *For some set of states $\mathcal{S}_0 \subseteq \mathcal{S}$, the* divergence radius $\chi(\mathcal{S}_0)$ *and* divergence centre $\sigma^*(\mathcal{S}_0)$ *are defined as*

$$\chi(\mathcal{S}_0) := \inf_{\sigma\in\mathcal{S}} \sup_{\rho\in\mathcal{S}_0} D(\rho\|\sigma), \qquad \sigma^*(\mathcal{S}_0) := \arg\min_{\sigma\in\mathcal{S}} \sup_{\rho\in\mathcal{S}_0} D(\rho\|\sigma). \tag{5.103}$$

*Similarly the ϵ-*hypothesis testing divergence radius $\chi^\epsilon_{\mathrm{h}}(\mathcal{S}_0)$ *is defined as*

$$\chi^\epsilon_{\mathrm{h}}(\mathcal{S}_0) := \inf_{\sigma\in\mathcal{S}} \sup_{\rho\in\mathcal{S}_0} D^\epsilon_h(\rho\|\sigma). \tag{5.104}$$

Whilst we have seen that the divergence radius captures the capacity of a channel, the ϵ-hypothesis testing divergence radius approximates the one-shot capacity.

Lemma 5.12 (Proposition 5 of [5.7])**.** *For $\mathcal{I} := \overline{\operatorname{im}\mathcal{W}}$, the maximum rate with error probability at most ϵ, $R^*(\epsilon)$, is upper bounded as*

$$R^*(\epsilon) \leq \chi^{2\epsilon}_{\mathrm{h}}(I) + \log\frac{2}{1-2\epsilon}. \tag{5.105}$$

Similarly for an error probability $1-\epsilon$, the maximum rate is upper bounded as

$$R^*(1-\epsilon) \leq \chi^{1-\epsilon/2}_{\mathrm{h}}(I) + \log\frac{2(2-\epsilon)}{\epsilon^2}. \tag{5.106}$$

If we take $\mathcal{I}_n := \overline{\operatorname{im}\mathcal{W}^{\otimes n}}$ to be the closure of the image of n uses of this channel, then we can extend this bound on the one-shot rate to the n-shot rate as

$$nR^*(n,\epsilon_n) \leq \chi^{2\epsilon_n}_{\mathrm{h}}(\mathcal{I}_n) + \log\frac{2}{1-\epsilon_n}, \tag{5.107}$$

$$nR^*(n,1-\epsilon_n) \leq \chi^{2\epsilon_n}_{\mathrm{h}}(\mathcal{I}_n) + \log\frac{2(2-\epsilon_n)}{\epsilon_n^2}. \tag{5.108}$$

As we are considering memoryless c-q channels, $\mathcal{I}_n$ simply consists of elementwise tensor products of $\mathcal{I}$

$$\mathcal{I}_n = \left\{ \bigotimes_{i=1}^n \rho_i \,\middle|\, \rho_i \in \mathcal{I} \right\}. \tag{5.109}$$

Once again we are going to take a_n to be an arbitrary moderate sequence, and $\epsilon_n := e^{-na_n^2}$. Expanding this out, this gives bounds on the rate of

$$R^*(n,\epsilon_n) \leq \inf_{\sigma^n} \sup_{\{\rho_i\}\subseteq\mathcal{I}} \frac{1}{n} D^{2\epsilon_n}_{\mathrm{h}}\left(\bigotimes_{i=1}^n \rho_i \middle\| \sigma^n\right) + \frac{1}{n}\log\frac{2}{1-\epsilon_n}, \tag{5.110}$$

$$R^*(n,1-\epsilon_n) \leq \inf_{\sigma^n} \sup_{\{\rho_i\}\subseteq\mathcal{I}} \frac{1}{n} D^{1-\epsilon_n/2}_{\mathrm{h}}\left(\bigotimes_{i=1}^n \rho_i \middle\| \sigma^n\right) + \frac{1}{n}\log\frac{2(2-\epsilon)}{\epsilon_n^2}. \tag{5.111}$$

A standard approach now is to pick a state σ^n, such that we can bound the above quantities for arbitrary sequences $\{\rho_i\}$ using the moderate deviation analysis of the hypothesis testing divergence presented in Sect. 5.3. To do this we need to consider two cases. The *high cases* are those in which the empirical relative entropy corresponding to $\{\rho_i\}_{i=1}^n$ is close to capacity, and the *low cases* are those in which the empirical relative entropy corresponding to $\{\rho_i\}_{i=1}^n$ is far from capacity. Specifically, for some constant γ that will be chosen later, the n which correspond to high and low cases are denoted by $H(\{\rho_i\},\gamma)$ and $L(\{\rho_i\},\gamma)$, respectively. They are defined as

$$H(\{\rho_i\},\gamma) := \left\{ n \,\middle|\, \frac{1}{n}\sum_{i=1}^n D(\rho_i\|\bar{\rho}_n) \geq C-\gamma \right\} \quad \text{and} \quad L(\{\rho_i\},\gamma) := \left\{ n \,\middle|\, \frac{1}{n}\sum_{i=1}^n D(\rho_i\|\bar{\rho}_n) < C-\gamma \right\} \tag{5.112}$$

such that $H(\{\rho_i\},\gamma)$ and $L(\{\rho_i\},\gamma)$ bipartition $\mathbb{N}$ for all γ.

Before employing a moderate deviation bound, we are going to construct a separable state σ^n that will allow us two different moderate deviation analyses for low and high sequences, such that we can obtain the required bounds in both cases. A convenient choice of σ^n would be $\sigma^n = \bar{\rho}_n^{\otimes n}$ where $\bar{\rho}_n := \frac{1}{n}\sum_{i=1}^n \rho_i$, but the order of the infimum and supremum require σ^n to be chosen to be independent of the sequence $\{\rho_i\}$. Instead we are going to construct σ^n from a mixture of states that lie in a covering of $\mathcal{S}$, and the divergence centre $\sigma^*(\mathcal{I})$.

The following lemma is based on a construction in Lemma II.4 of Ref. [5.43].

Lemma 5.13 (Lemma 18 of Ref. [5.7]). *For every $\delta \in (0,1)$, there exists a set $\mathcal{C}^\delta \subset \mathcal{S}$ of size*

$$\left|\mathcal{C}^\delta\right| \leq \left(\frac{20(2d+1)}{\delta}\right)^{2d^2} \left(\frac{8d(2d+1)}{\delta}+2\right)^{d-1} \leq \left(\frac{90d}{\delta^2}\right)^{2d^2} \tag{5.113}$$

such that, for every $\rho \in \mathcal{S}$ there exists a state $\tau \in \mathcal{C}^\delta$ such that

$$D(\rho\|\tau) \leq \delta \qquad \text{and} \qquad \lambda_{\min}(\tau) \geq \frac{\delta}{8d(2d+1)+\delta} \geq \frac{\delta}{25d^2}. \tag{5.114}$$

Given this covering upon states, we now want to take σ^n to be the separable state given by a mixture over such a covering, and the divergence centre

$$\sigma^n(\gamma) := \frac{1}{2}\sigma^*(I)^{\otimes n} + \frac{1}{2\left|\mathcal{C}^{\gamma/4}\right|}\sum_{\tau\in\mathcal{C}^{\gamma/4}} \tau^{\otimes n}. \tag{5.115}$$

Using the inequality

$$D_{\mathrm{h}}^\epsilon\big(\rho \,\big\|\, \mu\sigma + (1-\mu)\sigma'\big) \leq D_{\mathrm{h}}^\epsilon(\rho\|\sigma) - \log\mu \tag{5.116}$$

we will be able to bound divergences with respect to σ^n by those divergences with respect to either elements of $\mathcal{C}^{\gamma/4}$, or σ^*.

We will start by considering the low case. We will see that this case only accounts for hypothesis testing relative entropies which are below the capacity by a constant amount.

Lemma 5.14 (Low case)**.** *For any $\gamma > 0$, there exists a constant $N(\{a_i\}, \gamma)$ such that*

$$\frac{1}{n} D_{\mathrm{h}}^{\epsilon_n} \left(\bigotimes_{i=1}^{n} \rho_i \middle\| \sigma_n(\gamma) \right) \leq C - \gamma/4, \tag{5.117}$$

$$\frac{1}{n} D_{\mathrm{h}}^{1-\epsilon_n} \left(\bigotimes_{i=1}^{n} \rho_i \middle\| \sigma_n(\gamma) \right) \leq C - \gamma/4, \tag{5.118}$$

for any $\{\rho_i\}_i \subset I$, $n \in L(\{\rho_i\}, \gamma)$ and $n \geq N$.

Proof. We are going to start by considering the ϵ_n-hypothesis testing divergence. Take τ_n to be the closest element in $\mathcal{C}^{\gamma/4}$ to $\bar{\rho}_n$, such that $D(\bar{\rho}_n \| \tau_n) \leq \gamma/4$. Splitting out the τ_n term from $\sigma_n(\gamma)$, we have

$$D_{\mathrm{h}}^{\epsilon_n} \left(\bigotimes_{i=1}^{n} \rho_i \middle\| \sigma_n(\gamma) \right) \leq D_{\mathrm{h}}^{\epsilon_n} \left(\bigotimes_{i=1}^{n} \rho_i \middle\| \tau_n^{\otimes n} \right) + \log 2 \left| \mathcal{C}^{\gamma} \right| \tag{5.119}$$

$$\leq D_{\mathrm{h}}^{\epsilon_n} \left(\bigotimes_{i=1}^{n} \rho_i \middle\| \tau_n^{\otimes n} \right) + 2d^2 \log \left(\frac{120d}{\gamma^2} \right). \tag{5.120}$$

As the final term depending on $\left| \mathcal{C}^{\gamma/4} \right|$ is independent of n, there must exist a constant $N_1(\gamma)$ such that $2d^2 \log(120d/\gamma^2) \leq n\gamma/4$ for any $n \geq N_1$, and thus that

$$\frac{1}{n} D_{\mathrm{h}}^{\epsilon_n} \left(\bigotimes_{i=1}^{n} \rho_i \middle\| \sigma_n(\gamma) \right) \leq \frac{1}{n} D_{\mathrm{h}}^{\epsilon_n} \left(\bigotimes_{i=1}^{n} \rho_i \middle\| \tau_n^{\otimes n} \right) + \gamma/4. \tag{5.121}$$

Applying Proposition 5.8 to the ϵ_n-hypothesis testing relative entropy with respect to τ_n we get that there exists an $N_2(\{a_i\}, \gamma)$ such that

$$\frac{1}{n} D_{\mathrm{h}}^{\epsilon_n} \left(\bigotimes_{i=1}^{n} \rho_i \middle\| \tau_n^{\otimes} \right) \leq \frac{1}{n} \sum_{i=1}^{n} D(\rho_i \| \tau_n) + \gamma/4, \tag{5.122}$$

for any $n \geq N_2$. As for the divergence terms given with respect to τ_n, we can rearrange them in terms of divergences relative to the sequence mean $\bar{\rho}_n$ using the information geometric Pythagorean theorem, yielding

$$\sum_{i=1}^{n} D(\rho_i \| \tau_n) = \sum_{i=1}^{n} \operatorname{Tr} \rho_i (\log \rho_i - \log \bar{\rho}_n) + \sum_{i=1}^{n} \operatorname{Tr} \rho_i (\log \bar{\rho}_n - \log \tau_n) \tag{5.123}$$

$$= \sum_{i=1}^{n} D(\rho_i \| \bar{\rho}_n) + n D(\bar{\rho}_n \| \tau_n) \tag{5.124}$$

$$\leq \sum_{i=1}^{n} D(\rho_i \| \bar{\rho}_n) + n\gamma/4. \tag{5.125}$$

If we let $N(\{a_i\},\gamma) := \max\{N_1, N_2\}$, then pulling the above results together we see that for any $n \geq N$

$$\frac{1}{n} D_{\mathrm{h}}^{\epsilon_n}\left(\bigotimes_{i=1}^{n} \rho_i \middle\| \sigma_n(\gamma)\right) \leq \frac{1}{n} D_{\mathrm{h}}^{\epsilon_n}\left(\bigotimes_{i=1}^{n} \rho_i \middle\| \tau_n^{\otimes n}\right) + \gamma/4 \tag{5.126}$$

$$\leq \frac{1}{n} \sum_{i=1}^{n} D(\rho_i \| \tau_n) + 2\gamma/4 \tag{5.127}$$

$$\leq \frac{1}{n} \sum_{i=1}^{n} D(\rho_i \| \bar{\rho}_n) + 3\gamma/4. \tag{5.128}$$

Finally, since $n \in L(\{\rho_i\},\gamma)$ the average relative entropy is bounded away from capacity, and we arrive at the bound:

$$\frac{1}{n} D_{\mathrm{h}}^{\epsilon_n}\left(\bigotimes_{i=1}^{n} \rho_i \middle\| \sigma_n(\gamma)\right) \leq C - \gamma/4. \tag{5.129}$$

As we only relied on Proposition 5.8 to bound the regularised hypothesis testing divergence to within a constant of the average relative entropy, we could perform a similar analysis for the $(1-\epsilon_n)$-hypothesis testing divergence using Proposition 5.7 instead, which gives

$$\frac{1}{n} D_{\mathrm{h}}^{1-\epsilon_n}\left(\bigotimes_{i=1}^{n} \rho_i \middle\| \sigma_n(\gamma)\right) \leq C - \gamma/4. \tag{5.130}$$

□

Now that we have dealt with cases far from capacity, we turn our attention to the high cases.

Lemma 5.15 (High case)**.** *For any $\eta > 0$, there exist constants $\Gamma(\eta)$ and $N(\{a_i\},\eta)$, such that*

$$\frac{1}{n} D_{\mathrm{h}}^{\epsilon_n}\left(\bigotimes_{i=1}^{n} \rho_i \middle\| \sigma_n(\Gamma)\right) \leq C - \sqrt{2V_{\min}}\, a_n + \eta a_n \tag{5.131}$$

for any $\{\rho_i\}_i \subset I$, $n \in H(\{\rho_i\},\Gamma)$ and $n \geq N$. Similarly, the $(1-\epsilon_n)$-hypothesis testing relative entropy is bounded

$$\frac{1}{n} D_{\mathrm{h}}^{1-\epsilon_n}\left(\bigotimes_{i=1}^{n} \rho_i \middle\| \sigma_n(\Gamma)\right) \leq C + \sqrt{2V_{\max}} a_n + \eta a_n. \tag{5.132}$$

Proof. Splitting out the σ^* factor within $\sigma_n(\gamma)$ gives

$$D_{\mathrm{h}}^{\epsilon_n}\left(\bigotimes_{i=1}^{n} \rho_i \middle\| \sigma_n(\gamma)\right) \leq D_{\mathrm{h}}^{\epsilon_n}\left(\bigotimes_{i=1}^{n} \rho_i \middle\| \sigma^{*\otimes n}\right) + \log 2, \tag{5.133}$$

$$D_{\mathrm{h}}^{1-\epsilon_n}\left(\bigotimes_{i=1}^{n} \rho_i \middle\| \sigma_n(\gamma)\right) \leq D_{\mathrm{h}}^{1-\epsilon_n}\left(\bigotimes_{i=1}^{n} \rho_i \middle\| \sigma^{*\otimes n}\right) + \log 2. \tag{5.134}$$

As $\frac{1}{n}\log 2 = o(a_n)$, there exists an $N_1(\{a_i\})$ such that $n \geq N_1$ implies

$$D_{\mathrm{h}}^{\epsilon_n}\left(\bigotimes_{i=1}^{n}\rho_i \middle\| \sigma_n(\gamma)\right) \leq D_{\mathrm{h}}^{\epsilon_n}\left(\bigotimes_{i=1}^{n}\rho_i \middle\| \sigma^{*\otimes n}\right) + \eta a_n/3, \tag{5.135}$$

$$D_{\mathrm{h}}^{1-\epsilon_n}\left(\bigotimes_{i=1}^{n}\rho_i \middle\| \sigma_n(\gamma)\right) \leq D_{\mathrm{h}}^{1-\epsilon_n}\left(\bigotimes_{i=1}^{n}\rho_i \middle\| \sigma^{*\otimes n}\right) + \eta a_n/3. \tag{5.136}$$

We now wish to employ a moderate deviation result. We will start by addressing the ϵ_n-hypothesis testing divergence. For the weaker bound of Proposition 5.8 we will have no required bounds on $\frac{1}{n}\sum_{i=1}^{n} V(\rho_i\|\sigma^*)$, but for the stronger bound we will need a uniform lower bound.

If $V_{\min} \leq \eta^2/18$, then the weakened bound of Proposition 5.8 is sufficient, giving an $N_2(\{a_n\},\eta)$ such that $n \geq N_2$ implies

$$D_{\mathrm{h}}^{\epsilon_n}\left(\bigotimes_{i=1}^{n}\rho_i \middle\| \sigma_n(\gamma)\right) \leq \frac{1}{n} D_{\mathrm{h}}^{\epsilon_n}\left(\bigotimes_{i=1}^{n}\rho_i \middle\| \sigma^{*\otimes n}\right) + \eta a_n/3 \tag{5.137}$$

$$\leq \frac{1}{n}\sum_{i=1}^{n} D(\rho_i\|\sigma^*) + 2\eta a_n/3 \tag{5.138}$$

$$\leq \frac{1}{n}\sum_{i=1}^{n} D(\rho_i\|\sigma^*) - \sqrt{2V_{\min}}\, a_n + \eta a_n \tag{5.139}$$

$$\leq C - \sqrt{2V_{\min}}\, a_n + \eta a_n. \tag{5.140}$$

Next we need to consider the case where $V_{\min} > \eta^2/18$. To do this, we will need to establish a lower bound on $\frac{1}{n}\sum_{i=1}^{n} V(\rho_i\|\sigma^*)$, which places it near $V_{\min}$. The min-dispersion is defined for distributions which exactly achieve capacity; we will now consider an analogous quantity for distributions which are *near* capacity. Specifically

$$V_{\min}(\gamma) := \inf_{P\in\mathcal{P}(\mathcal{I})}\left\{\int \mathrm{d}P(\rho)\; V\left(\rho\|\sigma^*\right) \;\middle|\; \int \mathrm{d}P(\rho)\; D\left(\rho \middle\| \int \mathrm{d}P(\rho')\;\rho'\right) \geq C - \gamma\right\}. \tag{5.141}$$

By definition of the channel dispersion we have that $V_{\min}(0) = V_{\min}$. By Lemma 22 of Ref. [5.7] we can strengthen this to $\lim_{\gamma\to 0^+} V_{\min}(\gamma) = V_{\min}$, and so for any $\eta > 0$ there must exist a constant $\Gamma(\eta)$ such that

$$\sqrt{2V_{\min}(\Gamma)} \geq \sqrt{2V_{\min}} - \eta/3. \tag{5.142}$$

As $V_{\min} \geq \eta^2/18$, this implies that $V_{\min}(\Gamma) > 0$.

Next, let P_n be the empirical distribution corresponding to the set $\{\rho_i\}_{i=1}^n$, i.e. $P_n(\rho) := \frac{1}{n}\sum_{i=1}^{n}\delta(\rho-\rho_i)$. For all $n \in H(\{\rho_i\},\Gamma)$, these distributions are near capacity

$$\int \mathrm{d}P_n(\rho)\; D\left(\rho \middle\| \int \mathrm{d}P_n(\rho')\;\rho'\right) = \frac{1}{n}\sum_{i=1}^{n} D(\rho_i\|\bar{\rho}_n) \geq C - \Gamma, \tag{5.143}$$

and so we can lower bound the average variance with respect to the divergence centre

$$\frac{1}{n}\sum_{i=1}^{n} V(\rho_i\|\sigma^*) = \int \mathrm{d}P(\rho)\; V(\rho\|\sigma^*) \geq V_{\min}(\Gamma) > 0. \tag{5.144}$$

Using this lower bound, we can apply the stronger bound from Proposition 5.8 to give a constant $N_3(\{a_i\},\eta)$, such that, for every $n \in H(\{\rho_i\},\Gamma)$ and $n \geq N_3$, the hypothesis testing divergence is upper bounded

$$D_{\mathrm{h}}^{\epsilon_n}\left(\bigotimes_{i=1}^{n}\rho_i \,\middle\|\, \sigma_n(\gamma)\right) \leq \frac{1}{n} D_{\mathrm{h}}^{\epsilon_n}\left(\bigotimes_{i=1}^{n}\rho_i \,\middle\|\, \sigma^{*\otimes n}\right) + \eta a_n/3 \tag{5.145}$$

$$\leq \frac{1}{n}\sum_{i=1}^{n} D(\rho_i\|\sigma^*) - \sqrt{\frac{2}{n}\sum_{i=1}^{n} V(\rho_i\|\sigma^*)}\, a_n + 2\eta a_n/3 \tag{5.146}$$

$$\leq \frac{1}{n}\sum_{i=1}^{n} D(\rho_i\|\sigma^*) - \sqrt{2V_{\min}(\Gamma)}\, a_n + 2\eta a_n/3 \tag{5.147}$$

$$\leq C - \sqrt{2V_{\min}}\, a_n + \eta a_n. \tag{5.148}$$

Performing a similar argument for $V_{\max}$, we construct a function

$$V_{\max}(\gamma) := \sup_{P\in\mathcal{P}(\mathcal{I})} \left\{ \int \mathrm{d}P(\rho)\; V\left(\rho\|\sigma^*\right) \;\middle|\; \int \mathrm{d}P(\rho)\; D\left(\rho \,\middle\|\, \int \mathrm{d}P(\rho')\; \rho'\right) \geq C - \gamma \right\}, \tag{5.149}$$

and define a Γ such that

$$\sqrt{2V_{\max}(\Gamma)} \leq \sqrt{2V_{\max}} + \eta/3. \tag{5.150}$$

Following through the rest of the argument, and employing Proposition 5.7, we also get a bound on the $(1-\epsilon_n)$-hypothesis testing divergence

$$\frac{1}{n} D_{\mathrm{h}}^{1-\epsilon_n}\left(\bigotimes_{i=1}^{n}\rho_i \,\middle\|\, \sigma_n(\gamma)\right) \leq C + \sqrt{2V_{\max}} a_n + \eta a_n. \tag{5.151}$$

□

Proposition 5.16 (Channel coding: Optimality)**.** *For any moderate sequence $\{a_n\}_n$ and error probability $\epsilon_n := e^{-na_n^2}$, the rate is upper bounded as*

$$R^*\left(n,\epsilon_n\right) \leq C - \sqrt{2V_{\min}}\, a_n + o(a_n). \tag{5.152}$$

For error probability $(1-\epsilon_n)$ the rate is similarly upper bound as

$$R^*\left(n,1-\epsilon_n\right) \leq C + \sqrt{2V_{\max}}\, a_n + o(a_n). \tag{5.153}$$

Proof. Applying Lemmas 5.14 and 5.15, we get that there exist constants $\Gamma(\eta)$ and $N_1(\{a_i\},\eta)$ such that

$$\frac{1}{n}D_{\mathrm{h}}^{\epsilon_n}\left(\bigotimes_{i=1}^{n}\rho_i\middle\|\sigma_n(\Gamma)\right)\leq\begin{cases}C-\Gamma/4 & n\in L(\{\rho_i\},\Gamma)\\ C-\sqrt{2V_{\min}}\,a_n+\eta a_n & n\in H(\{\rho_i\},\Gamma)\end{cases} \tag{5.154}$$

for any $n\geq N_1$. As Γ is a constant, there must exist some $N_2(\{a_i\},\eta)$ such that $\Gamma/4\geq\sqrt{2V_{\min}}a_n$. As such, for any $n\geq\max\{N_1,N_2\}$, high or low, we have

$$\frac{1}{n}D_{\mathrm{h}}^{\epsilon_n}\left(\bigotimes_{i=1}^{n}\rho_i\middle\|\sigma_n(\Gamma)\right)\leq C-\sqrt{2V_{\min}}\,a_n+\eta a_n. \tag{5.155}$$

Pulling this bound back to Eq. 5.110, we have

$$R^*(n,\epsilon_n)\leq\sup_{\{\rho_i\}\subseteq I}\frac{1}{n}D_{\mathrm{h}}^{2\epsilon_n}\left(\bigotimes_{i=1}^{n}\rho_i\middle\|\sigma^n(\Gamma)\right)+\frac{1}{n}\log\frac{2}{1-\epsilon_n} \tag{5.156}$$

$$\leq C-\sqrt{2V_{\min}}a_n+\eta a_n+\frac{1}{n}\log\frac{2}{1-\epsilon_n}. \tag{5.157}$$

Finally, noting that $1/n=o(a_n)$, there must exist a constant $N_3(\{a_i\},\eta)$ such that $n\geq N_3$ implies

$$\frac{1}{n}\log\frac{2}{1-\epsilon_n}\leq\eta a_n. \tag{5.158}$$

We can therefore conclude that, for $n\geq\max\{N_1,N_2,N_3\}$, we get the overall upper bound

$$R^*(n,\epsilon_n)\leq C-\sqrt{2V_{\min}}\,a_n+2\eta a_n. \tag{5.159}$$

As this is true for arbitrary $\eta>0$, we can take $\eta\searrow 0$ and conclude

$$R^*(n,\epsilon_n)\leq C-\sqrt{2V_{\min}}a_n+o(a_n) \tag{5.160}$$

as required. A similar analysis for the $(1-\epsilon_n)$-error regime shows

$$R^*\left(n,1-\epsilon_n\right)\leq C+\sqrt{2V_{\max}}a_n+o(a_n). \tag{5.161}$$

□

5.5 Conclusion

The main result of this paper is to give a second order approximation of the non-asymptotic fundamental limit for classical information transmission over a quantum channel in the moderate deviations regime, as in Eqs. 5.3 and 5.4:

$$\frac{1}{n}\log M^*(\mathcal{W};n,\epsilon_n)=C(\mathcal{W})-\sqrt{2V_{\min}(\mathcal{W})}\,x_n+o(x_n)\,, \tag{5.162}$$

$$\frac{1}{n}\log M^*(\mathcal{W};n,1-\epsilon_n)=C(\mathcal{W})+\sqrt{2V_{\max}(\mathcal{W})}\,x_n+o(x_n)\,. \tag{5.163}$$

Along the lines of third and fourth order approximations for classical channel coding in the fixed error regime (see, e.g., Refs. [5.44–5.46]), a natural question to ask is whether we can expand this further and resolve the term $o(x_n)$. A preliminary investigation suggests the conjecture that $o(x_n) = O(x_n^2) + O(\log n)$ and that at least some of the implicit constants can be determined precisely. We leave this for future work.

Due to the central importance of binary asymmetric quantum hypothesis testing we expect our techniques to have applications also to other quantum channel coding tasks. In particular, source coding [5.47, 5.48], entanglement-assisted classical coding [5.49] as well as quantum [5.50] and private coding [5.51] over quantum channels have recently been analysed in the small deviations regime by relating the problem to quantum hypothesis testing. An extension of these results to moderate deviations using our techniques thus appears feasible.

Acknowledgements. MT is funded by an Australian Research Council Discovery Early Career Researcher Award (DECRA) fellowship (Grant Nos. CE110001013, DE160100821). Both MT and CTC acknowledge support from the ARC Centre of Excellence for Engineered Quantum Systems (EQuS). We also thank Hao-Chung Cheng and Min-Hsiu Hsieh for useful discussions and insightful comments.

5.A Moderate deviation tail bounds

5.A.1 Lower bound

Here we apply the lower bound of Ref. [5.52], which gives a Berry–Esseen-type inequality with multiplicative error.

Lemma 5.17 (Theorem B2, Ref. [5.52])**.** *There exists universal constants κ_1, κ_2 such that, for independent zero-mean variables $X_1, \ldots, X_n$ with*

$$V_n := \frac{1}{n}\sum_{i=1}^{n} \mathrm{Var}\left[X_i\right] \qquad \text{and} \qquad T_n := \frac{1}{n}\sum_{i=1}^{n} \mathbb{E}\left[|X_i|^3\right], \tag{5.164}$$

and a t_n bounded

$$\sqrt{\frac{V_n}{n}} \leq t_n \leq \frac{V_n^2}{T_n}, \tag{5.165}$$

the probability that the average variable $\frac{1}{n}\sum_{i=1}^{n} X_i$ deviates above the mean by t_n is lower bounded

$$\ln \Pr\left[\frac{1}{n}\sum_{i=1}^{n} X_i \geq t_n\right] \geq \ln \Phi\left(-\sqrt{\frac{nt_n^2}{V_n}}\right) - \frac{\kappa_1 T_n n t_n^3}{V_n^3} + \ln\left(1 - \frac{\kappa_2 T_n t_n}{V_n^2}\right). \tag{5.166}$$

Given this Lemma, we can now prove the desired lower bound on the moderately deviating tail.

Proof of Lemma 5.1. First we note that the bound on the average third absolute moment also imposes a bound on the average variance

$$V_n = \frac{1}{n}\sum_{i=1}^{n}\mathbb{E}\left[X^2\right] \tag{5.167}$$

$$< \frac{1}{n}\sum_{i=1}^{n}\mathbb{E}\left[|X|^3+1\right] \tag{5.168}$$

$$= T_n + 1 \tag{5.169}$$

$$\leq \tau + 1. \tag{5.170}$$

As $\{t_i\}_i$ is moderate, and the moments are bounded $\nu \leq V_n \leq \tau+1$ and $T_n \leq \tau$, there must exist an $N_1(\{t_i\},\nu,\tau)$ such that

$$\sqrt{\frac{\tau+1}{n}} \leq t_n \leq \frac{\nu^2}{\tau} \tag{5.171}$$

for $n \geq N_1$. Applying Lemma 5.17, we have that for $n \geq N_1$

$$\ln \Pr\left(\frac{1}{n}\sum_{i=1}^{n}X_{i,n} \geq t\right) \geq \ln\Phi\left(-\sqrt{\frac{nt_n^2}{V_n}}\right) - \frac{\kappa_1\tau}{\nu^3}nt_n^3 + \ln\left(1-\frac{\kappa_2\tau}{\nu^{3/2}}t_n\right). \tag{5.172}$$

As $nt_n^2 \to \infty$ and $V_n \leq \tau+1$, there must exist a constant $N_2(\{t_i\},\tau)$ such that $n \geq N_2$ implies $nt_n^2/V_n \geq 1$. Using the standard bound $\ln\Phi(-x) \geq -x^2/2 - \ln\sqrt{8\pi}x$ for $x \geq 1$, we find

$$\ln \Pr\left(\frac{1}{n}\sum_{i=1}^{n}X_{i,n} \geq t\right) \geq -\frac{nt_n^2}{2V_n} - \ln\sqrt{8\pi\frac{nt_n^2}{V_n}} - \frac{\kappa_1\tau}{\nu^3}nt_n^3 + \ln\left(1-\frac{\kappa_2\tau}{\nu^{3/2}}t_n\right). \tag{5.173}$$

As t_n is moderate, we have that the first term $-nt_n^2/2V_n$ dominates as $n\to\infty$ in the above. As such, for any $\eta > 0$, there must exist an $N_3(\{t_n\},\nu,\tau,\eta)$ such that, for all $n \geq N_3$, the other terms are smaller than this dominant term by a multiplicative factor of $\eta > 0$, such that

$$\ln\sqrt{8\pi\frac{nt_n^2}{V_n}} + \frac{\kappa_1\tau}{V_n^3}nt_n^3 - \ln\left(1-\frac{\kappa_2\tau}{V_n^{3/2}}t_n\right) \leq \eta\frac{nt_n^2}{2V_n}. \tag{5.174}$$

We conclude that for $n \geq N(\{t_i\},\nu,\tau,\eta) := \max\{N_1,N_2,N_3\}$

$$\ln\Pr\left[\frac{1}{n}\sum_{i=1}^{n}X_{i,n} \geq t_n\right] \geq -(1+\eta)\frac{nt_n^2}{2V_n}. \tag{5.175}$$

□

5.A.2 Upper bound

For the upper bound we are going to use a proof technique similar to that used to prove Cramér's and Gartner-Ellis theorems in the large deviation regime (see, e.g., Ref. [5.8]), and for Lemma 4 of Ref. [5.53] in the iid moderate deviation regime. However, our approach differs from that in Ref. [5.53] because we do not want to assume that the average variance, V_n, is bounded away from zero.

Proof of Lemma 5.2. Let $h_n(s)$ be the average cumulant generating function

$$h_n(s) := \frac{1}{n}\sum_{i=1}^{n} \ln \mathbb{E}\left[e^{sX_{i,n}}\right], \tag{5.176}$$

such that

$$h_n(0) = 0, \qquad h_n'(0) = \frac{1}{n}\sum_{i=1}^{n} \mathbb{E}\left[X_{i,n}\right] = 0, \qquad h_n''(0) = \frac{1}{n}\sum_{i=1}^{n} \mathrm{Var}\left[X_{i,n}\right] = V_n. \tag{5.177}$$

For our tail bound we are going to employ a Chernoff bound. Specifically for any $\alpha > 0$, an application of the Markov inequality gives

$$\Pr\left[\frac{1}{n}\sum_{i=1}^{n} X_{i,n} \geq t_n\right] = \Pr\left[e^{\alpha t_n \sum_{i=1}^{n} X_{i,n}} \geq e^{\alpha n t_n^2}\right] \leq \frac{\mathbb{E}\left[e^{\alpha t_n \sum_{i=1}^{n} X_{i,n}}\right]}{e^{\alpha n t_n^2}}. \tag{5.178}$$

Using the independence of $\{X_i\}$, the above bound can be expressed in terms of the average cumulant generating function as

$$\ln \Pr\left[\frac{1}{n}\sum_{i=1}^{n} X_{i,n} \geq t_n\right] \leq -n\big(\alpha t_n^2 - h_n(\alpha t_n)\big). \tag{5.179}$$

In general our choice of α will depend on n. If we assume for the moment that α is bounded then, as $t_n \to 0$, there exists a constant $N_1(\{t_i\}, \alpha)$ such that $n \geq N_1$ implies $\alpha t_n \leq 1/2$. Applying Taylor's theorem for such n, specifically a second-order expansion with the error in Lagrange form, gives that there exists an $s \in [0, \alpha t_n] \subseteq [0, 1/2]$ such that

$$h_n(s) = h_n(0) + \alpha t_n h_n'(0) + \alpha^2 t_n^2 h_n''(0)/2 + \alpha^3 t_n^3 h_n'''(s)/6 \tag{5.180}$$

$$\leq \alpha^2 t_n^2 V_n/2 + \alpha^3 t_n^3 \gamma/6. \tag{5.181}$$

Plugging this Taylor expansion in to our Chernoff bound above gives

$$\frac{1}{n t_n^2} \ln \Pr\left[\frac{1}{n}\sum_{i=1}^{n} X_{i,n} \geq t_n\right] \leq \left(\alpha^2 V_n/2 - \alpha\right) + \alpha^2 \gamma t_n/6. \tag{5.182}$$

We now need to choose our value of α. An obvious choice would be $\alpha = 1/V_n$, which gives the tightest possible asymptotic bound. As we have not imposed a lower bound on V_n, this value is not necessarily bounded, and therefore could render the previous Taylor expansion invalid. Instead we will slightly

modify this choice such that the Taylor expansion is still valid, whilst only changing the final bound by the introducing of an η. Specifically we will take

$$\alpha^{-1} := \sqrt{V_n + \eta/4}\left(\sqrt{V_n + \eta/4} + \sqrt{\eta/4}\right). \tag{5.183}$$

As required, this choice of α is bounded independent V_n as $\alpha \leq 2/\eta$, meaning that the previous Taylor expansion was indeed valid, and that $N_1 = N_1(\{t_i\}, \eta)$. Plugging this choice in to Eq. refeqn:chernoff gives

$$\ln \Pr\left[\frac{1}{n}\sum_{i=1}^{n} X_{i,n} \geq t_n\right] \leq -\frac{nt_n^2}{2V_n + \eta/2} + \frac{2\gamma}{3\eta^2} n t_n^3. \tag{5.184}$$

Similar to Lemma 5.1, the bound on the third derivative of cumulant function bounds the variances as $\mathrm{Var}[X_i] \leq \gamma + 1$. Given this, there must exist a constant $N_2(\{t_i\}, \gamma, \eta)$ such that $n \geq N_2$ implies

$$-\frac{1}{2V_n + \eta/2} + \frac{2\gamma}{3\eta^3} t_n \leq -\frac{1}{2V_n + \eta}. \tag{5.185}$$

We conclude therefore that for any $n \geq N(\{t_i\}, \gamma, \eta) := \max\{N_1, N_2\}$ we have the desired tail bound

$$\ln \Pr\left[\frac{1}{n}\sum_{i=1}^{n} X_i \leq t_n\right] \leq -\frac{nt_n^2}{2V_n + \eta}. \tag{5.186}$$

□

5.A.3 Dimensionless bound

The non-dimensional bound follows as a corollary of the two previous bound, where we explicitly use the possible dependence of our random variables $X_{i,n}$ on n.

Proof of Corollary 5.3. Starting with random variables $\{X_{i,n}\}_{i \leq n}$, define rescaled variables as $\tilde{X}_{i,n} := X_{i,n}/\sqrt{V_n}$ for all $i \leq n$. This scaling has the property that it normalises the average variance

$$\tilde{V}_n := \frac{1}{n}\sum_{i=1}^{n} \mathrm{Var}[\tilde{X}_{i,n}] = 1. \tag{5.187}$$

As well as this, we can see the dimensionless assumption on $X_{i,n}$

$$\frac{1}{nV_n^{3/2}} \sum_{i=1}^{n} \sup_{s \in [0,1/2]} \left| \frac{\mathrm{d}^3}{\mathrm{d}s^3} \ln \mathbb{E}\left[e^{sX_{i,n}}\right] \right| \leq \gamma, \tag{5.188}$$

is equivalent to the bound on $\tilde{X}_{i,n}$

$$\frac{1}{n} \sum_{i=1}^{n} \sup_{s \in [0,1/2]} \left| \frac{\mathrm{d}^3}{\mathrm{d}s^3} \ln \mathbb{E}\left[e^{s\tilde{X}_{i,n}}\right] \right| \leq \gamma. \tag{5.189}$$

Noticing that

$$\Pr\left[\frac{1}{n}\sum_{i=1}^{n} X_{i,n} \leq t_n\sqrt{V_n}\right] = \Pr\left[\frac{1}{n}\sum_{i=1}^{n} X_{i,n}/\sqrt{V_n} \leq t_n\right] = \Pr\left[\frac{1}{n}\sum_{i=1}^{n} \tilde{X}_{i,n} \leq t_n\right], \tag{5.190}$$

we can simply apply the existing tail bounds of Lemmas 5.1 and 5.2 to $\tilde{X}_{i,n}$, giving that, for any $\eta > 0$, there must exist a constant $N(\{t_i\}, \gamma, \eta)$ such that

$$-(1+\eta)\frac{nt_n^2}{2} \leq \ln \Pr\left[\frac{1}{n}\sum_{i=1}^{n} X_{i,n} \leq t_n\sqrt{V_n}\right] \leq -(1-\eta)\frac{nt_n^2}{2}. \tag{5.191}$$

□

5.B Proof of Lemma 5.4

Lemma 5.18. *Let A and B be two real functions both defined on the same domain , with*

$$\hat{A}(b) := \inf_t \{A(t)|B(t) \leq b\} \qquad \text{and} \qquad \hat{B}(b) := \inf_t \{B(t)|A(t) \leq a\}, \tag{5.192}$$

then for any $a \geq \inf_t A(t)$ and $\delta > 0$

$$\hat{A}\big(\hat{B}(a)+\delta\big) \leq a \tag{5.193}$$

$$\hat{A}\big(\hat{B}(a)-\delta\big) \geq a. \tag{5.194}$$

Proof. By the definition of the infimum in $\hat{B}(\cdot)$, there must exist a $t^\star$ such that $A(t^\star) \leq a$ and $B(t^\star) \leq \hat{B}(a) + \delta$. Hence we can upper bound

$$\hat{A}\big(\hat{B}(a)+\delta\big) = \inf_s \Big\{A(s)\Big|B(s) \leq \hat{B}(a)+\delta\Big\} \leq A(t^\star) \leq a. \tag{5.195}$$

Next, suppose $\hat{A}\big(\hat{B}(a)-\delta\big) \leq a - \epsilon$ for some $\epsilon > 0$. By definition of the infimum in $\hat{A}(\cdot)$, there must therefore exist an $s^\star$ such that $B(s^\star) \leq \hat{B}(a) - \delta$ and $A(s^\star) \leq a$. This in turn allows us to upper bound

$$\hat{B}(a) = \inf_t \{B(t)|A(t) \leq a\} \leq B(s^\star). \tag{5.196}$$

We can therefore conclude that $\hat{B}(x) \leq \hat{B}(a) - \delta$, proving $\hat{A}\big(\hat{B}(a)-\delta\big) > a - \epsilon$ by contradiction. As this is true for arbitrarily small ϵ, we therefore conclude

$$\hat{A}\big(\hat{B}(x)-\delta\big) \geq a. \tag{5.197}$$

□

Proof of Lemma 5.4. By swapping both A_n and B_n we can see that the forward and backwards directions of this proof are equivalent, as such we will only consider the forward direction. First, we assume that

$$\lim_{n\to\infty} \frac{\hat{B}_n(a_n)}{a_n} = 1, \quad \forall a_n \text{ moderate.} \tag{5.198}$$

Next we split the proof of the limit into upper bounding the limit superior, and lower bounding the limit inferior.

Take any moderate sequence b_n, and let $a_n := b_n/2$ and $b'_n := \hat{B}_n(a_n) + b_n/n$. By Lemma 5.18 we have that $\hat{A}_n(b'_n) = \hat{A}_n\big(\hat{B}_n(a_n) + b_n/n\big) \leq a_n$. By assumption we then have that

$$\lim_{n\to\infty} \frac{b'_n}{b_n} = \lim_{n\to\infty} \frac{\hat{B}_n(a_n) + 2a_n/n}{2a_n} = \lim_{n\to\infty} \frac{\hat{B}_n(a_n)}{2a_n} = \frac{1}{2}. \tag{5.199}$$

As a result we have, for sufficiently large n, that $b'_n \leq b_n$. Using this we can bound the limit superior:

$$\limsup_{n\to\infty} \frac{\hat{A}_n(b_n)}{b_n} \leq \lim_{n\to\infty} \frac{a_n}{b'_n} \tag{5.200}$$

$$= \lim_{n\to\infty} \frac{a_n}{\hat{B}_n(a_n)} = 1. \tag{5.201}$$

Moreover, if we take $a_n := 2b_n$ and $b'_n := \hat{B}_n(a_n) - b_n/n$ then, by an analogous argument, Lemma 5.18 gives us $\hat{A}_n(b'_n) \geq a_n$, and the assumption gives us $b'_n \geq b_n$ for sufficiently large n. As such we can also bound the limit inferior

$$\liminf_{n\to\infty} \frac{\hat{A}_n(b_n)}{b_n} \geq \lim_{n\to\infty} \frac{a_n}{b'_n} = \lim_{n\to\infty} \frac{a_n}{\hat{B}_n(a_n)} = 1. \tag{5.202}$$

□

5.C Proof of Lemma 5.6

Proof. Consider the moment generating function $m(t) := \mathbb{E}\left[e^{tZ}\right]$, such that $h(t) = \ln m(t)$. Similar to the relationship between cumulants and central moments, the derivatives of $h(t)$ can be expressed in terms of derivatives of $m(t)$:

$$h = \ln m \tag{5.203}$$

$$h' = \frac{m'}{m} \tag{5.204}$$

$$h'' = \frac{m''m - m'm'}{m^2} \tag{5.205}$$

$$h''' = \frac{m'''mm - 3m''m'm + 2m'm'm'}{m^3} \tag{5.206}$$

$$\vdots \tag{5.207}$$

As such, if we bound $m(t)$ away from zero, proving that the derivatives of $m(t)$ are uniformly bounded would imply the same about $h(t)$. Noticing that $\sum_a r_a = 1$ implies $\sum_a r_a^2 \geq 1/d$ and $\lambda \leq 1/d$, we can see that m is bounded away from zero for any $t \leq 1$:

$$m(t) = \sum_{a,b} r_a \left|\langle \phi_a | \psi_b \rangle\right|^2 (r_a/s_b)^t \tag{5.208}$$

$$\geq \frac{1}{\lambda^t} \sum_{a,b} \left|\langle \phi_a | \psi_b \rangle\right|^2 r_a^{t+1} \tag{5.209}$$

$$\geq \frac{1}{\lambda^t} \sum_{a} r_a^{1+t} \tag{5.210}$$

$$\geq \frac{1}{\lambda^t} \sum_{a} r_a^{2} \tag{5.211}$$

$$\geq \frac{1}{d\lambda} \geq 1 \tag{5.212}$$

Next we can use the bound $\sup_{x\in[0,1]} \sqrt{x} \ln^k(1/x) = (2k/e)^k$, to bound the derivatives of the moment generating function for $|t| \leq 1/2$

$$\left|m^{(k)}(t)\right| = \left|\sum_{a,b} r_a \left|\langle \phi_a | \psi_b \rangle\right|^2 (r_a/s_b)^t \ln^k (r_a/s_b)\right| \tag{5.213}$$

$$\leq \sum_{a,b} r_a \left|\langle \phi_a | \psi_b \rangle\right|^2 (r_a/s_b)^t \left|\ln^k (r_a/s_b)\right| \tag{5.214}$$

$$\leq \sum_{a,b} r_a \left|\langle \phi_a | \psi_b \rangle\right|^2 (r_a/s_b)^t \left[\ln^k(1/s_b) + \ln^k(1/r_a)\right] \tag{5.215}$$

$$\leq \sum_{a,b} \left|\langle \phi_a | \psi_b \rangle\right|^2 \sqrt{r_a/s_b} \left[\ln^k(1/s_b) + \ln^k(1/r_a)\right] \tag{5.216}$$

$$\leq \max_{\substack{\lambda\leq s\leq 1\\ 0\leq r\leq 1}} \sqrt{r/s} \left[\ln^k(1/s) + \ln^k(1/r)\right] \tag{5.217}$$

$$\leq \frac{1}{\sqrt{\lambda}} \max_{\substack{\lambda\leq s\leq 1\\ 0\leq r\leq 1}} \left[\ln^k(1/s) + \sqrt{r} \ln^k(1/r)\right] \tag{5.218}$$

$$\leq \frac{1}{\sqrt{\lambda}} \left[\ln^k\left(\frac{1}{\lambda}\right) + (2k/e)^k\right] =: C_k. \tag{5.219}$$

□

Bibliography

5.1 A. S. Holevo, "The Capacity of the Quantum Channel with General Signal States," *IEEE Transactions on Information Theory* **44**, 269–273, arXiv:quant-ph/9611023, (1998).

5.2 A. S. Holevo, "Bounds for the Quantity of Information Transmitted by a Quantum Communication Channel," *Problems of Information Transmission* **9**, 3, 177–183, (1973).

5.3 B. Schumacher and M. Westmoreland, "Sending Classical Information via Noisy Quantum Channels," *Physical Review A* **56**, 131–138, (1997).

5.4 A. Holevo, "Reliability function of general classical-quantum channel," *IEEE Transactions on Information Theory* **46**, 2256–2261, arXiv:quant-ph/9907087, (2000).

5.5 M. Hayashi, "Error Exponent in Asymmetric Quantum Hypothesis Testing and Its Application to Classical-Quantum Channel coding," *Physical Review A* **76**, 062301, arXiv:quant-ph/0611013, (2006).

5.6 M. Dalai, "Lower Bounds on the Probability of Error for Classical and Classical-Quantum Channels," *IEEE Transactions on Information Theory* **59**, 8027–8056, arXiv:1201.5411, (2013).

5.7 M. Tomamichel and V. Y. F. Tan, "Second-Order Asymptotics for the Classical Capacity of Image-Additive Quantum Channels," *Communications in Mathematical Physics* **338**, 103–137, arXiv:1308.6503, (2015).

5.8 A. Dembo and O. Zeitouni, *Large Deviations Techniques and Applications*, Stochastic Modelling and Applied Probability, Springer, (1998).

5.9 H. Umegaki, "Conditional expectation in an operator algebra. IV. Entropy and information," *Kodai Mathematical Seminar Reports* **14**, 2, 59–85, (1962).

5.10 M. Tomamichel and M. Hayashi, "A Hierarchy of Information Quantities for Finite Block Length Analysis of Quantum Tasks," *IEEE Transactions on Information Theory* **59**, 7693–7710, arXiv:1208.1478, (2013).

5.11 K. Li, "Second-Order Asymptotics for Quantum Hypothesis Testing," *Annals of Statistics* **42**, 171–189, arXiv:1208.1400, (2014).

5.12 Y. Altug and A. B. Wagner, "Moderate Deviations in Channel Coding," *IEEE Transactions on Information Theory* **60**, 4417–4426, arXiv:1208.1924, (2014).

5.13 Y. Polyanskiy and S. Verdu, "Channel dispersion and moderate deviations limits for memoryless channels," *2010 48th Annual Allerton Conference on Communication, Control, and Computing (Allerton)*, 1334–1339, IEEE, (2010).

5.14 W. Mathews, "A linear program for the finite block length converse of Polyanskiy-Poor-Verdú via non-signalling codes," *IEEE Transactions on Information Theory* **58**, 7036–7044, arXiv:1109.5417, (2012).

5.15 W. Mathews and S. Wehner, "Finite blocklength converse bounds for quantum channels," *IEEE Transactions on Information Theory* **60**, 7317–7329, arXiv:1210.4722, (2014).

5.16 L. Wang and R. Renner, "One-Shot Classical-Quantum Capacity and Hypothesis Testing," *Physical Review Letters* **108**, 200501, arXiv:1007.5456, (2012).

5.17 H.-C. Cheng and M.-H. Hsieh, "Moderate Deviation Analysis for Classical-Quantum Channels and Quantum Hypothesis Testing," arXiv:1701.03195, (2016).

5.18 W. Hoeffding, "Asymptotically Optimal Tests for Multinomial Distributions," *Annals of Mathematical Statistics* **36**, 2, 369–401, (1965).

5.19 R. G. Gallager, Information Theory and Reliable Communication, Wiley, (1968).

5.20 I. Csiszár and J. Körner, Information Theory: Coding Theorems for Discrete Memoryless Systems, Cambridge University Press, (2011).

5.21 H. Nagaoka, "The Converse Part of The Theorem for Quantum Hoeffding Bound," arXiv:quant-ph/0611289, (2006).

5.22 I. Sason, "Moderate deviations analysis of binary hypothesis testing," *2012 IEEE International Symposium on Information Theory Proceedings*, 821–825, IEEE, arXiv:1111.1995, (2012).

5.23 V. Strassen, "Asymptotische Abschätzungen in Shannons Informationstheorie," *Trans. Third Prague Conference on Information Theory*, Prague, 689–723, (1962).

5.24 M. Hayashi, "Information Spectrum Approach to Second-Order Coding Rate in Channel Coding," *IEEE Transactions on Information Theory* **55**, 4947–4966, arXiv:0801.2242, (2009).

5.25 Y. Polyanskiy, H. V. Poor, and S. Verdú, "Channel Coding Rate in the Finite Blocklength Regime," *IEEE Transactions on Information Theory* **56**, 2307–2359, (2010).

5.26 I. Csiszár and G. Longo, "On the Error Exponent for Source Coding and for Testing Simple Statistical Hypotheses," *Studia Scientiarum Mathematicarum Hungarica* **6**, 181–191, (1971).

5.27 T. S. Han and K. Kobayashi, "The Strong Converse Theorem for Hypothesis Testing," *IEEE Transactions on Information Theory* **35**, 178–180, (1989).

5.28 S. Arimoto, "On the converse to the coding theorem for discrete memoryless channels," *IEEE Transactions on Information Theory* **19**, 357–359, (1973).

5.29 G. Dueck and J. Korner, "Reliability function of a discrete memoryless channel at rates above capacity (Corresp.)," *IEEE Transactions on Information Theory* **25**, 82–85, (1979).

5.30 M. Mosonyi and T. Ogawa, "Quantum Hypothesis Testing and the Operational Interpretation of the Quantum Rényi Relative Entropies," *Communications in Mathematical Physics* **334**, 1617–1648, arXiv:1309.3228, (2015).

5.31 M. Mosonyi and T. Ogawa, "Two approaches to obtain the strong converse exponent of quantum hypothesis testing for general sequences of quantum states," *IEEE Transactions on Information Theory* **61**, 6975–6994, arXiv:1407.3567, (2014).

5.32 M. Mosonyi and T. Ogawa, "Strong Converse Exponent for Classical-Quantum Channel Coding," arXiv:1409.3562, (2014).

5.33 K. R. Parthasarathy, Probability Measures on Metric Spaces. Academic Press, New York and London, (1967).

5.34 A. Winter, "Coding Theorem and Strong Converse for Quantum Channels," *IEEE Transactions on Information Theory* **45**, 2481–2485, arXiv:1409.2536, (2014).

5.35 T. Ogawa and H. Nagaoka, "Strong converse to the quantum channel coding theorem," *IEEE Transactions on Information Theory* **45**, 2486–2489, arXiv:quant-ph/9808063, (1999).

5.36 B. Schumacher and M. D. Westmoreland, "Optimal signal ensembles," *Physical Review A* **63**, 022308, arXiv:quant-ph/9912122, (1999).

5.37 F. Dupuis, L. Kraemer, P. Faist, J. M. Renes, and R. Renner, "Generalized Entropies," *Proc. XVIIth International Congress on Mathematical Physics*, 134–153, arXiv:1211.3141, (2012).

5.38 D. Petz, "Quasi-entropies for finite quantum systems," *Reports on Mathematical Physics* **23**, 57–65, (1986).

5.39 M. S. Lin and M. Tomamichel, "Investigating Properties of a Family of Quantum Renyi Divergences," *Quantum Information Processing* **14**, 1501–1512, arXiv:1408.6897, (2014).

5.40 M. Nussbaum and A. Szkoła, "The Chernoff lower bound for symmetric quantum hypothesis testing," *The Annals of Statistics* **37**, 1040–1057, arXiv:quant-ph/0607216, (2009).

5.41 M. Fukuda, I. Nechita, and M. M. Wolf, "Quantum Channels With Polytopic Images and Image Additivity," *IEEE Transactions on Information Theory* **61**, 1851–1859, arXiv:1408.2340, (2015).

5.42 M. Hayashi and H. Nagaoka, "General Formulas for Capacity of Classical-Quantum Channels," *IEEE Transactions on Information Theory* **49**, 1753–1768, arXiv:quant-ph/0206186, (2003).

5.43 P. Hayden, D. Leung, P. W. Shor, and A. Winter, "Randomizing Quantum States: Constructions and Applications," *Communications in Mathematical Physics* **250**, 371–391, arXiv:quant-ph/03071, (2004).

5.44 Y. Polyanskiy, *Channel Coding: Non-Asymptotic Fundamental Limits*, PhD thesis, Princeton University, (2010).

5.45 M. Tomamichel and V. Y. F. Tan, "A Tight Upper Bound for the Third-Order Asymptotics for Most Discrete Memoryless Channels," *IEEE Transactions on Information Theory* **59**, 7041–7051, arXiv:1212.3689, (2013).

5.46 P. Moulin, "The Log-Volume of Optimal Codes for Memoryless Channels, Asymptotically Within A Few Nats," *IEEE Transactions on Information Theory*, **63**, 2278 – 2313, arXiv:1311.0181, (2017).

5.47 N. Datta and F. Leditzky, "Second-Order Asymptotics for Source Coding, Dense Coding, and Pure-State Entanglement Conversions," *IEEE Transactions on Information Theory* **61**, 582–608, arXiv:1403.2543, (2015).

5.48 F. Leditzky and N. Datta, "Second-Order Asymptotics of Visible Mixed Quantum Source Coding via Universal Codes," *IEEE Transactions on Information Theory* **62**, 4347–4355, arXiv:1407.6616, (2016).

5.49 N. Datta, M. Tomamichel, and M. M. Wilde, "On the second-order asymptotics for entanglement-assisted communication," *Quantum Information Processing* **15**, 2569–2591, arXiv:1405.1797, (2016).

5.50 M. Tomamichel, M. Berta, and J. M. Renes, "Quantum coding with finite resources," *Nature Communications* **7**:11419, arXiv:1504.04617, (2016).

5.51 M. M. Wilde, M. Tomamichel, and M. Berta, "Converse bounds for private communication over quantum channels," *IEEE Transactions on Information Theory*, **63**, 1792–1817, arXiv:1602.08898, (2017).

5.52 L. V. Rozovsky, "Estimate from Below for Large-Deviation Probabilities of a Sum of Independent Random Variables with Finite Variances," *Journal of Mathematical Sciences* **109**, 6, 2192–2209, (2002).

5.53 S.-H. Lee, V. Y. F. Tan, and A. Khisti, "Streaming Data Transmission in the Moderate Deviations and Central Limit Regimes," *IEEE Transactions on Information Theory* **62**, 6816–6830, arXiv:1512.06298, (2016).

Chapter 6

Beyond the thermodynamic limit: finite-size corrections to state interconversion rates

Christopher T. Chubb[1], Marco Tomamichel[2], and Kamil Korzekwa[1]

[1] *Centre for Engineered Quantum Systems, School of Physics, University of Sydney, Sydney NSW 2006, Australia*

[2] *Centre for Quantum Software and Information, School of Software, University of Technology Sydney, Sydney NSW 2007, Australia*

Quantum **2**, 108 (2018)
Doi: 10.22331/q-2018-11-27-108
ArXiv: 1711.01193

Abstract

Thermodynamics is traditionally constrained to the study of macroscopic systems whose energy fluctuations are negligible compared to their average energy. Here, we push beyond this thermodynamic limit by developing a mathematical framework to rigorously address the problem of thermodynamic transformations of finite-size systems. More formally, we analyse state interconversion under thermal operations and between arbitrary energy-incoherent states. We find precise relations between the optimal rate at which interconversion can take place and the desired infidelity of the final state when the system size is sufficiently large. These so-called second-order asymptotics provide a bridge between the extreme cases of single-shot thermodynamics and the asymptotic limit of infinitely large systems. We illustrate the utility of our results with several examples. We first show how thermodynamic cycles are affected by irreversibility due to finite-size effects. We then provide a precise expression for the gap between the distillable work and work of formation that opens away from the thermodynamic limit. Finally, we explain how the performance of a heat engine gets affected when the working body is small. We find that while perfect work cannot generally be extracted at Carnot efficiency, there are conditions under which these finite-size effects vanish. In deriving our results we also clarify relations between different notions of approximate majorisation.

6.1 Introduction

Background. Thermodynamics forms an integral part of contemporary physics, providing us with invaluable rules that govern which transformations between macroscopic states are possible and which are not [6.1]. In modern parlance thermodynamics is an example of a resource theory [6.2]. These provide us with a general framework to study the question of state interconversion by leveraging the structure imposed by the free states and operations within a given theory. The resource-theoretic approach has recently garnered renewed attention in the field of quantum information (see [6.3] for a recent review) and allows us to quantify notions like entanglement [6.4], coherence [6.5–6.7] and asymmetry [6.8, 6.9]. In the study of entanglement, for example, local operations and classical communication constitute the free operations and separable states are treated as free states. In an analogous way, a rigorous resource-theoretic formulation of quantum thermodynamics was provided in Refs. [6.10–6.12], with thermal Gibbs states being free and the laws of thermodynamics being captured by the restricted set of free operations known as *thermal operations*.

One of the main problems studied within the resource theory of thermodynamics is single-shot state interconversion, i.e., identifying when it is possible to convert a given state to another using only free operations (or, alternatively, identifying what extra resources are necessary to enable such transformations). Although for general quantum states only partial results are known [6.13–6.15] (see also Ref. [6.16] for the most recent progress), the full solution was found for the restricted case of transforming states with no coherence between distinct energy eigenspaces. For such *energy-incoherent states* the necessary and sufficient conditions for single-shot state interconversion are given by a thermomajorisation relation between the initial and final states [6.11]. As a result the allowed transformations are in general irreversible, which is captured by the fact that the amount of work needed to create a given state, the *work of formation*, is larger than the amount of work one can extract from it, the *distillable work*. We note that formally the thermomajorisation condition is strongly related to the majorisation condition appearing within the resource theory of entanglement when studying transformations between pure bipartite states [6.17].

An important variant of the interconversion problem, lying on the opposite extreme of the single-shot case, is asymptotic state interconversion. In this case one considers having access to arbitrarily many copies of the initial state, and asks for the maximal conversion rate at which it is possible to transform instances of one state to another with asymptotically vanishing error. It was found that this rate is given by the ratio of non-equilibrium free energies of the initial and target states [6.12]. Thus, the asymptotic interconversion rate is directly linked with the amount of useful work that can be extracted on average from a given state. Moreover, in this regime all transformations become fully reversible, as work of formation and distillable work asymptotically coincide. Again, this result closely resembles that obtained in pure state entanglement theory, where the optimal interconversion rate is given by the ratio of the entanglement entropies of the initial and target states [6.18].

In this work we study the interconversion problem in an intermediate regime, between the single-shot and asymptotic cases described above. We thus consider transformations of a finite number n of instances of the input state and tolerate a non-zero error, which affects the optimal conversion rate. By developing the notions of approximate majorisation [6.19] and thermomajorisation, we extend a formal relationship between the resource theories of pure state entanglement and thermodynamics of incoherent states to the approximate case. This allows us to adapt recently developed tools for approximate entanglement transformations in Ref. [6.20] to study corrections to the asymptotic rates

of thermodynamic transformations, the so-called second-order asymptotics (see, e.g., Refs [6.21–6.28] for other recent studies of second-order asymptotics in quantum information). The crucial technical difference between Ref. [6.20] and our work is the reversed direction of the majorisation relation, resulting in free states being given by uniform states rather than pure states. The second-order corrections were known to scale as $1/\sqrt{n}$ [6.12], and our main technical contribution is identifying the exact constant (including its dependence on the error) and interpreting its thermodynamic meaning.

Motivation. Probably the most famous thermodynamic result concerns the irreversible nature of thermodynamic transformations, and is often captured by the oversimplified statement "entropy has to grow". The dynamics of a system interacting with a thermal bath is irreversible since transformations performed at finite speed lead to heat dissipation, resulting in a loss of information about the system. One thus often studies idealised scenarios, when the system undergoes changes so slowly that it stays approximately in thermal equilibrium at all times. In this quasi-static limit one recovers reversible dynamics. However, thermodynamic reversibility actually requires one more assumption that is usually made implicitly. Namely, the thermodynamic description is only valid when applied to systems whose energy fluctuations are much smaller than their average energy. This is true for macroscopic systems composed of $n \to \infty$ particles, in the so-called thermodynamic limit, and is reflected by the reversibility of the interconversion problem in the asymptotic limit. However, for finite n the macroscopic results do not hold anymore, leading to another source of irreversibility.

In the emerging field of quantum thermodynamics (see Ref. [6.29] and references therein) a focus is placed on possible transformations of small quantum systems interacting with a thermal environment. The necessity to go beyond classical thermodynamics is motivated by the fact that at the nanoscale quantum effects, like coherence [6.13, 6.14, 6.30–6.33] and entanglement [6.34–6.36], start playing an important role. However, beyond these phenomena, in the quantum regime one also deals with systems composed of a finite number n of particles. Hence, thermodynamic transformations of such systems are affected by the effective irreversibility discussed in the previous paragraph. The results we present in this paper provide a mathematical framework to rigorously address this problem. We thus provide a bridge between the extreme case of single-shot thermodynamics with $n = 1$ and the asymptotic limit of $n \to \infty$, allowing us to study the irreversibility of thermodynamic processes in the intermediate regime of large but finite n.

Main Results. In order to state our main results we first need to introduce some concepts that will be defined more formally in Sections 6.2 and 6.3. Let us consider a finite-dimensional quantum system, characterised by its Hamiltonian H, in the presence of a thermal bath at fixed temperature T. The initial state of the system is in general out of thermal equilibrium, and the bath can be governed by an arbitrary Hamiltonian. Energy-conserving operations that interact the system with a bath in thermal equilibrium are then known as thermal operations. A simple example of a thermal operation just swaps the system with a bath (governed by the same Hamiltonian), replacing the initial state with a thermal Gibbs state. Gibbs states for H, denoted by γ, are thus free states in the resource-theoretic formulation of thermodynamics.

In the following we focus on a system comprised of a finite number n of non-interacting subsystems, each governed by the Hamiltonian H. Let us consider a pair of subsystem states ρ and σ that both

commute with H.[1] Our results will be expressed in terms of two information quantities: the relative entropy [6.37] with the Gibbs state, $D(\cdot\|\gamma)$, and the relative entropy variance [6.21, 6.22] with the Gibbs state, $V(\cdot\|\gamma)$. These quantities can also be interpreted thermodynamically: $k_B T \cdot D(\rho\|\gamma)$ is the difference between the generalised free energies of ρ and γ (with k_B denoting the Boltzmann constant); and $V(\rho\|\gamma)$ is proportional to a generalised heat capacity of the system. The latter interpretation is justified since for $\rho = \gamma'$ being a Gibbs state at temperature $T' \neq T$, the quantity $V(\gamma'\|\gamma)$ is proportional to the heat capacity at T'. We also note that $D(\rho\|\gamma)$ vanishes if and only if $\rho = \gamma$, whereas $V(\rho\|\gamma)$ vanishes whenever ρ is proportional to the Gibbs state on the support of ρ, e.g., when ρ is pure.

Let us now consider the problem of thermodynamic state interconversion between a finite number of instances of a state and for a fixed inverse temperature β of the background bath. Formally, we are looking for the maximal rate R for which there exists a thermal operation $\mathcal{E}^\beta$ such that $\mathcal{E}^\beta(\rho^{\otimes n}) = \tilde{\sigma}$ for some state $\tilde{\sigma}$ on Rn subsystems that is sufficiently close to $\sigma^{\otimes Rn}$. To measure the proximity of two quantum states we will use infidelity, i.e. we require that $F(\sigma^{\otimes Rn}, \tilde{\sigma}) \geq 1 - \epsilon$ for some accuracy parameter $\epsilon \in (0, 1)$, where $F(\cdot, \cdot)$ denotes Uhlmann's fidelity [6.38]. The maximal conversion rate, denoted by $R^*(n, \epsilon)$, depends on both the number of subsystems n and the accuracy ϵ. We can assume that neither the initial state ρ nor the target state σ are the thermal state γ, as otherwise the interconversion problem is trivial. We then find the following expansions of $R^*(n, \epsilon)$ in n:

$$R^*(n,\epsilon) \simeq \frac{D(\rho\|\gamma)}{D(\sigma\|\gamma)}\left(1 + \sqrt{\frac{V(\rho\|\gamma)}{n\,D(\rho\|\gamma)^2}}\,Z^{-1}_{1/\nu}(\epsilon)\right) \simeq \frac{D(\rho\|\gamma)}{D(\sigma\|\gamma)}\left(1 + \sqrt{\frac{V(\sigma\|\gamma)}{n\,D(\rho\|\gamma)D(\sigma\|\gamma)}}\,Z_\nu^{-1}(\epsilon)\right), \tag{6.1a}$$

where Z_ν^{-1} is the inverse of the cumulative function of Rayleigh-normal distribution Z_ν introduced in Ref. [6.20] with ν given by

$$\nu = \frac{V(\rho\|\gamma)/D(\rho\|\gamma)}{V(\sigma\|\gamma)/D(\sigma\|\gamma)}, \tag{6.2}$$

and $\simeq$ denotes equality up to terms of order $o(1/\sqrt{n})$. We note that $Z_0 = \Phi$ is the cumulative normal distribution function and Z_1 is the cumulative Rayleigh distribution function. The inverse of the cumulative Rayleigh-normal distribution is typically negative for small values of ϵ (unless $\nu = 1$), and thus the finite-size correction term that scales as $1/\sqrt{n}$ is generally negative. For the special case $V(\rho\|\gamma) = V(\sigma\|\gamma) = 0$ (when ν is undefined) we provide an exact formula for $R^*(n, \epsilon)$, up to all orders in n.

In deriving our results we also prove an important relation between two different notions of approximate majorisation [6.19]. More precisely, we show that *pre-* and *post-majorisation*, which hold when the majorisation relation holds up to the smoothing of the major*ising* or major*ised* distribution, are equivalent. We further extend these concepts to thermomajorisation, which allows us to rigorously address the problem of approximate thermodynamic transformations.

[1] In a particular case of trivial Hamiltonian $H \propto \mathbb{1}$, ρ and σ can be arbitrary quantum states.

Discussion. One of the main applications of our result is to the study of thermodynamic irreversibility. In the asymptotic limit, $n \to \infty$, the optimal conversion rate R^* from ρ to σ is equal to the inverse of the conversion rate from σ to ρ [6.12]. We can thus transform $\rho^{\otimes n}$ through $\sigma^{\otimes R^* n}$ back to $\rho^{\otimes n}$, so that the rate of concatenated transformations R_r^* is equal to 1 and the process can be performed reversibly. However, using Eq. (6.1a) twice, one finds the correction term to reversibility rate R_r^*, which is proportional to $1/\sqrt{n}$. Moreover, if $\nu \neq 1$ this correction term is negative for small errors. In fact, it diverges when the error approaches zero, preventing a perfect reversible cycle. However, pairs of states with equal ratios of relative entropy and relative entropy variance with respect to the thermal state (such that $\nu = 1$) are reversibly interconvertible up to second-order asymptotic corrections, mirroring a recent result in entanglement theory [6.39]. Thus, ν can be interpreted as the irreversibility parameter that quantifies the amount of infidelity of an approximate cyclic process.

One particular consequence of the discussed irreversibility is the difference between the distillable work, W_D, and the work of formation, W_F, for a given state ρ [6.11]. The former is defined as the maximal amount of free energy in the form of pure energy eigenstates ψ that can be obtained per copy of ρ; the latter as the minimal amount of free energy in the form of pure energy eigenstates ψ needed per copy to create the target state ρ. We note that in the special case when ψ is chosen to be the ground state, the distillation process can be considered as Landauer erasure (resetting to zero energy pure states), whereas the formation process can be seen as the action of a Szilard engine (creating states out of information). In single-shot thermodynamics W_D and W_F were shown to be proportional to max- and min-relative entropies with respect to the thermal state [6.11]; while in the asymptotic scenario they are both equal to $W = k_B T \cdot D(\rho\|\gamma)$, the non-equilibrium free energy of a state [6.12]. Here, using appropriately modified Eq. (6.1a), we show that for large n the values of distillable work and work of formation per particle lie symmetrically around the asymptotic value, $W_D \simeq W - \Delta W$ and $W_F \simeq W + \Delta W$, and provide the exact expression for the gap ΔW. Moreover, in the special case when the investigated state ρ is itself a thermal state at some temperature different from the background temperature, ΔW can be directly related to the relative strength of energy fluctuations of the system.

Finally, we investigate how one can investigate the performance of heat engines with finite working bodies using an appropriately chosen interconversion scenario. This allows us to study finite-size corrections to the efficiency of a heat engine and the quality of work it performs. More precisely, we consider a heat engine operating between two thermal baths at temperatures T_h and T_c, and a finite working body composed of n particles initially at temperature T_c. We show that, unless the irreversibility parameter satisfies $\nu = 1$, near-perfect work can be performed only with efficiency lower than the Carnot efficiency η_C. However, allowing for imperfect work allows one to achieve and even surpass η_C [6.40]. Moreover, we find that it is possible for a finite working body to have two thermal states at different temperatures, T_c and $T_{\mathrm{c}'}$, such that the irreversibility parameter for them is equal to 1. Thus, in a particularly engineered setting, it is possible to achieve Carnot efficiency and perform perfect work, while the finite working body changes temperature from T_c to $T_{\mathrm{c}'}$.

Overview. The remainder of this paper is organised as follows. We first describe the resource-theoretic approach to thermodynamics in Section 6.2 and introduce necessary mathematical concepts used within the paper in Section 6.3. In Section 6.4 we state our main result concerning state interconversion under thermal operations, and discuss its thermodynamic interpretation and possible

applications. We then proceed to Section 6.5, where we present auxiliary results concerning approximate majorisation and thermomajorisation, which we believe may be of independent interest. The technical proof of the main result can be found in Section 6.6. We conclude with an outlook in Section 6.7.

6.2 Thermodynamic setting

6.2.1 Thermal operations

We begin by describing the resource-theoretic approach to the thermodynamics of finite-dimensional quantum systems in the presence of a single heat bath at temperature T [6.11,6.41]. The investigated system is described by a Hamiltonian $H = \sum_i E_i |E_i\rangle\langle E_i|$ and prepared in a general state ρ, whereas the bath, with a Hamiltonian H_B, is in a thermal equilibrium state,

$$\gamma_B = \frac{e^{-\beta H_B}}{\operatorname{tr} e^{-\beta H_B}}, \tag{6.3}$$

where $\beta = 1/k_B T$ is the inverse temperature with k_B denoting the Boltzmann constant.[2] The evolution of the joint system is assumed to be closed, so that it is described by a unitary operator U, which additionally conserves the total energy,

$$[U, H + H_B] = 0. \tag{6.4}$$

The central question now is: what are the possible final states that a given initial state ρ can be transformed into?

More formally, one defines the set of *thermal operations* [6.10], which describes the free operations of the resource theory of thermodynamics, i.e., all possible transformations of the system that can be performed without the use of additional resources (beyond the single heat bath). These are defined as follows:

Definition 6.1 (Thermal operations)**.** *Given a fixed inverse temperature β, the set of thermal operations $\{\mathcal{E}^\beta\}$ consists of completely positive trace-preserving (CPTP) maps that act on a system ρ with Hamiltonian H as*

$$\mathcal{E}^\beta(\rho) = \operatorname{Tr}_B \left(U \left(\rho \otimes \gamma_B \right) U^\dagger \right), \tag{6.5}$$

with U satisfying Eq. (6.4), γ_B given by Eq. (6.3), and H_B being arbitrary.

Note that energy conservation condition, Eq. (6.4), can be interpreted as encoding the first law of thermodynamics; whereas the fact that the bath is in thermal equilibrium leads to $\mathcal{E}^\beta(\gamma) = \gamma$, with γ being the thermal Gibbs state of the system (i.e., given by Eq. (6.3) with H_B replaced by H), thus encoding the second law.

[2] Note that within the paper we will refer interchangeably to systems at temperature T_x or inverse temperature β_x.

6.2.2 Thermodynamic state interconversion

The thermodynamic interconversion problem is stated as follows: given a system (i.e., fixing H) together with initial and target states, ρ and σ, does there exist a thermal operation $\mathcal{E}^\beta$ (for a fixed β) such that $\mathcal{E}^\beta(\rho) = \sigma$? The general answer for such a question is not known beyond the simplest qubit case [6.13, 6.14] (however, we note that the problem has very recently been solved for a larger class of free operations given by *generalised thermal processes* [6.16] for coherent state interconversion). Nevertheless, for a restricted problem involving only *energy-incoherent states*, i.e., ρ and σ commuting with H, the set of necessary and sufficient conditions was found [6.11]. First, note that within this incoherent subtheory a quantum state can be equivalently represented by a probability distribution. For a non-degenerate Hamiltonian H, the initial and target states, ρ and σ, that commute with H are diagonal in the energy eigenbasis, so we can identify them with probability distributions $\boldsymbol{p}$ and $\boldsymbol{q}$, with $p_i = \langle E_i| \rho |E_i\rangle$ and $q_i = \langle E_i| \sigma |E_i\rangle$. For degenerate Hamiltonians we note that unitaries within a degenerate energy subspace are thermal operations, so one can always diagonalise a state within such subspace for free. Therefore, in a general case the components of $\boldsymbol{p}$ and $\boldsymbol{q}$ representing ρ and σ are simply given by the eigenvalues of ρ and σ. Next, in Ref. [6.10] (see also Refs. [6.11, 6.42] for an expanded discussion) the existence of a thermal operation between incoherent states was linked to the existence of a particular stochastic map via the following theorem.

Theorem 6.1 (Theorem 5 of Ref. [6.10])**.** *Let ρ and σ be quantum states commuting with the system Hamiltonian H, and γ its thermal equilibrium state. Denote their eigenvalues by $\boldsymbol{p}$, $\boldsymbol{q}$ and $\boldsymbol{\gamma}$, respectively. Then there exists a thermal operation $\mathcal{E}^\beta$ such that $\mathcal{E}^\beta(\rho) = \sigma$ if and only if there exists a stochastic map Λ^β such that*

$$\Lambda^\beta \boldsymbol{\gamma} = \boldsymbol{\gamma}, \quad \Lambda^\beta \boldsymbol{p} = \boldsymbol{q}\,. \tag{6.6}$$

As a result, studying thermodynamic interconversion problem between energy-incoherent states, one can replace CPTP maps and density matrices with stochastic matrices and probability vectors. We will fully address this simplified problem in Section 6.3.

In this paper we study a particular variant of the general interconversion problem: the limit of asymptotically many copies of input and output states. Informally, we want to find the optimal rate R^* allowing one to transform n copies of an energy-incoherent state ρ into R^*n copies of another energy-incoherent state σ, as n becomes large. Since the dimension of the input and output spaces must be the same[3], we note that one can append any number of states in thermal equilibrium, $\gamma^{\otimes m}$, to both the initial state $\rho^{\otimes n}$, and target state $\sigma^{\otimes Rn}$. Physically, it is motivated by the fact that thermal states are free resources; mathematically, it comes from the fact that the bath Hamiltonian H_B is arbitrary (so, in particular, it may contain m copies of the system, effectively adding $\gamma^{\otimes m}$ to the initial state) and that transforming any copy of the system into γ is a thermal operation (so that the part of the final state beyond $\sigma^{\otimes R^*n}$ can always be replaced by $\gamma^{\otimes m}$). Therefore, we ask for the maximal value of R^* (in the limit $n \to \infty$) for which there exists a thermal operation $\mathcal{E}^\beta$ satisfying

$$\mathcal{E}^\beta(\rho^{\otimes n} \otimes \gamma^{\otimes R^*n}) \approx \sigma^{\otimes R^*n} \otimes \gamma^{\otimes n}, \tag{6.7}$$

where $\approx$ denotes closeness in some distance measure, e.g. infidelity or trace norm.

[3] Strictly speaking this is only true because of our choice of Definition 6.1. More generally the partial trace in Eq. (6.5) may be taken over arbitrary subsystem, not necessarily B. However, as it does not affect our analysis, we believe that it is more compelling to keep the dimension of the system under study fixed.

As already mentioned in the Introduction, our focus here is on R^* for large but finite n, i.e., we look for corrections of order $1/\sqrt{n}$ to the optimal conversion rate coming from the finite number of systems involved in the thermodynamic process.

6.3 Mathematical preliminaries

6.3.1 Majorisation and embedding

Unless otherwise stated we consider d-dimensional probability distributions and their products. We define the uniform state $\boldsymbol{\eta}$ and the thermal Gibbs state $\boldsymbol{\gamma}$ at inverse temperature β as

$$\boldsymbol{\eta} := \frac{1}{d}[1, \ldots, 1], \tag{6.8a}$$

$$\boldsymbol{\gamma} := \frac{1}{\mathcal{Z}}\left[e^{-\beta E_1}, \ldots, e^{-\beta E_d}\right], \tag{6.8b}$$

with E_i denoting the eigenvalues of H and $\mathcal{Z} = \sum_i e^{-\beta E_i}$ being the partition function of the system. Moreover, we we call a distribution $\boldsymbol{f}$ *flat* if all its non-zero entries are equal. Note that in the infinite and zero temperature limits, $\beta \to 0$ and $\beta \to \infty$ respectively, the thermal state $\boldsymbol{\gamma}$ becomes flat. Specifically, in the former case $\boldsymbol{\gamma} \to \boldsymbol{\eta}$, and in the latter $\boldsymbol{\gamma} \to \boldsymbol{s} := [1, 0, \ldots, 0]$ for Hamiltonians with non-degenerate ground spaces.

The most general transformation between two probability distributions is given by a stochastic matrix Λ satisfying $\Lambda_{ij} \geq 0$ and $\sum_i \Lambda_{ij} = 1$. We denote by Λ^β a *Gibbs-preserving* stochastic matrix with a thermal fixed point, i.e., $\Lambda^\beta \boldsymbol{\gamma} = \boldsymbol{\gamma}$. In particular, a matrix Λ^0 that preserves the uniform distribution $\boldsymbol{\eta}$ is also called *bistochastic*. A probability vector $\boldsymbol{p}$ is said to majorise $\boldsymbol{q}$, denoted by $\boldsymbol{p} \succ \boldsymbol{q}$, if and only if

$$\sum_{i=1}^{k} p_i^{\downarrow} \geq \sum_{i=1}^{k} q_i^{\downarrow}, \tag{6.9}$$

for all $k \in \{1, \ldots, d\}$, with $\boldsymbol{p}^{\downarrow}$ denoting a probability vector with entries of $\boldsymbol{p}$ arranged in a non-increasing order. We then have the following central result that is used in the study of state interconversion:

Theorem 6.2 (Theorem II.1.10 of Ref. [6.43])**.** *There exists a bistochastic matrix mapping from $\boldsymbol{p}$ to $\boldsymbol{q}$ if and only if $\boldsymbol{p} \succ \boldsymbol{q}$, i.e.*

$$\exists \Lambda^0 : \ \Lambda^0 \boldsymbol{\eta} = \boldsymbol{\eta} \quad \text{and} \quad \Lambda^0 \boldsymbol{p} = \boldsymbol{q} \qquad \Longleftrightarrow \qquad \boldsymbol{p} \succ \boldsymbol{q}. \tag{6.10}$$

The above theorem can be generalised from bistochastic matrices Λ^0 to Gibbs-preserving matrices Λ^β with arbitrary β. To achieve this we first need to introduce the following embedding map.

Definition 6.2 (Embedding map)**.** *Given a thermal distribution $\boldsymbol{\gamma}$ with rational entries, $\gamma_i = D_i/D$ and $D_i, D \in \mathbb{N}$, the embedding map Γ^β sends a d-dimensional probability distribution $\boldsymbol{p}$ to a D-dimensional probability distribution $\hat{\boldsymbol{p}} := \Gamma^\beta(\boldsymbol{p})$ as follow:*

$$\hat{\boldsymbol{p}} = \Bigg[\underbrace{\frac{p_1}{D_1}, \ldots, \frac{p_1}{D_1}}_{D_1 \text{ times}}, \ \ldots \ , \underbrace{\frac{p_d}{D_d}, \ldots, \frac{p_d}{D_d}}_{D_d \text{ times}} \Bigg]. \tag{6.11}$$

The potentially irrational values of γ_i can be approached with arbitrarily high accuracy by choosing D large enough.

Note that $\hat{\boldsymbol{\gamma}} = \boldsymbol{\eta}^D$, i.e., embedding maps a thermal distribution into a uniform distribution over D entries, and that Γ^β is injective, implying the existence of a left inverse $(\Gamma^\beta)^{-1}$. The action of $(\Gamma^\beta)^{-1}$ on a D-dimensional vector $\boldsymbol{r}$ is given by summing up all the entries belonging to the same block of D_i entries. Moreover, $\Gamma^\beta(\Gamma^\beta)^{-1}$ is a bistochastic map, that transforms each block of D_i entries into a uniform distribution, i.e., given an index j belonging to a block D_i, denoted by $j \in [D_i]$, we have

$$[\Gamma^\beta(\Gamma^\beta)^{-1}\boldsymbol{r}]_j = \frac{\sum_{k\in[D_i]} r_k}{D_i}. \tag{6.12}$$

We can also introduce the embedded version of a matrix Λ^β,

$$\hat{\Lambda}^\beta := \Gamma^\beta \Lambda^\beta (\Gamma^\beta)^{-1}. \tag{6.13}$$

Notice that $\hat{\Lambda}^\beta$ is a bistochastic matrix, as it clearly maps the set of D-dimensional probability distributions into itself, and it preserves the uniform state $\boldsymbol{\eta}^D$,

$$\hat{\Lambda}^\beta \boldsymbol{\eta}^D = \Gamma^\beta \Lambda^\beta (\Gamma^\beta)^{-1} \hat{\boldsymbol{\gamma}} = \Gamma^\beta \Lambda^\beta \boldsymbol{\gamma} = \Gamma^\beta \boldsymbol{\gamma} = \boldsymbol{\eta}^D. \tag{6.14}$$

Using the notion of embedding we can define *thermomajorisation* relation [6.11] (originally introduced in Ref. [6.44] as d-majorisation). A probability vector $\boldsymbol{p}$ is said to thermomajorise $\boldsymbol{q}$, denoted by $\boldsymbol{p} \succ^\beta \boldsymbol{q}$, if and only if the majorisation relation holds between the embedded versions of $\boldsymbol{p}$ and $\boldsymbol{q}$, i.e.,

$$\boldsymbol{p} \succ^\beta \boldsymbol{q} \quad \Longleftrightarrow \quad \hat{\boldsymbol{p}} \succ \hat{\boldsymbol{q}}. \tag{6.15}$$

Note that for $\beta = 0$ thermomajorisation becomes standard majorisation, as the embedding map is the identity matrix. Now, we have the following equivalence

$$\Lambda^\beta \boldsymbol{p} = \boldsymbol{q} \quad \Longleftrightarrow \quad \hat{\Lambda}^\beta \hat{\boldsymbol{p}} = \hat{\boldsymbol{q}}, \tag{6.16}$$

which, since $\hat{\Lambda}^\beta$ is bistochastic, allows us to use Theorem 6.2 to obtain

Corollary 6.1 (Thermodynamic interconversion)**.** *There exists a Gibbs-preserving matrix mapping from $\boldsymbol{p}$ to $\boldsymbol{q}$ if and only if $\boldsymbol{p} \succ^\beta \boldsymbol{q}$, i.e.*

$$\exists \Lambda^\beta : \ \Lambda^\beta \boldsymbol{\gamma} = \boldsymbol{\gamma} \quad \text{and} \quad \Lambda^\beta \boldsymbol{p} = \boldsymbol{q} \qquad \Longleftrightarrow \qquad \boldsymbol{p} \succ^\beta \boldsymbol{q}. \tag{6.17}$$

Due to Theorem 6.1, Corollary 6.1 specifies the necessary and sufficient conditions for energy-incoherent state interconversion under thermal operations. In Fig. 6.1 we present the chain of equivalence relations leading to this result.

6.3.2 Information-theoretic notions and their thermodynamic interpretation

The *relative entropy* or Kullback-Leibler divergence of a probability distribution $\boldsymbol{p}$ with $\boldsymbol{q}$ is defined as

$$D(\boldsymbol{p}\|\boldsymbol{q}) := \sum_{i=1}^{d} p_i \log \frac{p_i}{q_i} \tag{6.18}$$

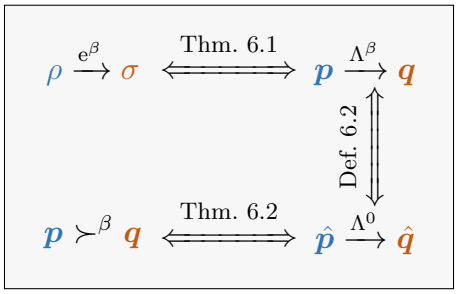


Figure 6.1: *Interconversion equivalence.* Quantum states ρ and σ are energy-incoherent, and their eigenvalues are given by $\boldsymbol{p}$ and $\boldsymbol{q}$, respectively. The arrow between states (distributions) symbolises the existence of a given map.

whenever the support of $\boldsymbol{q}$ contains the support of $\boldsymbol{p}$ (otherwise the divergence is set to $+\infty$). Denoting the average of a random variable X in a state $\boldsymbol{p}$ by

$$\langle X\rangle_{\boldsymbol{p}} = \sum_i p_i X_i, \tag{6.19}$$

we can introduce a random variable L with

$$\Pr\left[L = \log\frac{p_i}{q_i}\right] = p_i, \tag{6.20}$$

so that the divergence can be interpreted as the expectation value of the log-likelihood ratio, $D(\boldsymbol{p}\|\boldsymbol{q}) = \langle L\rangle_{\boldsymbol{p}}$. Similarly, we define the corresponding variance, the *relative entropy variance*, as

$$V(\boldsymbol{p}\|\boldsymbol{q}) := \mathrm{Var}_{\boldsymbol{p}}(L), \tag{6.21}$$

where

$$\mathrm{Var}_{\boldsymbol{p}}(X) := \langle (X - \langle X\rangle_{\boldsymbol{p}})^2\rangle_{\boldsymbol{p}}. \tag{6.22}$$

The following equalities are an immediate consequence of the embedding map introduced in Eq. (6.11):

$$D(\hat{\boldsymbol{p}}\|\hat{\boldsymbol{q}}) = D(\boldsymbol{p}\|\boldsymbol{q}) \qquad \text{and} \qquad V(\hat{\boldsymbol{p}}\|\hat{\boldsymbol{q}}) = V(\boldsymbol{p}\|\boldsymbol{q})\,. \tag{6.23}$$

In this work we will mostly encounter these quantities in the special case when $\boldsymbol{q} = \boldsymbol{\gamma}$ is the thermal distribution corresponding to inverse temperature β,

$$\gamma_i = \frac{e^{-\beta E_i}}{\mathcal{Z}}, \quad \mathcal{Z} = \sum_{i=1}^{d} e^{-\beta E_i}. \tag{6.24}$$

Then, both $D(\boldsymbol{p}\|\boldsymbol{\gamma})$ and $V(\boldsymbol{p}\|\boldsymbol{\gamma})$ can be interpreted thermodynamically. First note that

$$D(\boldsymbol{p}\|\boldsymbol{\gamma}) = \beta\langle E\rangle_{\boldsymbol{p}} - H(\boldsymbol{p}) + \log\mathcal{Z}, \tag{6.25}$$

with $\langle E\rangle_{\boldsymbol{p}}$ being the average energy and

$$H(\boldsymbol{p}) := -\sum_{i=1}^{d} p_i \log p_i \tag{6.26}$$

denoting the Shannon entropy of $\boldsymbol{p}$ (as a function of a distribution $\boldsymbol{p}$ it should not to be confused with the Hamiltonian H). Now, recall that the classical expression for free energy reads $U - TS$, with U being the average energy of the system, T the background temperature and S the thermodynamic entropy; and that the free energy of the thermal state is $-k_B T \log \mathcal{Z}$. We thus see that $D(\boldsymbol{p}||\boldsymbol{\gamma})/\beta$ can be interpreted as a non-equilibrium generalisation of free energy difference between an incoherent state ρ (represented by a probability distribution $\boldsymbol{p}$) and a thermal state γ.

Now, to interpret $V(\boldsymbol{p}\|\boldsymbol{\gamma})$ let us first introduce a covariance matrix M for the log-likelihood $\log p$ and energy in the units of temperature βE:

$$M = \begin{bmatrix} \mathrm{Cov}_{\boldsymbol{p}}(\beta E, \beta E) & \mathrm{Cov}_{\boldsymbol{p}}(\beta E, \log p) \\ \mathrm{Cov}_{\boldsymbol{p}}(\log p, \beta E) & \mathrm{Cov}_{\boldsymbol{p}}(\log p, \log p) \end{bmatrix}, \tag{6.27}$$

where $\mathrm{Cov}_{\boldsymbol{p}}(X, Y) = \langle XY\rangle_{\boldsymbol{p}} - \langle X\rangle_{\boldsymbol{p}}\langle Y\rangle_{\boldsymbol{p}}$. The relative entropy variance can then be expressed as

$$V(\boldsymbol{p}\|\boldsymbol{\gamma}) = \sum_{i,j=1}^{2} M_{ij}. \tag{6.28}$$

In a particular case when the distribution $\boldsymbol{p}$ is a thermal distribution $\boldsymbol{\gamma}'$ at some different temperature $T' \neq T$, the expression becomes

$$V(\boldsymbol{\gamma}'||\boldsymbol{\gamma}) = \left(1 - \frac{T'}{T}\right)^2 \cdot \frac{c_{T'}}{k_B}, \tag{6.29}$$

where

$$c_{T'} = \frac{\partial \langle E\rangle_{\boldsymbol{\gamma}'}}{\partial T'} \tag{6.30}$$

is the specific heat capacity of the system in a thermal state at temperature T'.

Finally, let us note that one can define quantum generalisations of both the relative entropy, $D(\rho||\sigma)$ [6.37], and the relative entropy variance, $V(\rho||\sigma)$ [6.21,6.22]. Moreover, for a quantum state ρ commuting with the Hamiltonian, these relative quantities with respect to the thermal state γ coincide with classical expressions,

$$D(\rho||\gamma) = D(\boldsymbol{p}||\boldsymbol{\gamma}), \qquad V(\rho||\gamma) = V(\boldsymbol{p}||\boldsymbol{\gamma}), \tag{6.31}$$

where $\boldsymbol{p}$ denotes the vector of eigenvalues of ρ.

6.3.3 Approximate interconversion

As already mentioned in Section 6.2 we will focus on approximate interconversions, allowing the final state $\tilde{\boldsymbol{q}}$ to differ from the target state $\boldsymbol{q}$, as long as it is close enough. We measure distance between states using the *infidelity*,

$$\delta(\boldsymbol{p}, \boldsymbol{q}) := 1 - F(\boldsymbol{p}, \boldsymbol{q}), \tag{6.32}$$

where the *fidelity* (or Bhattacharyya coefficient) is

$$F(\boldsymbol{p},\boldsymbol{q}) := \left(\sum_{i=1}^{d}\sqrt{p_i q_i}\right)^2. \tag{6.33}$$

We will also use the fidelity $\mathcal{F}$ between two (continuous) probability density functions $f(x)$ and $g(x)$, defined as

$$\mathcal{F}(f,g) := \left(\int_{-\infty}^{\infty}\sqrt{f(x)g(x)}\,\mathrm{d}x\right)^2. \tag{6.34}$$

The two important properties of the infidelity that we will use throughout the paper are as follows. First, since fidelity is non-decreasing under stochastic maps we have

$$\delta(\Lambda^0\boldsymbol{p},\Lambda^0\boldsymbol{q}) \leq \delta(\boldsymbol{p},\boldsymbol{q}). \tag{6.35}$$

Second, the distance δ between two probability vectors is the same as between their embedded versions, i.e.,

$$\delta(\boldsymbol{p},\boldsymbol{q}) = \delta(\hat{\boldsymbol{p}},\hat{\boldsymbol{q}}), \tag{6.36}$$

which can be verified by direct calculation.

Although we will be mainly concerned with "smoothing" the final distribution (allowing it to differ from the desired target one), it is useful to introduce two dual definitions of approximate majorisation and thermomajorisation.

Definition 6.3 (Pre- and post-thermomajorisation)**.** *A distribution $\boldsymbol{p}$ ϵ-*pre-thermomajorises *a distribution $\boldsymbol{q}$, which we denote $\boldsymbol{p}\,{}_{\epsilon}\!\succ^{\beta}\boldsymbol{q}$, if there exists a $\tilde{\boldsymbol{p}}$ such that*

$$\tilde{\boldsymbol{p}} \succ^{\beta} \boldsymbol{q} \qquad \textit{and} \qquad \delta(\boldsymbol{p},\tilde{\boldsymbol{p}}) \leq \epsilon. \tag{6.37}$$

*A distribution $\boldsymbol{p}$ ϵ-*post-thermomajorises *a distribution $\boldsymbol{q}$, which we denote $\boldsymbol{p} \succ^{\beta}_{\epsilon} \boldsymbol{q}$, if there exists a $\tilde{\boldsymbol{q}}$ such that*

$$\boldsymbol{p} \succ^{\beta} \tilde{\boldsymbol{q}} \qquad \textit{and} \qquad \delta(\boldsymbol{q},\tilde{\boldsymbol{q}}) \leq \epsilon. \tag{6.38}$$

In a particular case of $\beta = 0$, when thermomajorisation coincides with majorisation, we will speak of pre- and post-majorisation, denoted by ${}_{\epsilon}\!\succ$ and $\succ_{\epsilon}$, respectively.

Let us make a few comments about the above definition. First, notice that due to Corollary 6.1, $\boldsymbol{p}\,{}_{\epsilon}\!\succ^{\beta}\boldsymbol{q}$ means that in the vicinity of $\boldsymbol{p}$ there exists a state $\tilde{\boldsymbol{p}}$ and Λ^{β} that maps it to $\boldsymbol{q}$. Similarly, $\boldsymbol{p} \succ^{\beta}_{\epsilon} \boldsymbol{q}$ means that there exists Λ^{β} that maps $\boldsymbol{p}$ to $\tilde{\boldsymbol{q}}$, which lies in the vicinity of $\boldsymbol{q}$. We illustrate this in Fig. 6.2. Next note that both $\succ^{\beta}_{0}$ and ${}_{0}\!\succ^{\beta}$ reduce to thermomajorisation $\succ^{\beta}$, specifically $\succ^{0}_{0}$ and ${}_{0}\!\succ^{0}$ are equivalent to the standard majorisation relation $\succ$. Let us also mention that the concept of majorisation between smoothed distributions has been recently studied in Refs. [6.19, 6.45, 6.46]. Moreover, as with exact thermomajorisation, approximate thermomajorisation specifies the existence of a thermal operation between two energy-incoherent states. More precisely, due to Theorem 6.1, Corollary 6.1 and Definition 6.3, we have the following:

Corollary 6.2. *Let ρ and σ be quantum states commuting with the system Hamiltonian H. Denote their eigenvalues by $\boldsymbol{p}$ and $\boldsymbol{q}$, respectively. Then there exists a thermal operation $\mathcal{E}^\beta$ such that $\delta(\mathcal{E}^\beta(\rho), \sigma) \leq \epsilon$ if and only if $\boldsymbol{p} \succ^\beta_\epsilon \boldsymbol{q}$.*

Finally, let us make an important comment concerning approximate majorisation. Consider two distributions, $\boldsymbol{p}$ and $\boldsymbol{q}$, such that $\boldsymbol{p} \succ_\epsilon \boldsymbol{q}$. By definition there exists $\tilde{\boldsymbol{q}}$ close to $\boldsymbol{q}$ that is majorised by $\boldsymbol{p}$. As majorisation is invariant under permutations, $\boldsymbol{p}$ also majorises any distribution $\Pi\tilde{\boldsymbol{q}}$, where Π is arbitrary permutation. However, the fidelity between $\boldsymbol{q}$ and $\Pi\tilde{\boldsymbol{q}}$ is, in general, permutation-dependent. It is the largest, when the i-th largest entries of $\boldsymbol{q}$ and $\Pi\tilde{\boldsymbol{q}}$ coincide for all i, and so it is equal to $F(\boldsymbol{q}^\downarrow, \tilde{\boldsymbol{q}}^\downarrow)$. Therefore, for a given $\tilde{\boldsymbol{q}}$ satisfying some majorisation relation, we know that for every state $\boldsymbol{q}$ there exists $\Pi\tilde{\boldsymbol{q}}$ satisfying the same relation, and with

$$F(\Pi\tilde{\boldsymbol{q}}, \boldsymbol{q}) = F(\tilde{\boldsymbol{q}}^\downarrow, \boldsymbol{q}^\downarrow). \tag{6.39}$$

Thus, in the context of approximate majorisation, while calculating fidelities between any two states we will assume, without loss of generality, that they are ordered.

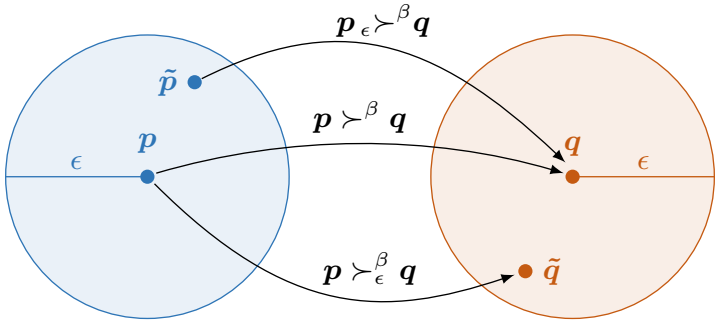


Figure 6.2: *Pre- and post-thermomajorisation.* Arrows depict the existence of Gibbs-preserving maps between corresponding distributions, whereas ϵ-circles represent sets of probability distributions whose distance δ from $\boldsymbol{p}$ and from $\boldsymbol{q}$ is smaller than ϵ.

6.3.4 Asymptotic notation

As we will be interested in approximating the optimal rate up to terms of order $O(1/\sqrt{n})$, we will adopt the following asymptotic notation for sequences $\{a_n\}_n$ and $\{b_n\}_n$ in $n \in \mathbb{N}$.

$$a_n \simeq b_n \quad \Longleftrightarrow \quad a_n - b_n = o(1/\sqrt{n})\,, \tag{6.40a}$$

$$a_n \lesssim b_n \quad \Longleftrightarrow \quad a_n - b_n \leq f_n = o(1/\sqrt{n})\,, \tag{6.40b}$$

$$a_n \gtrsim b_n \quad \Longleftrightarrow \quad a_n - b_n \geq g_n = o(1/\sqrt{n})\,, \tag{6.40c}$$

where $\{f_n\}_n$ and $\{g_n\}_n$ are auxiliary sequences that we usually do not introduce explicitly.

6.3.5 Rayleigh-normal distributions

The dependence of the finite-size corrections to optimal interconversion rate on the infidelity is given by generalisations of the Gaussian distribution known as the Rayleigh-normal distributions. This family of distributions was first introduced in Ref. [6.20] in the context of LOCC entanglement conversion. In order to define it, let us first denote the Gaussian cumulative distribution function, with mean value μ and variance ν, by $\Phi_{\mu,\nu}$,

$$\Phi_{\mu,\nu}(x) = \frac{1}{\sqrt{2\pi\nu}} \int_{-\infty}^{x} e^{-\frac{(t-\mu)^2}{2\nu}} \, \mathrm{d}t. \tag{6.41}$$

As a shorthand notation we will also use Φ to denote $\Phi_{0,1}$. Following Ref. [6.20] we can now define

Definition 6.4 (Rayleigh-normal distributions)**.** *For any $\nu > 0$ the Rayleigh-normal distribution is a distribution on $\mathbb{R}$, whose cumulative function is given by*

$$Z_\nu(\mu) := 1 - \sup_{A \geq \Phi} \mathcal{F}\left(A', \Phi'_{\mu,\nu}\right), \tag{6.42}$$

where the supremum is taken over all monotone increasing and continuously differentiable $A : \mathbb{R} \to [0,1]$ such that $A \geq \Phi$ pointwise; and $f'(x)$ denotes the derivative of $f(x)$.

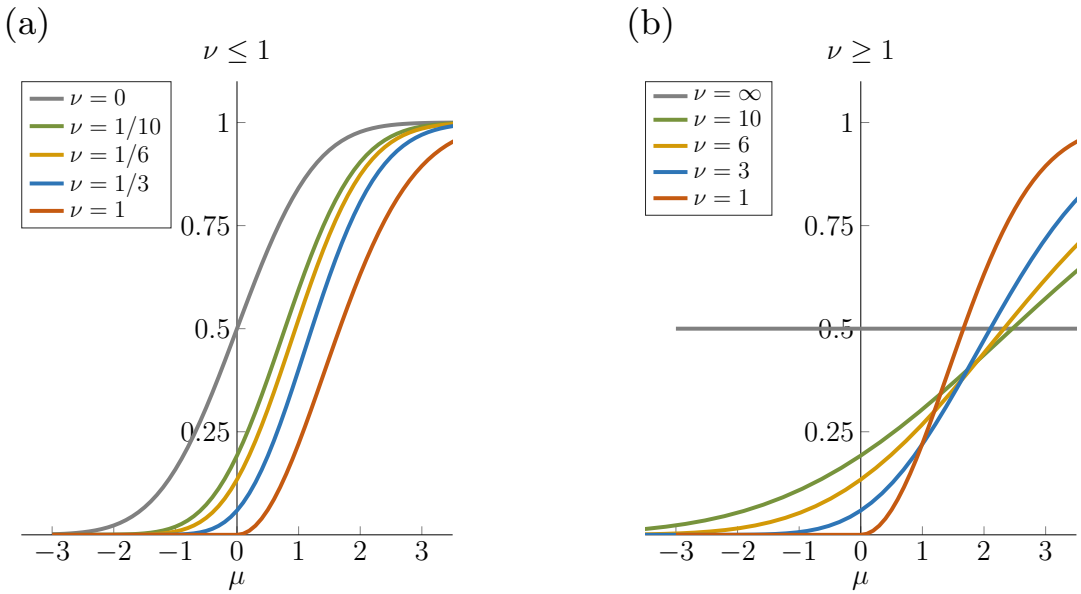


Figure 6.3: *Rayleigh-normal distributions.* Plots of Rayleigh-normal cumulative probability distributions introduced in Ref. [6.20]. Note the difference between the above graphs and those presented in Fig. 1 of Ref. [6.20]. The parameter is chosen in the ranges (a) $\nu \in [0,1]$ and (b) $\nu \in [1,\infty]$. Due to the duality property, Eq. (6.45), the plots can be directly related to the ones presented in panel (a).

We now present some relevant properties of the Rayleigh-normal distributions.

Lemma 6.3 (Section 2 of Ref. [6.20])**.** *The Rayleigh-normal distributions have the following properties:*

- *The $\nu \to 0$ case converges in distribution to the normal Gaussian,*

$$\lim_{\nu \to 0} Z_\nu(\mu) = \Phi(\mu). \tag{6.43}$$

- *The $\nu = 1$ case reduces to the Rayleigh distribution of scale parameter $\sigma = \sqrt{2}$,*

$$Z_1(\mu) = R_{\sqrt{2}}(\mu) := \begin{cases} 1 - e^{-\mu^2/4} & \mu \geq 0, \\ 0 & \mu \leq 0. \end{cases} \tag{6.44}$$

- *The Rayleigh-normal distributions possess a duality under inversion of the parameter ν of the form*

$$Z_{1/\nu}(\mu) = Z_\nu(\sqrt{\nu}\mu). \tag{6.45}$$

As well as these properties, an explicit form for the Rayleigh-normal distribution can be given. If we define $\alpha_{\mu,\nu}$ as the unique solution [6.20, Lemma 3] to

$$\frac{\Phi'(x)}{\Phi'_{\mu,\nu}(x)} = \frac{\Phi(x)}{\Phi_{\mu,\nu}(x)}, \tag{6.46}$$

and let

$$A_{\mu,\nu} := \begin{cases} \Phi_{\mu,\nu}(x) \dfrac{\Phi(\alpha_{\mu,\nu})}{\Phi_{\mu,\nu}(\alpha_{\mu,\nu})} & x \leq \alpha_{\mu,\nu}, \\ \Phi(x) & x \geq \alpha_{\mu,\nu}, \end{cases} \tag{6.47}$$

then for $\nu > 1$ we have

$$\arg\max_{A \geq \Phi} \mathcal{F}\left(A', \Phi'_{\mu,\nu}\right) = A'_{\mu,\nu}, \tag{6.48}$$

and so $Z_\nu(\mu) = 1 - \mathcal{F}\left(A'_{\mu,\nu}, \Phi'_{\mu,\nu}\right)$ [6.20, Theorem 4]. Using the duality property, Eq. (6.45), a similar expression can be given for $\nu \in (0, 1)$. We present plots of Rayleigh-normal distributions for a few selected values of ν in Fig. 6.3.

6.4 Second-order asymptotics for thermodynamic interconversion

6.4.1 Statement of the main result

We are now ready to state our main result concerning the second-order analysis of the approximate interconversion rates between independent and identically distributed (i.i.d.) states under thermal operations. We focus on initial and target states, ρ and σ, that commute with the Hamiltonian H, so that we can represent them as probability distributions, $\boldsymbol{p}$ and $\boldsymbol{q}$, over their eigenvalues. For two fixed distributions $\boldsymbol{p}$ and $\boldsymbol{q}$ we will be interested in the trade-off between three parameters in the asymptotic $n \to \infty$ regime: the rate of conversion R, the infidelity ϵ, and the inverse temperature of the bath β[4].

[4]Note that for fixed H the inverse temperature β fully specifies the thermal Gibbs distribution $\boldsymbol{\gamma}$.

Specifically, we will be interested in the triples (β, ϵ, R) for which there exist Gibbs-preserving maps Λ^β such that

$$\delta\Big(\Lambda^\beta\left(\boldsymbol{p}^{\otimes n}\otimes\boldsymbol{\gamma}^{\otimes Rn}\right),\left(\boldsymbol{q}^{\otimes Rn}\otimes\boldsymbol{\gamma}^{\otimes n}\right)\Big)\leq\epsilon, \tag{6.49}$$

where $\boldsymbol{\gamma}$ denotes the Gibbs state at inverse temperature β. By Corollary 6.1 and Definition 6.3 this condition is equivalent to approximate post-thermomajorisation,

$$\boldsymbol{p}^{\otimes n}\otimes\boldsymbol{\gamma}^{\otimes Rn}\succ_\epsilon^\beta\boldsymbol{q}^{\otimes Rn}\otimes\boldsymbol{\gamma}^{\otimes n}, \tag{6.50}$$

and, by Corollary 6.2, there exists a thermal operation transforming $\rho^{\otimes n}$ into a state ϵ away in the infidelity measure from $\sigma^{\otimes Rn}$.

We then define the optimal interconversion rate $R^*_\beta(n,\epsilon;\boldsymbol{p},\boldsymbol{q})$ and the optimal infidelity of interconversion $\epsilon^*_\beta(n,R;\boldsymbol{p},\boldsymbol{q})$ as

$$R^*_\beta := \max\left\{R\,\middle|\,\boldsymbol{p}^{\otimes n}\otimes\boldsymbol{\gamma}^{\otimes Rn}\succ_\epsilon^\beta\boldsymbol{q}^{\otimes Rn}\otimes\boldsymbol{\gamma}^{\otimes n}\right\}, \tag{6.51a}$$

$$\epsilon^*_\beta := \min\left\{\epsilon\,\middle|\,\boldsymbol{p}^{\otimes n}\otimes\boldsymbol{\gamma}^{\otimes Rn}\succ_\epsilon^\beta\boldsymbol{q}^{\otimes Rn}\otimes\boldsymbol{\gamma}^{\otimes n}\right\}. \tag{6.51b}$$

When it is clear from context we will drop the explicit dependence on $\boldsymbol{p}$ and $\boldsymbol{q}$. Our main result is then given by the following theorem.

Theorem 6.3 (Second-order asymptotic interconversion rates)**.** *Let ρ and σ be energy-incoherent initial and target states with eigenvalues given by $\boldsymbol{p}$ and $\boldsymbol{q}$, respectively. Then, for inverse temperature β and infidelity $\epsilon\in(0,1)$, the optimal interconversion rate has the following second-order expansions*

$$R^*(n,\epsilon)\simeq\frac{D(\boldsymbol{p}\|\boldsymbol{\gamma})}{D(\boldsymbol{q}\|\boldsymbol{\gamma})}\left(1+\sqrt{\frac{V(\boldsymbol{p}\|\boldsymbol{\gamma})}{n\,D(\boldsymbol{p}\|\boldsymbol{\gamma})^2}}\,Z^{-1}_{1/\nu}(\epsilon)\right) \tag{6.52a}$$

$$\simeq\frac{D(\boldsymbol{p}\|\boldsymbol{\gamma})}{D(\boldsymbol{q}\|\boldsymbol{\gamma})}\left(1+\sqrt{\frac{V(\boldsymbol{q}\|\boldsymbol{\gamma})}{n\,D(\boldsymbol{p}\|\boldsymbol{\gamma})D(\boldsymbol{q}\|\boldsymbol{\gamma})}}\,Z^{-1}_{\nu}(\epsilon)\right), \tag{6.52b}$$

where

$$\nu=\frac{V(\boldsymbol{p}\|\boldsymbol{\gamma})/D(\boldsymbol{p}\|\boldsymbol{\gamma})}{V(\boldsymbol{q}\|\boldsymbol{\gamma})/D(\boldsymbol{q}\|\boldsymbol{\gamma})} \tag{6.53}$$

is the irreversibility parameter.

The full proof of Theorem 6.3 can be found in Section 6.6. Before presenting it, we will discuss some of its consequences and applications in Section 6.4.2, and prove auxiliary results concerning approximate majorisation in Section 6.5. But first, let us make a few technical remarks about the above theorem. Note that Eq. (6.52a)-(6.52b) are simply related by the duality property of Rayleigh-normal distribution, Eq. (6.45). The reason to state both formulas is that this way one covers each of the special cases, $V(\boldsymbol{p}\|\boldsymbol{\gamma})=0$ and $V(\boldsymbol{q}\|\boldsymbol{\gamma})=0$, avoiding the use of Z^{-1}_∞, which is undefined. The special case when both relative entropy variances vanish is covered separately in Section 6.6.2, where

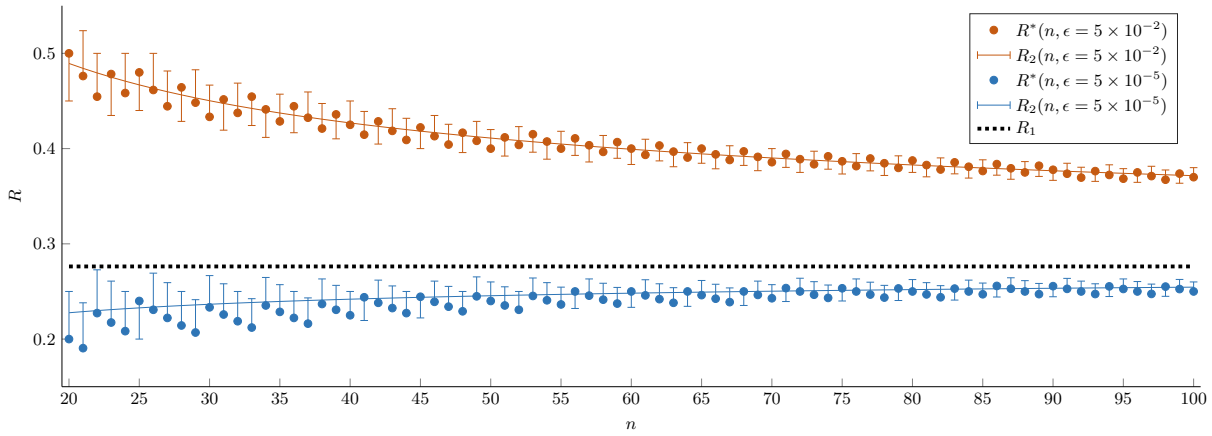


Figure 6.4: Comparison between the second-order approximation R_2 and exact thermal interconversion rates R^*, when converting from $\rho = \frac{7}{10}|0\rangle\langle 0| + \frac{3}{10}|1\rangle\langle 1|$ to $\sigma = \frac{8}{10}|0\rangle\langle 0| + \frac{2}{10}|1\rangle\langle 1|$, with Hamiltonian $H = |1\rangle\langle 1|$ and access to a thermal bath at temperature $1/\beta = 3$. The circles indicate exact conversion rates (c.f. Section 6.C), and the lines the second-order approximation given by Eq. (1). As the exact interconversion rate is always a multiple of $1/n$, we have also indicated the rounding of the second-order approximation to the nearest multiples of $1/n$ with error bars. The colours indicate the infidelity tolerance, with $\epsilon = 5 \times 10^{-2}$ for red and $\epsilon = 10^{-5}$ for blue. The dotted line indicates the asymptotic interconversion rate R_1. We plot the results for $n \leq 20$ in Figure 6.10.

an exact expression for $R^*(n, \epsilon)$ is provided (the asymptotic expansion of which coincides with the appropriate limit of Eq. (6.52a)).

Furthermore, since all the involved states are energy-incoherent, one can replace probability distributions $\boldsymbol{p}, \boldsymbol{q}, \boldsymbol{\gamma}$ in Theorem 6.3 with density matrices ρ, σ, γ. This way one can study interconversion between non-commuting states ρ and σ, as long as they both commute with H. For example, if the Hamiltonian is trivial, $H \propto \mathbb{1}$, ρ and σ may be arbitrary states. Thus, Theorem 6.3 yields a complete second-order analysis of interconversion under noisy operations [6.47], as for trivial Hamiltonians thermal operations coincide with noisy operations.

Finally, using results originally derived in Ref. [6.48] (see Section 6.B), one can numerically evaluate the optimal interconversion rates. In Section 6.C we show that this algorithm can be executed with a runtime that is efficient in the system size. Using this, in Figures 6.4 and 6.10 we can compare our second-order expansion to the exact interconversion rates. We find that even for relatively small system sizes, the second-order asymptotic expansion gives a remarkably good approximation to the optimal interconversion rates, especially when compared to the first-order asymptotics.

6.4.2 Discussion and applications

Although general state interconversion may seem to be a rather abstract problem, we will now show how the formalism can be applied to study more familiar thermodynamic scenarios. Since asymptotic conversion rates allow for reversible interconversion cycles and the results presented in the previous section describe finite-size corrections to these rates, our considerations will mainly revolve around irreversibility. We will first quantify it directly, by calculating the rate at which n copies of a system can be transformed from initial state ρ, through σ, and back to ρ. We will then discuss the gap

between work of formation and distillable work that opens when one processes finite number n of systems. Finally, we will apply our results to study the performance of heat engines operating with finite-size working bodies.

Finite-size reversibility

We start by considering the following thermodynamic process

$$\rho^{\otimes n} \to \sigma^{\otimes Rn} \to \rho^{\otimes R'Rn}, \tag{6.54}$$

with optimal interconversion rates given by

$$R = R^*_\beta(n, \epsilon_1; \boldsymbol{p}, \boldsymbol{q}), \quad R' = R^*_\beta(Rn, \epsilon_2; \boldsymbol{q}, \boldsymbol{p}), \tag{6.55}$$

where ρ and σ commute with the Hamiltonian and their eigenvalues are given by $\boldsymbol{p}$ and $\boldsymbol{q}$, respectively. Without the second-order asymptotic corrections derived in this work, the reversibility rate $R^*_r := RR'$ is equal to 1, and Eq. (6.54) describes a perfect cyclic process illustrated in Fig. 6.5a. However, including finite-size corrections, from Eq. (6.52a) we get

$$R^*_r \simeq \left(1 + \sqrt{\frac{V(\boldsymbol{p}||\boldsymbol{\gamma})}{nD(\boldsymbol{p}||\boldsymbol{\gamma})^2}} Z^{-1}_{1/\nu}(\epsilon_1)\right) \times \left(1 + \sqrt{\frac{V(\boldsymbol{q}||\boldsymbol{\gamma})}{RnD(\boldsymbol{q}||\boldsymbol{\gamma})^2}} Z^{-1}_{\nu}(\epsilon_2)\right),$$

with the irreversibility parameter ν given by Eq. (6.53). Now, using the duality of Rayleigh-normal distribution, Eq. (6.45), and ignoring the terms of order $o(1/\sqrt{n})$ we obtain

$$R^*_r \simeq 1 + \sqrt{\frac{V(\boldsymbol{p}||\boldsymbol{\gamma})}{nD(\boldsymbol{p}||\boldsymbol{\gamma})^2}} \left(Z^{-1}_{1/\nu}(\epsilon_1) + Z^{-1}_{1/\nu}(\epsilon_2)\right). \tag{6.56}$$

The error is accumulated during both transformations appearing in Eq. (6.54). However, since the infidelity δ is not a metric, we cannot simply add the errors. Instead, we note that $\sqrt{\delta}$ *is* a metric, and so it satisfies the triangle inequality. Thus, the total error ϵ, i.e., the infidelity between the final state and the target state $\rho^{\otimes R^*_r n}$, satisfies $\sqrt{\epsilon} \leq \sqrt{\epsilon_1} + \sqrt{\epsilon_2}$. Actually, for $\epsilon_1 + \epsilon_2 < 1$, one can obtain a tighter upper bound [6.49],

$$\epsilon \leq \left(\sqrt{\epsilon_1(1-\epsilon_2)} + \sqrt{\epsilon_2(1-\epsilon_1)}\right)^2. \tag{6.57}$$

Let us now introduce a threshold amount of infidelity,

$$\epsilon_0 := Z_{1/\nu}(0) = Z_\nu(0), \tag{6.58}$$

where the equality comes from duality of Rayleigh-normal distribution. Note that, if $\nu = 1$, resulting in $\epsilon_0 = 0$, then for any finite error one can eventually achieve $R^*_r > 1$, and a perfect transformation with $R^*_r = 1$ and arbitrarily small error can be achieved. Thus, pairs of states satisfying $\nu = 1$ are reversibly interconvertible up to second-order asymptotic corrections, analogously to a recent result in entanglement theory [6.39]. The use of such states in thermodynamic transformations is favourable, as it minimises the dissipation of free energy to the environment.

We will now show that the irreversibility parameter ν quantifies the incompatibility of two states (in that transformation from one state to the other leads to irreversibility) also beyond the special $\nu = 1$ case. Consider a process in which one requires that the number of systems n stays constant at all times. In other words, we require $R = R' = 1$, which implies

$$\epsilon_1 = \epsilon_2 = \epsilon_0. \tag{6.59}$$

Now, since $Z_\nu(0) \leq 1/2$, with equality achieved only for $\nu = 0$ and $\nu \to \infty$, the error rates satisfy Eq. (6.57) and the total error ϵ can be bounded by

$$\epsilon \leq 4\epsilon_0(1-\epsilon_0) = 4Z_\nu(0)\left(1 - Z_\nu(0)\right). \tag{6.60}$$

We present the above bound as a function of ν in Fig. 6.6a. We see that the closer ν is to 1, the less error will be induced while performing a thermodynamic transformation $\rho^{\otimes n} \to \sigma^{\otimes n} \to \rho^{\otimes n}$ or, in other words, the more reversible the process will be.

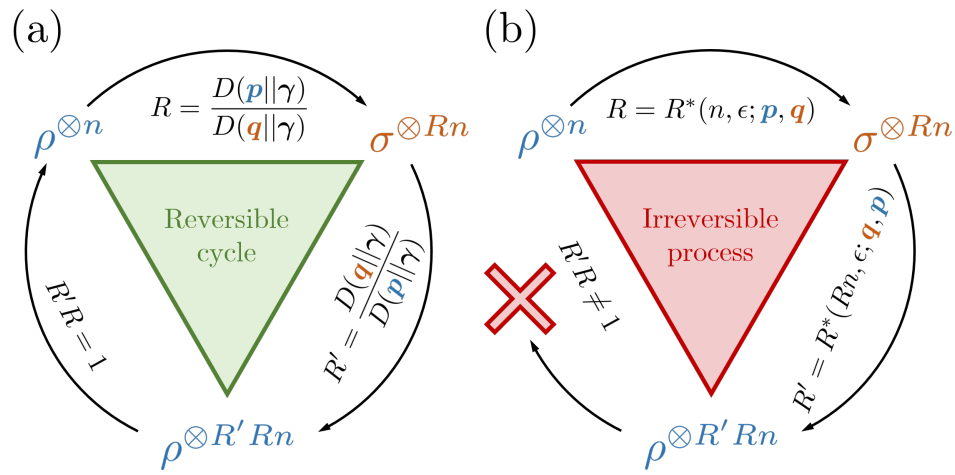


Figure 6.5: *Finite-size irreversibility.* (a) In the asymptotic limit, $n \to \infty$, the optimal conversion rate from ρ to σ is equal to the inverse of the conversion rate from σ to ρ. Therefore, reversible cycles can be performed. (b) In general, finite n corrections to conversion rates for near-perfect interconversion are negative, leading to irreversibility with $R'R < 1$.

Distillable work and work of formation gap

One particularly important consequence of irreversibility is the difference between distillable work and work of formation [6.11]. These quantify the amount of thermodynamically relevant resources that can be distilled from, or are needed to form, a given state. Similarly to the resource theory of entanglement, where Bell states act as standard units of entanglement resource [6.4], also within the resource theory of thermodynamics there are states acting as "gold standards" for measuring the amount of resources present in a state. These are given by pure energy eigenstates which, having zero entropy, have a clear energetic interpretation. The transformation requiring a change of an ancillary battery state $|w\rangle$, with energy w, into a state $|0\rangle$, with zero energy, is thus interpreted as performing work w; and

a transformation allowing for an opposite change corresponds to extracting work w. Hence, in order to assess the thermodynamic resourcefulness of n copies of a given energy-incoherent state, $\rho^{\otimes n}$, we will now investigate how much the energy of a pure battery system has to decrease per copy of ρ to construct $\rho^{\otimes n}$, and how much can it increase per copy of ρ while transforming $\rho^{\otimes n}$ to a thermal state?

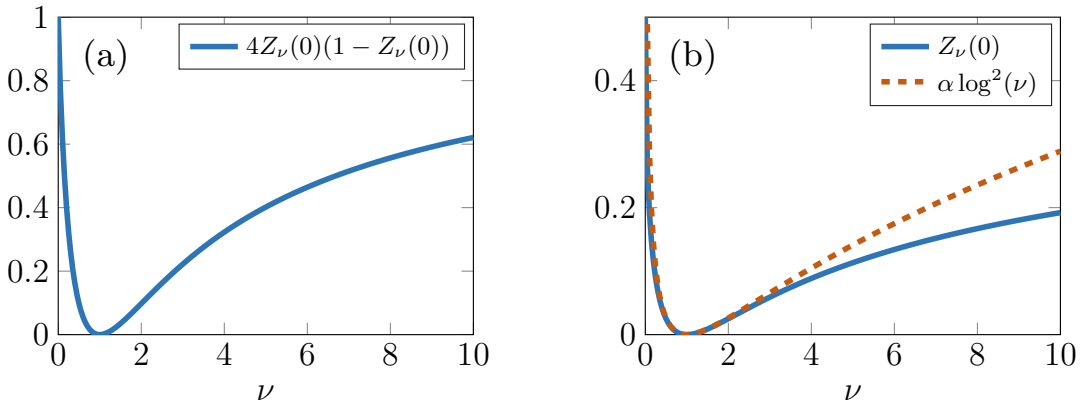


Figure 6.6: *Bounds on the infidelity of thermodynamic transformations.* All plots are symmetric under $\nu \to 1/\nu$ transformation. (a) The upper bound on the total error accumulated in a cyclic process $\rho^{\otimes n} \to \sigma^{\otimes n} \to \rho^{\otimes n}$ as a function of irreversibility parameter ν. (b) Infidelity generated by a heat engine working at the Carnot efficiency during a process that heats up the finite working body from T_c to $T_{c'}$ as a function of irreversibility parameter ν (that depends on both T_c and $T_{c'}$, as well as on the hot bath temperature through Eq. (6.53)). The optimal achievable infidelity during the process is plotted in solid line, while the bound on the infidelity generated during a continuous process (when the finite working body evolves through thermal states at all intermediate temperatures) is plotted in dashed line.

More formally, to calculate work of distillation W_D we want to find the maximal value w allowing for the thermodynamic transformation

$$(\rho \otimes |0\rangle\langle 0|)^{\otimes n} \to (\gamma \otimes |w\rangle\langle w|)^{\otimes n}\,, \tag{6.61}$$

where the second subsystem is a battery described by a Hamiltonian $H_B = 0\,|0\rangle\langle 0| + w\,|w\rangle\langle w|$. Similarly, to calculate work of formation W_F we want to find the minimal value w allowing for the thermodynamic transformation

$$(\gamma \otimes |w\rangle\langle w|)^{\otimes n} \to (\rho \otimes |0\rangle\langle 0|)^{\otimes n}\,. \tag{6.62}$$

Using Theorem 6.3 we can obtain the optimal rate for transformation described by Eq. (6.61) as a function of w, set it to 1 and solve for w, thus arriving at the approximate expression for the work of distillation:

$$W_D \approx k_B T \left(D(\boldsymbol{p}||\boldsymbol{\gamma}) + \sqrt{\frac{V(\boldsymbol{p}||\boldsymbol{\gamma})}{n}}\Phi^{-1}(\epsilon) \right), \tag{6.63}$$

with $\boldsymbol{p}$ denoting the eigenvalues of ρ and Φ^{-1} being the inverse of the normal Gaussian distribution. One can obtain the expression for the work of formation in an analogous way, this time looking for the optimal rate for transformation given in Eq. (6.62), resulting in

$$W_F \approx k_B T \left(D(\boldsymbol{p}||\boldsymbol{\gamma}) - \sqrt{\frac{V(\boldsymbol{p}||\boldsymbol{\gamma})}{n}}\Phi^{-1}(\epsilon) \right). \tag{6.64}$$

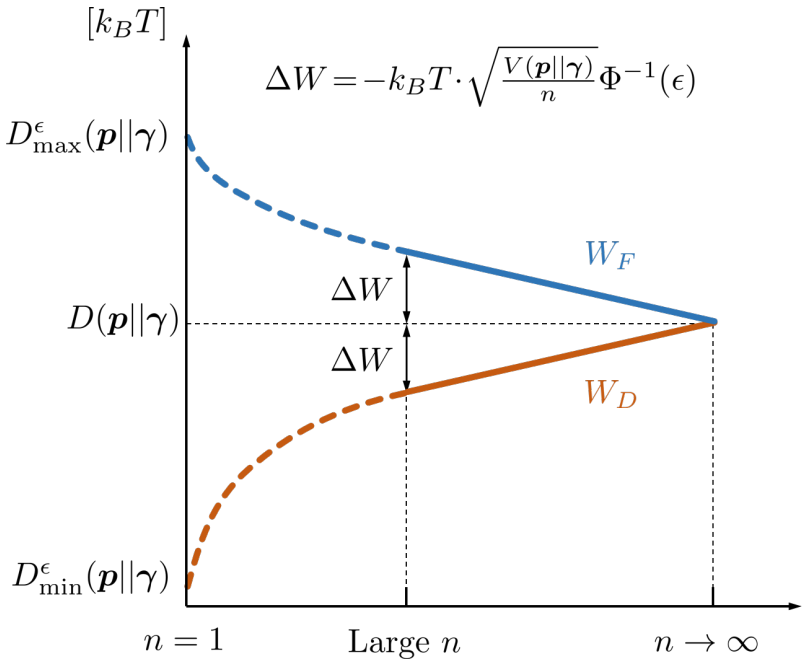


Figure 6.7: *Distillable work and work of formation gap.* The behaviour of distillable work W_D and work of formation W_F varies in different regimes. In single-shot scenarios they are proportional to min- and max-relative entropies [6.11]. In the intermediate regime of large but finite n studied in this work, the values of W_D and W_F lie symmetrically around the value achieved in the asymptotic limit, where W_D and W_F coincide and are equal to the non-equilibrium generalisation of free energy. Note that the y axis above is in the units of k_BT.

First of all, let us briefly comment on the effect that imperfect transformations (characterised by infidelity ϵ) have on the interpretation of distillable work and work of formation derived above. In the case of distillable work the non-zero infidelity means that the final battery state may differ from the pure state $|w\rangle\langle w|^{\otimes n}$, so one may actually distil less than W_D work per particle. Similarly, for the work of formation the final state of the battery may differ from the pure state $|0\rangle\langle 0|^{\otimes n}$, so one may actually use more than W_F work per particle (by using the purity of the battery). To overcome such problems, one may employ the idea of ϵ-deterministic work extraction [6.50] in the following way. After the distillation process (the argument for the formation process is analogous) one can simply measure the battery in its energy eigenbasis. With probability larger or equal to $1-\epsilon$ the battery state will collapse on $|w\rangle\langle w|^{\otimes n}$ (and so $n \cdot W_D$ work will be distilled), and with probability ϵ the work gain will differ from the derived value. Additionally, one has to take into account the thermodynamic cost of measuring the battery (erasing memory), proportional to the binary entropy of ϵ (note however that this cost is constant and so the cost per particle vanishes as $1/n$). Crucially, by choosing ϵ to be arbitrarily small, one can approach deterministic work distillation arbitrarily well, i.e., distil $n \cdot W_D$ work with probability arbitrarily close to 1.

Secondly, let us note that with the second-order asymptotic correction W_D and W_F lie symmetrically around the asymptotic value $W = k_BT \cdot D(\boldsymbol{p}||\boldsymbol{\gamma})$,

$$W_D \approx W - \Delta W, \quad W_F \approx W + \Delta W, \tag{6.65}$$

with

$$\Delta W := -k_B T \cdot \sqrt{\frac{V(\boldsymbol{p}||\boldsymbol{\gamma})}{n}}\Phi^{-1}(\epsilon). \tag{6.66}$$

Notice that the above correction term is positive for small values of infidelity ϵ, so that the resource cost of near-perfect formation of a state is always larger than the amount of resources than can be distilled from it. This symmetric gap that opens for finite n is illustrated in Fig. 6.7, where we also compare it with the values of W_D and W_F for the single-shot scenario $n = 1$ (where W_D and W_F generally lie asymmetrically around the asymptotic value W).

Furthermore, our second-order correction for distillable work exactly coincides with the one derived in Ref. [6.50] within an alternative thermodynamic framework, where state transformations are modelled by a sequence of energy level transformations (changes of Hamiltonian eigenvalues interpreted as performing/extracting work) and full thermalisations (replacing a state with the thermal state), rather than by thermal operations. This might have been expected, as the authors of Ref. [6.51] (and more recently of Ref. [6.52]) showed that any transformation between energy-incoherent states which can be achieved via a thermal operation, can also be achieved by a sequence of level transformations and partial level thermalisations.

Finally, let us analyse the special case when the state under scrutiny is itself a thermal equilibrium state $\boldsymbol{\gamma}'$, at some temperature T' different from the background temperature T. In the asymptotic limit $n \to \infty$, both distillable work W_D and work of formation W_F coincide with the standard thermodynamic result: the maximal (minimal) amount of work that can be extracted (needs to be invested) while changing the temperature of the system from T' to T (from T to T') is given by its free energy change. However, we also obtain the second-order asymptotic correction to W_D and W_F, given by

$$\Delta W = -|T - T'|\sqrt{\frac{k_B c_{T'}}{n}}\,\Phi^{-1}(\epsilon), \tag{6.67}$$

where we have used Eq. (6.29) to relate the relative entropy variance with $c_{T'}$, the heat capacity of the system at temperature T'. In order to interpret this correction term, we first note that standard thermodynamic results apply when fluctuations of energy are much smaller than the average energy of the system. Now, to quantify the relative strength of fluctuations, we introduce a fluctuation parameter f as a ratio of the total energy variation and the total energy itself,

$$f := \frac{\sqrt{n\mathrm{Var}_{\boldsymbol{\gamma}'}(E)}}{n\langle E\rangle_{\boldsymbol{\gamma}'}} = \sqrt{\frac{k_B c_{T'}}{n}} \cdot \frac{T'}{\langle E\rangle_{\boldsymbol{\gamma}'}}. \tag{6.68}$$

We then see that the correction term ΔW can be expressed as

$$\Delta W = -f w\,\Phi^{-1}(\epsilon), \qquad w := \langle E\rangle_{\boldsymbol{\gamma}'}\left|1 - \frac{T}{T'}\right|, \tag{6.69}$$

so that ΔW is directly related to the relative strength of fluctuations f, and disappears when the standard thermodynamic assumption, $f = 0$, holds. Note that $\Phi^{-1}(\epsilon)$ is negative for $\epsilon < 1/2$. Moreover, w is the amount of work performed by an engine operating at Carnot efficiency between two heat baths at temperatures T and T', when the amount of heat equal to $\langle E\rangle_{\boldsymbol{\gamma}'}$ flows in to, or out of, the bath at temperature T' (the former for $T > T'$, the latter for $T' > T$).

Corrections to efficiency of heat engines

One of the consequences of studying thermodynamics in the quantum regime is that it may not always be plausible to assume that working bodies and thermal reservoirs are infinite. Thus, recent studies focused on the effects finite-size baths have on standard thermodynamic results like fluctuation theorems [6.53], Landauer's principle [6.54], second [6.55] and third law of thermodynamics [6.56]. Moreover, the performance of heat engines operating between finite-size baths was investigated in Refs. [6.57, 6.58] (by directly focusing on the behaviour of thermodynamic quantities and not on the interconversion problem) and Ref. [6.59] (where the main focus was on the energetic structure of the finite bath and not on its size). Here, we will show how our results can be employed to investigate the performance of heat engines with finite working bodies, by studying the appropriately chosen interconversion problem. As we will discuss systems in equilibrium at different temperatures, we will indicate the (inverse) temperature in the subscript. More precisely, a system at temperature T_{x} (at inverse temperature β_{x}) will be denoted by γ_{x}, and the corresponding partition function by $\mathcal{Z}_{\mathrm{x}}$. Also, note that equilibrium states are diagonal in the energy eigenbasis, so our results are applicable.

We consider two infinite baths at temperatures $T_{\mathrm{h}} > T_{\mathrm{c}}$, and a finite working body composed of n particles initially at a cold temperature T_{c} (analogous considerations hold for the initial temperature being T_{h}). As in the previous subsection, we also include a battery system comprised of n two-level systems, each described by Hamiltonian $H_B = w\,|w\rangle\langle w|$, initially in a zero energy eigenstate $|0\rangle\langle 0|^{\otimes n}$. We now couple the working body at temperature T_{c} and the battery to the hot bath, allowing us to perform a thermal operation with respect to temperature T_{h}. In particular, we consider the following transformation

$$(\gamma_{\mathrm{c}} \otimes |0\rangle\langle 0|)^{\otimes n} \to (\gamma_{\mathrm{c}'} \otimes |w\rangle\langle w|)^{\otimes n}\,. \tag{6.70}$$

This transformation can be understood as a result of heat Q_{in} flowing from the hotter background bath into an engine; part of it, Q_{out}, then heats up the working body composed of n particles from T_{c} to $T_{\mathrm{c}'}$, while the remaining energy is used to perform work $n \cdot w$ on n particles comprising the battery. We schematically present this thermodynamic process in Fig. 6.8.

The heat Q_{out} flowing into the working body is given by the change of energy,

$$Q_{\mathrm{out}} = n\Delta E, \tag{6.71}$$

while the optimal amount of performed work $W = n \cdot w$ can be calculated similarly as in the previous subsection (by setting the rate from Eq. (6.52a) for the transformation given by Eq. (6.70) to 1 and solving for w), yielding

$$w \approx k_B T_{\mathrm{h}} \left(\Delta D + \sqrt{\frac{V(\boldsymbol{\gamma}_{\mathrm{c}}||\boldsymbol{\gamma}_{\mathrm{h}})}{n}}\, Z^{-1}_{1/\nu}(\epsilon) \right), \tag{6.72}$$

where we have introduced the following shorthand notation

$$\Delta E := \langle E\rangle_{\gamma_{\mathrm{c}'}} - \langle E\rangle_{\gamma_{\mathrm{c}}}, \tag{6.73a}$$
$$\Delta D := D(\boldsymbol{\gamma}_{\mathrm{c}}||\boldsymbol{\gamma}_{\mathrm{h}}) - D(\boldsymbol{\gamma}_{\mathrm{c}'}||\boldsymbol{\gamma}_{\mathrm{h}}), \tag{6.73b}$$

and $\nu = V(\boldsymbol{\gamma}_{\mathrm{c}}||\boldsymbol{\gamma}_{\mathrm{h}})/V(\boldsymbol{\gamma}_{\mathrm{c}'}||\boldsymbol{\gamma}_{\mathrm{h}})$. Now, using energy conservation,

$$Q_{\mathrm{in}} = Q_{\mathrm{out}} + W, \tag{6.74}$$

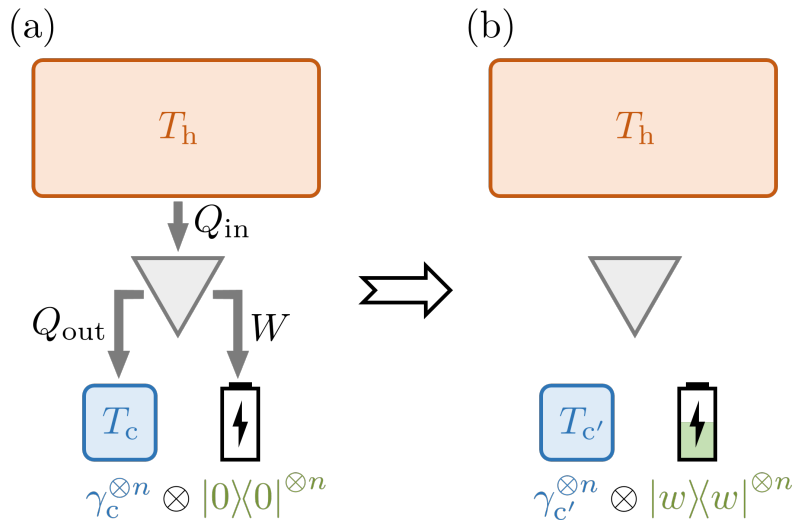


Figure 6.8: *Performance of a heat engine as an interconversion problem.* (a) The transformation of the working body (initially at cold temperature T_c) and the battery (initially in the ground state), $(\gamma_\mathrm{c} \otimes |0\rangle\langle 0|)^{\otimes n}$, can be seen as heat flowing in and out of the engine that performs work on the battery. (b) As a result of such interconversion, i.e., the heat engine performance, the finite working body ends up at the intermediate temperature $T_{\mathrm{c}'}$, while the battery gets transformed to an excited state $|w\rangle\langle w|^{\otimes n}$.

we can calculate the efficiency of the considered process to be

$$\eta = \frac{W}{Q_\mathrm{in}} = \left(1 + \frac{Q_\mathrm{out}}{W}\right)^{-1}, \tag{6.75}$$

with Q_out and W given by Eqs. (6.71) and (6.72), respectively.

In order to interpret the obtained expression let us first analyse the limiting case. Ignoring the second-order asymptotic correction (sending $n \to \infty$), the extracted work is just equal to the change of the free energy of the working body. In Appendix 6.A we show that this is exactly the amount of work that would be extracted by an engine operating at Carnot efficiency,

$$\eta_\mathrm{C}(T_\mathrm{x}) := 1 - \frac{T_\mathrm{x}}{T_\mathrm{h}}, \tag{6.76}$$

between an infinite bath at fixed temperature T_h and a colder finite bath that heats up during the process from $T_\mathrm{x} = T_\mathrm{c}$ to $T_\mathrm{x} = T_{\mathrm{c}'}$. In other words, without the $1/\sqrt{n}$ correction we obtain an *integrated* Carnot efficiency $\eta_\mathrm{C}^\mathrm{int}$ that arises from an *instantaneous* Carnot efficiency η_C at all times,

$$\eta_\mathrm{C}^\mathrm{int}(T_\mathrm{c} \to T_{\mathrm{c}'}) = \left(1 + \frac{\Delta E}{k_B T_\mathrm{h} \Delta D}\right)^{-1}. \tag{6.77}$$

The relation becomes even more evident when we consider the limit $\Delta T \to 0$. Then

$$\Delta E \xrightarrow{\Delta T \to 0} c_{T_\mathrm{c}} \Delta T, \tag{6.78a}$$

$$\Delta D \xrightarrow{\Delta T \to 0} c_{T_\mathrm{c}} \Delta T \left(\frac{1}{k_B T_\mathrm{c}} - \frac{1}{k_B T_\mathrm{h}}\right), \tag{6.78b}$$

with c_{T_c} denoting the heat capacity of the system at temperature T_c, and so

$$\eta_C^{\text{int}}(T_c \to T_{c'}) \xrightarrow{\Delta T \to 0} \eta_C(T_c). \tag{6.79}$$

Now, the finite-size correction leads to a modified expression for integrated efficiency, given by

$$\eta^{\text{int}} \approx \eta_C^{\text{int}} + \frac{k_B(T_h - T_c)\Delta E}{(k_B T_h \Delta D + \Delta E)^2}\sqrt{\frac{c_{T_c}}{nk_B}} Z_{1/\nu}^{-1}(\epsilon), \tag{6.80}$$

where we have used Eq. (6.29) again to relate the relative entropy variance with the heat capacity of the system. We first note that there exists a threshold amount of infidelity ϵ_0, given by Eq. (6.58), below which the correction term is negative. Since the infidelity between final and target states can be interpreted as performing imperfect work, near-perfect work can be performed only with efficiency strictly smaller than η_C^{int}. On the other hand, accepting infidelity $\epsilon \geq \epsilon_0$ allows one to achieve and even go beyond the integrated efficiency corresponding to instantaneous Carnot efficiency. This is in accordance with a recent result showing that the Carnot efficiency can be surpassed by extracting imperfect work [6.40].

As in the asymptotic limit, we also want to investigate the instantaneous efficiency, when $T_{c'}$ is very close to T_c. In particular, we will focus on the quality of performed work when the engine works at instantaneous Carnot efficiency. We thus require that the error $\Delta\epsilon$ accumulated during an infinitesimal step that changes the temperature by $\Delta T \to 0$ is equal to the threshold error ϵ_0. Since then the correction term vanishes, we have $\eta^{\text{int}} \approx \eta_C^{\text{int}}$ and we know that for small ΔT this yields η_C. Because the two considered thermal states are close, ν is close to unity, and as such the infidelity of the process can be expanded as

$$\Delta\epsilon = Z_\nu(0) = Z_{1+\Delta\nu}(0) \approx \alpha\Delta\nu^2, \tag{6.81}$$

where $\alpha \approx 0.0545$ can be numerically evaluated. The expansion of $\Delta\nu$ in terms of ΔT is given by

$$\Delta\nu \approx g(T_c)\Delta T, \tag{6.82}$$

with

$$g(T_c) := \frac{\mathrm{d}}{\mathrm{d}T_x}\left[\log V(\boldsymbol{\gamma}_x||\boldsymbol{\gamma}_h)\right]\Big|_{T_x=T_c}. \tag{6.83}$$

As discussed in Section 6.4.2, it is not infidelity, but its square root that satisfies the triangle inequality. We thus have that the instantaneous rate of accumulating square root infidelity is given by

$$\frac{\mathrm{d}\sqrt{\epsilon}}{\mathrm{d}T_x} \approx \sqrt{\alpha}|g(T_x)|, \tag{6.84}$$

and so one can achieve the instantaneous Carnot efficiency by paying the price of an instantaneous rate of accumulating error. This can be then translated into the bound on the total accumulated error in the following way

$$\epsilon \leq \left(\int_{T_c}^{T_{c'}} \frac{\mathrm{d}\sqrt{\epsilon}}{\mathrm{d}T_x}\,\mathrm{d}T_x\right)^2 = \alpha\left(\int_{T_c}^{T_{c'}} |g(T_x)|\,\mathrm{d}T_x\right)^2. \tag{6.85}$$

It is straightforward to show that this upper bound is larger than $\alpha \log^2 \nu$, which in turn is larger than $Z_\nu(0)$. This shows that the error accumulated in a continuous process (with the working body continuously passing through all intermediate temperatures) is in general larger than that of an optimal "one-step" process. We illustrate this in Fig. 6.6b.

Finally, let us comment on a special case when $\nu = 1$. For initial and target states being thermal equilibrium states at distinct temperatures (and different from background temperature T_h), the value of ν depends on the Hamiltonian of the investigated system. If it happens that for a given Hamiltonian there exist T_x and $T_{\mathrm{x}'}$ such that $\nu = 1$, then it is possible to achieve perfect work extraction at integrated Carnot efficiency $\eta_\mathrm{C}^{\mathrm{int}}(T_\mathrm{x} \to T_{\mathrm{x}'})$. Interestingly, for any Hamiltonian there always exist such pairs of temperatures. To see this note that for both $T_\mathrm{x} = 0$ and $T_\mathrm{x} = T_\mathrm{h}$ the relative entropy variance vanishes, $V(\boldsymbol{\gamma}_\mathrm{x}||\boldsymbol{\gamma}_\mathrm{h}) = 0$. Since it is a continuous function of temperature, we get that for any T_x in the interval $(0, T_\mathrm{h})$ there exists at least one other temperature $T_{\mathrm{x}'}$ such that $V(\boldsymbol{\gamma}_\mathrm{x}||\boldsymbol{\gamma}_\mathrm{h}) = V(\boldsymbol{\gamma}_{\mathrm{x}'}||\boldsymbol{\gamma}_\mathrm{h})$, resulting in $\nu = 1$. This shows that by appropriately choosing the temperatures between which the heat engine operates, one may decrease or even avoid irreversible losses.

6.5 Results on approximate majorisation

We now proceed to the presentation of a few technical lemmas that may be of independent interest. These concern relations between different notions of approximate majorisation and thermomajorisation introduced in Section 6.3.3. We first need the following auxiliary result.

Lemma 6.4. *For fixed probability vectors $\boldsymbol{p}$ and $\boldsymbol{q}$ denote by $\tilde{\boldsymbol{p}}$ any distribution that majorises $\boldsymbol{q}$, and by $\tilde{\boldsymbol{q}}$ any distribution that is majorised by $\boldsymbol{p}$. Then the maximum fidelity between $\tilde{\boldsymbol{p}}$ and $\boldsymbol{p}$ over all such $\tilde{\boldsymbol{p}}$ is equal to the maximum fidelity between $\boldsymbol{q}$ and $\tilde{\boldsymbol{q}}$ over all such $\tilde{\boldsymbol{q}}$, i.e.,*

$$\max_{\tilde{\boldsymbol{p}}:\ \tilde{\boldsymbol{p}} \succ \boldsymbol{q}} F(\boldsymbol{p}, \tilde{\boldsymbol{p}}) = \max_{\tilde{\boldsymbol{q}}:\ \boldsymbol{p} \succ \tilde{\boldsymbol{q}}} F(\boldsymbol{q}, \tilde{\boldsymbol{q}}). \tag{6.86}$$

The proof of the above lemma is based on the results first derived in Ref. [6.48] and can be found in Appendix 6.B. Moreover, the proof includes an explicit construction of the state $\tilde{\boldsymbol{p}}^\star$ maximising the left hand side of Eq. (6.86), so that one can calculate the value of optimal achievable fidelities appearing in Lemma 6.4.

Now we can prove the following crucial result concerning pre- and post-majorisation, i.e., approximate thermomajorisation for $\beta = 0$.

Lemma 6.5. *Pre- and post-majorisation are equivalent, i.e., $\boldsymbol{p} \,{}_\epsilon\!\succ \boldsymbol{q}$ if and only if $\boldsymbol{p} \succ_\epsilon \boldsymbol{q}$.*

Proof. First, assume that $\boldsymbol{p} \,{}_\epsilon\!\succ\, \boldsymbol{q}$. This means that there exists $\tilde{\boldsymbol{p}}$ such that $\tilde{\boldsymbol{p}} \succ \boldsymbol{q}$ and $\delta(\boldsymbol{p}, \tilde{\boldsymbol{p}}) \leq \epsilon$. By Theorem 6.2 this implies that there exists a bistochastic Λ^0 such that $\Lambda^0 \tilde{\boldsymbol{p}} = \boldsymbol{q}$. Let $\tilde{\boldsymbol{q}} := \Lambda^0 \boldsymbol{p}$, so that $\boldsymbol{p} \succ \tilde{\boldsymbol{q}}$. Using the fact that fidelity is non-decreasing under stochastic maps, we then have

$$\delta(\boldsymbol{q}, \tilde{\boldsymbol{q}}) = \delta(\Lambda^0 \tilde{\boldsymbol{p}}, \Lambda^0 \boldsymbol{p}) \leq \delta(\tilde{\boldsymbol{p}}, \boldsymbol{p}) \leq \epsilon, \tag{6.87}$$

which means that $\boldsymbol{p} \,{}_\epsilon\!\succ\, \boldsymbol{q} \implies \boldsymbol{p} \succ_\epsilon \boldsymbol{q}$.

Now, assume that $\boldsymbol{p} \succ_\epsilon \boldsymbol{q}$. This means that there exists $\tilde{\boldsymbol{q}}$ such that $\boldsymbol{p} \succ \tilde{\boldsymbol{q}}$ and $\delta(\boldsymbol{q}, \tilde{\boldsymbol{q}}) \leq \epsilon$. Let

$$\tilde{\boldsymbol{p}}^\star := \operatorname*{arg\,max}_{\tilde{\boldsymbol{p}}:\ \tilde{\boldsymbol{p}} \succ \boldsymbol{q}} F(\boldsymbol{p}, \tilde{\boldsymbol{p}}). \tag{6.88}$$

By definition $\tilde{\boldsymbol{p}}^\star \succ \boldsymbol{q}$ and, by Lemma 6.4, we have

$$F(\boldsymbol{p}, \tilde{\boldsymbol{p}}^\star) = \max_{\boldsymbol{q}' \colon \boldsymbol{p} \succ \boldsymbol{q}'} F(\boldsymbol{q}, \boldsymbol{q}') \geq F(\boldsymbol{q}, \tilde{\boldsymbol{q}}), \tag{6.89}$$

so that $\delta(\boldsymbol{p}, \tilde{\boldsymbol{p}}^\star) \leq \epsilon$. Thus $\boldsymbol{p} \mathrel{{}_\epsilon\!\succ} \boldsymbol{q} \impliedby \boldsymbol{p} \succ_\epsilon \boldsymbol{q}$. □

The next lemma links post-majorisation of embedded vectors with post-thermomajorisation for $\beta \neq 0$.

Lemma 6.6. *Post-majorisation between embedded vectors is equivalent to post-thermomajorisation between the original vectors, i.e.,*

$$\hat{\boldsymbol{p}} \succ_\epsilon \hat{\boldsymbol{q}} \qquad \iff \qquad \boldsymbol{p} \succ_\epsilon^\beta \boldsymbol{q} \tag{6.90}$$

Proof. First, assume $\hat{\boldsymbol{p}} \succ_\epsilon \hat{\boldsymbol{q}}$. This means that there exists a bistochastic matrix Λ^0 such that $\Lambda^0 \hat{\boldsymbol{p}} = \tilde{\hat{\boldsymbol{q}}}$ with $\delta(\hat{\boldsymbol{q}}, \tilde{\hat{\boldsymbol{q}}}) \leq \epsilon$. This, in turn, means that

$$\Lambda^\beta \boldsymbol{p} = \Gamma^{-1} \tilde{\hat{\boldsymbol{q}}}, \tag{6.91}$$

with $\Lambda^\beta = \Gamma^{-1} \Lambda^0 \Gamma$ being a Gibbs-preserving matrix and Γ a shorthand notation of the embedding map Γ^β. We thus conclude that $\boldsymbol{p} \succ^\beta \Gamma^{-1} \tilde{\hat{\boldsymbol{q}}}$. It remains to show that $\delta(\boldsymbol{q}, \Gamma^{-1} \tilde{\hat{\boldsymbol{q}}}) \leq \epsilon$. To achieve this we use the facts that embedding is fidelity-preserving, $\Gamma\Gamma^{-1}$ is bistochastic and fidelity is non-decreasing under stochastic maps, so that

$$\begin{aligned} F(\boldsymbol{q}, \Gamma^{-1} \tilde{\hat{\boldsymbol{q}}}) &= F(\Gamma^{-1} \hat{\boldsymbol{q}}, \Gamma^{-1} \tilde{\hat{\boldsymbol{q}}}) \\ &= F(\Gamma\Gamma^{-1} \hat{\boldsymbol{q}}, \Gamma\Gamma^{-1} \tilde{\hat{\boldsymbol{q}}}) \geq F(\hat{\boldsymbol{q}}, \tilde{\hat{\boldsymbol{q}}}). \end{aligned} \tag{6.92}$$

Therefore $\delta(\boldsymbol{q}, \Gamma^{-1} \tilde{\hat{\boldsymbol{q}}}) \leq \epsilon$, which results in $\boldsymbol{p} \succ_\epsilon^\beta \boldsymbol{q}$.

Now, assume that $\boldsymbol{p} \succ_\epsilon^\beta \boldsymbol{q}$. This means that there exists a Gibbs-preserving matrix Λ^β such that $\Lambda^\beta \boldsymbol{p} = \tilde{\boldsymbol{q}}$ with $\delta(\boldsymbol{q}, \tilde{\boldsymbol{q}}) \leq \epsilon$. Through embedding this is equivalent to the existence of a bistochastic $\hat{\Lambda}^\beta$ such that $\hat{\Lambda}^\beta \hat{\boldsymbol{p}} = \tilde{\hat{\boldsymbol{q}}}$, resulting in $\hat{\boldsymbol{p}} \succ \tilde{\hat{\boldsymbol{q}}}$. It remains to show that $\delta(\hat{\boldsymbol{q}}, \tilde{\hat{\boldsymbol{q}}}) \leq \epsilon$. This, however, follows directly from the fact that the embedding map is fidelity-preserving, as $\delta(\hat{\boldsymbol{q}}, \tilde{\hat{\boldsymbol{q}}}) = \delta(\boldsymbol{q}, \tilde{\boldsymbol{q}}) \leq \epsilon$. We thus conclude that $\hat{\boldsymbol{p}} \succ_\epsilon \hat{\boldsymbol{q}}$. □

The statement of Lemma 6.6 can be rephrased as

$$\exists \tilde{\hat{\boldsymbol{q}}} : \ \hat{\boldsymbol{p}} \succ \tilde{\hat{\boldsymbol{q}}}, \ \delta(\hat{\boldsymbol{q}}, \tilde{\hat{\boldsymbol{q}}}) \leq \epsilon \qquad \iff \qquad \exists \tilde{\boldsymbol{q}} : \ \hat{\boldsymbol{p}} \succ \hat{\tilde{\boldsymbol{q}}}, \ \delta(\hat{\boldsymbol{q}}, \hat{\tilde{\boldsymbol{q}}}) \leq \epsilon, \tag{6.93}$$

so that it can be interpreted as the fact that embedding (denoted by hat) and smoothing (denoted by tilde) commute when applied to the target state.

Finally, we present a result that links pre-majorisation of embedded vectors with pre-thermomajorisation for $\beta \neq 0$.

Lemma 6.7. *Pre-majorisation between embedded vectors is implied by the pre-thermomajorisation between the original vectors, but it does not imply it, i.e.,*

$$\hat{\boldsymbol{p}} \mathrel{{}_\epsilon\!\succ} \hat{\boldsymbol{q}} \qquad \impliedby \qquad \boldsymbol{p} \mathrel{{}_\epsilon\!\succ^\beta} \boldsymbol{q}, \tag{6.94a}$$

$$\hat{\boldsymbol{p}} \mathrel{{}_\epsilon\!\succ} \hat{\boldsymbol{q}} \qquad \not\Longrightarrow \qquad \boldsymbol{p} \mathrel{{}_\epsilon\!\succ^\beta} \boldsymbol{q}. \tag{6.94b}$$

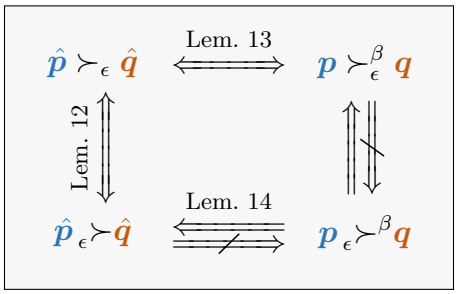


Figure 6.9: *Relations between approximate pre- and post-thermomajorisation relations.*

Proof. We first prove Eq. (6.94a). Assuming $\boldsymbol{p} \,{}_{\epsilon}\!\succ^{\beta} \boldsymbol{q}$ (and recalling that embedding is fidelity preserving) means that there exists $\hat{\tilde{\boldsymbol{p}}}$ majorising $\hat{\boldsymbol{q}}$ and satisfying $\delta(\hat{\boldsymbol{p}}, \hat{\tilde{\boldsymbol{p}}}) \leq \epsilon$. Then, by simply choosing $\tilde{\hat{\boldsymbol{p}}} = \hat{\tilde{\boldsymbol{p}}}$, we get $\tilde{\hat{\boldsymbol{p}}} \succ \hat{\boldsymbol{q}}$ and $\delta(\hat{\boldsymbol{p}}, \tilde{\hat{\boldsymbol{p}}}) \leq \epsilon$, so that $\hat{\boldsymbol{p}} \,{}_{\epsilon}\!\succ \hat{\boldsymbol{q}}$.

Now, in order to prove Eq. (6.94b) we will construct a particular counterexample. Consider the distributions

$$\boldsymbol{p} = [1, 0], \quad \boldsymbol{q} = [1/2, 1/2], \quad \boldsymbol{\gamma} = [3/4, 1/4]. \tag{6.95}$$

Their embedded versions are then given by

$$\hat{\boldsymbol{p}} = [1/3, 1/3, 1/3, 0], \quad \hat{\boldsymbol{q}} = [1/6, 1/6, 1/6, 1/2]. \tag{6.96}$$

We note that $\tilde{\hat{\boldsymbol{p}}} = [1/2, 1/4, 1/4, 0]$ majorises $\hat{\boldsymbol{q}}$ and that $\delta(\hat{\boldsymbol{p}}, \tilde{\hat{\boldsymbol{p}}}) = (3 - 2\sqrt{2})/6 =: \epsilon_0$. Therefore, $\hat{\boldsymbol{p}} \,{}_{\epsilon_0}\!\succ \hat{\boldsymbol{q}}$. On the other hand, for a general two-dimensional distribution, $\tilde{\boldsymbol{p}} = [\tilde{p}, 1 - \tilde{p}]$, we have $\delta(\boldsymbol{p}, \tilde{\boldsymbol{p}}) = 1 - \tilde{p}$, and the smallest $\tilde{p}$ for which $\tilde{\boldsymbol{p}} \succ^{\beta} \boldsymbol{q}$ is equal to $1/2$. This means that the optimal ϵ_1 for which $\tilde{\boldsymbol{p}} \,{}_{\epsilon_1}\!\succ^{\beta} \boldsymbol{q}$ holds is $\epsilon_1 = 1/2 > \epsilon_0$. We thus conclude that $\hat{\boldsymbol{p}} \,{}_{\epsilon}\!\succ \hat{\boldsymbol{q}}$ does not imply $\boldsymbol{p} \,{}_{\epsilon}\!\succ^{\beta} \boldsymbol{q}$. □

The results of this section are collectively presented in Fig. 6.9. We also would like to make a couple of remarks. First, using the equivalence between $\hat{\boldsymbol{p}} \succ_{\epsilon} \hat{\boldsymbol{q}}$ and $\boldsymbol{p} \succ_{\epsilon}^{\beta} \boldsymbol{q}$ one can, in principle, calculate the optimal fidelity (equivalently: minimal distance δ) between the final and target state under thermodynamic interconversion. More precisely, given initial distribution $\boldsymbol{p}$ and final $\boldsymbol{q}$, the optimal fidelity $F(\boldsymbol{q}, \tilde{\boldsymbol{q}})$ among $\tilde{\boldsymbol{q}}$ that are thermomajorised by $\boldsymbol{p}$ is equal to the optimal fidelity $F(\hat{\boldsymbol{q}}, \tilde{\hat{\boldsymbol{q}}})$ among $\tilde{\hat{\boldsymbol{q}}}$ that are majorised by $\hat{\boldsymbol{p}}$. This in turn, via Lemma 6.4, is equal to the optimal fidelity $F(\hat{\boldsymbol{p}}, \tilde{\hat{\boldsymbol{p}}})$ among $\tilde{\hat{\boldsymbol{p}}}$ that majorise $\hat{\boldsymbol{q}}$. But such an optimal state $\tilde{\hat{\boldsymbol{p}}}^{\star}$ is given by the explicit construction presented in Appendix 6.B, and thus we can directly calculate

$$\max_{\tilde{\boldsymbol{q}}:\ \boldsymbol{p} \succ^{\beta} \tilde{\boldsymbol{q}}} F(\boldsymbol{q}, \tilde{\boldsymbol{q}}) = F(\hat{\boldsymbol{p}}, \tilde{\hat{\boldsymbol{p}}}^{\star}). \tag{6.97}$$

In Section 6.C we discuss applying these very concepts to numerically compare our approximations of the optimal conversion rate to the true optimum for small system sizes. We give examples of such numerics in Figures 6.4 and 6.10.

Second, we want to point out that Lemmas 6.5 through 6.7 still hold if one applies them to the concept of approximate thermomajorisation based on total variation distance, i.e., if one replaces $\delta(\boldsymbol{p}, \boldsymbol{q})$ with $\frac{1}{2}||\boldsymbol{p} - \boldsymbol{q}||$ in Definition 6.3. The required modifications of the proofs are rather straightforward (with the exception of Lemma 6.5 which requires some fiddling), and we discuss them in Appendix 6.D.

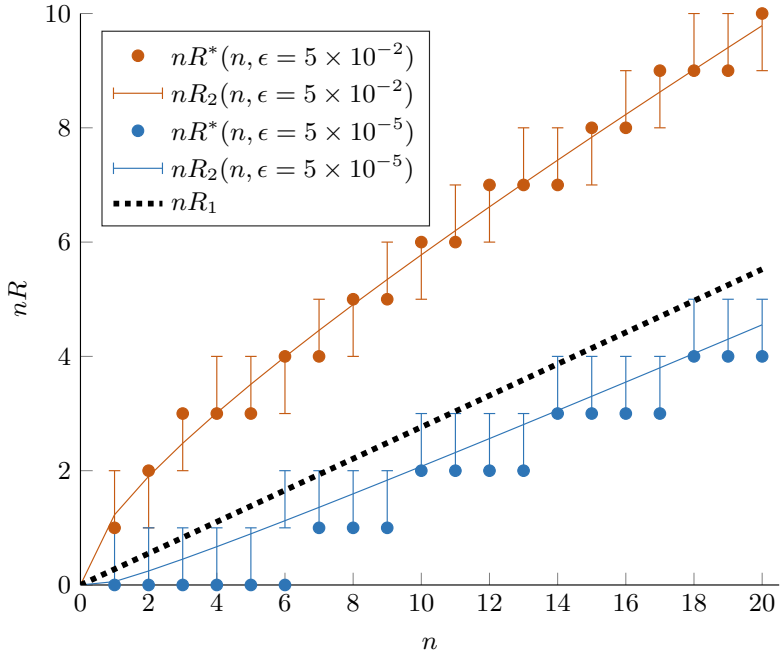


Figure 6.10: Comparison between the second-order approximation R_2 and exact thermal interconversion rates R^* for small system sizes, converting from $\rho = \frac{7}{10}|0\rangle\langle 0| + \frac{3}{10}|1\rangle\langle 1|$ to $\sigma = \frac{8}{10}|0\rangle\langle 0| + \frac{2}{10}|1\rangle\langle 1|$, with Hamiltonian $H = |1\rangle\langle 1|$ and access to a thermal bath at temperature $1/\beta = 3$, as in Figure 6.4. The circles indicate number of states produced (c.f. Section 6.C), and the lines those given by the second-order expansion from Eq. (1). As the exact number of states produced is always an integer, we have also indicated the rounding of the second-order approximation both up and down with error bars. The colours indicate the infidelity tolerance, with $\epsilon = 5 \times 10^{-2}$ for red and $\epsilon = 10^{-5}$ for blue. The dotted line indicates the number of produced states predicted by the asymptotic interconversion rate R_1.

6.6 Proofs of the main result

We will now present a proof of our main result, Theorem 6.3. We will do this by first showing a reduction to special case of bistochastic interconversion, which corresponds to infinite temperature. We recall that as we are considering energy-incoherent initial and target states, ρ and σ, we only need to consider their eigenvalues, denoted by $\boldsymbol{p}$ and $\boldsymbol{q}$, with the embedded versions of these given by $\hat{\boldsymbol{p}}$ and $\hat{\boldsymbol{q}}$. Also note that the embedded thermal state $\hat{\boldsymbol{\gamma}}$ simply corresponds to the uniform state $\boldsymbol{\eta}$. We can thus use the equivalence between approximate post-thermomajorisation and embedded majorisation, Lemma 6.6, to obtain:

$$R^*_\beta(n, \epsilon; \boldsymbol{p}, \boldsymbol{q}) = R^*_0(n, \epsilon; \hat{\boldsymbol{p}}, \hat{\boldsymbol{q}}), \tag{6.98a}$$

$$\epsilon^*_\beta(n, R; \boldsymbol{p}, \boldsymbol{q}) = \epsilon^*_0(n, R; \hat{\boldsymbol{p}}, \hat{\boldsymbol{q}}). \tag{6.98b}$$

Conveniently, the second-order expansion only depends on the states through the relative entropy and relative entropy variance, both of which are invariant under embedding, as noted earlier in Eq. (6.23). We can thus solve the problem for embedded distributions and $\beta = 0$, and in the final result exchange all $D(\hat{\boldsymbol{a}}||\boldsymbol{\eta})$ and $V(\hat{\boldsymbol{a}}||\boldsymbol{\eta})$ with $D(\boldsymbol{a}||\boldsymbol{\gamma})$ and $V(\boldsymbol{a}||\boldsymbol{\gamma})$, respectively. For the remainder of this section we will

henceforth drop the embedding hats and define $D(\boldsymbol{a}) := D(\boldsymbol{a}\|\boldsymbol{\eta})$. We also note that $V(\boldsymbol{a}\|\boldsymbol{\eta}) = V(\boldsymbol{a})$, where

$$V(\boldsymbol{a}) = \langle(-\log a_i - H(\boldsymbol{a}))^2\rangle_{\boldsymbol{a}} \tag{6.99}$$

is the entropy variance, and that $V(\boldsymbol{a}) = 0$ if and only if $\boldsymbol{a}$ is flat. Specifically, the irreversibility parameter given in Eq. (6.53) can be expressed as

$$\nu = \frac{V(\boldsymbol{p})/D(\boldsymbol{p})}{V(\boldsymbol{q})/D(\boldsymbol{q})}. \tag{6.100}$$

We will split the infinite-temperature proof into four parts, based on whether the relative entropy variances of the initial and target states are non-zero.

	$V(\boldsymbol{q}) = 0$	$V(\boldsymbol{q}) > 0$
$V(\boldsymbol{p}) = 0$	Flat-to-flat	Formation
$V(\boldsymbol{p}) > 0$	Distillation	Interconversion

We start with the case where both initial and target states have zero relative entropy variance. We refer to this as flat-to-flat interconversion, since both the initial and target state are flat. Recalling that states with embedded distributions being flat are proportional to the thermal state on their support, we note that this case contains the conversion between sharp energy states as a special case. We will then consider the cases of distillation and formation, in which either the target or initial states are flat respectively. These are so-named because they contain both the distillation of, and formation from, sharp energy states. Finally we will consider the general interconversion problem, in which neither state is flat. We refer to the non-flat distribution case as general, because it in fact implies the three other results by using the limiting behaviours of the Rayleigh-normal distributions given in Eq. (6.43).

6.6.1 Central limit theorem

Before we present our proofs, we first formulate the main mathematical tool needed for such a small deviation analysis: a central limit theorem. Specifically, we want to give tail bounds on i.i.d. product distributions. Considering the standard central limit theorem, one can derive the following tail bound.

Lemma 6.8. *For any distribution $\boldsymbol{a}$ such that $V(\boldsymbol{a}) > 0$,*

$$\lim_{n\to\infty} \sum_i \left\{ \left(\boldsymbol{a}^{\otimes n}\right)_i \middle| \left(\boldsymbol{a}^{\otimes n}\right)_i \geq 1/k_n(x) \right\} = \Phi\left(x\right), \tag{6.101}$$

where $k_n(x) := \left\lfloor \exp\left(H(\boldsymbol{a}^{\otimes n}) + x\sqrt{V(\boldsymbol{a}^{\otimes n})}\right)\right\rfloor$.

For completeness we provide the proof of the above known result in Appendix 6.E. We will also rely on an alternate form of central limit theorem.

Lemma 6.9 (Lemma 12 of Ref. [6.20]). *For any distribution $\boldsymbol{a}$ such that $V(\boldsymbol{a}) > 0$,*

$$\lim_{n\to\infty} \sum_{i=1}^{k_n(x)} \left(\boldsymbol{a}^{\otimes n}\right)_i^{\downarrow} = \Phi(x), \tag{6.102}$$

where $k_n(x) := \left\lfloor \exp\left(H(\boldsymbol{a}^{\otimes n}) + x\sqrt{V(\boldsymbol{a}^{\otimes n})}\right)\right\rfloor$.

We now want to convert the above result into the specific form of a bound we will use. For some rate $R(n)$, we define the *total* initial and target states as

$$\boldsymbol{P}^n := \boldsymbol{p}^{\otimes n} \otimes \boldsymbol{\eta}^{\otimes nR(n)}, \tag{6.103a}$$

$$\boldsymbol{Q}^n := \boldsymbol{q}^{\otimes nR(n)} \otimes \boldsymbol{\eta}^{\otimes n}. \tag{6.103b}$$

We note that generally R depends on n, but will henceforth omit this explicit dependence. We also introduce a quantity analogous to $k_n(x)$,

$$K_n(x) := \left\lfloor \exp\left(H(\boldsymbol{Q}^n) + x\sqrt{V(\boldsymbol{Q}^n)}\right)\right\rfloor, \tag{6.104}$$

that will be crucial in all our proofs. The central limit theorem for these distributions is given by the following result.

Lemma 6.10 (Central limit theorem for $\boldsymbol{P}^n$ and $\boldsymbol{Q}^n$). *If $V(\boldsymbol{q}) > 0$ and R is bounded away from zero, then $\boldsymbol{Q}^n$ has the tail bound*

$$\lim_{n\to\infty} \sum_{i=1}^{K_n(x)} Q_i^{n\downarrow} = \Phi(x). \tag{6.105}$$

Moreover, if we consider a rate of the form

$$R_\mu(n) = \frac{1}{D(\boldsymbol{q})}\left[D(\boldsymbol{p}) + \sqrt{\frac{D(\boldsymbol{p})}{D(\boldsymbol{q})}\frac{V(\boldsymbol{q})}{n}}\,\mu\right], \tag{6.106}$$

for some $\mu \in \mathbb{R}$, then $\boldsymbol{P}^n$ also has a corresponding tail bound

$$\lim_{n\to\infty} \sum_{i=1}^{K_n(x)} P_i^{n\downarrow} = \Phi_{\mu,\nu}(x), \tag{6.107}$$

with ν given by Eq. (6.100).

Proof. We start by noticing that Lemma 6.9 remains true if the product distribution is "smeared out". Specifically, for any flat state $\boldsymbol{f}$ we have

$$\lim_{n\to\infty} \sum_{i=1}^{k_n(x)} \left(\boldsymbol{a}^{\otimes n} \otimes \boldsymbol{f}\right)_i^{\downarrow} = \Phi(x), \tag{6.108}$$

where

$$k_n(x) := \left\lfloor \exp\left(H(\boldsymbol{a}^{\otimes n} \otimes \boldsymbol{f}) + x\sqrt{V(\boldsymbol{a}^{\otimes n} \otimes \boldsymbol{f})} \right) \right\rfloor . \tag{6.109}$$

Using this, if we make the substitutions $\boldsymbol{a}^{\otimes n} \leftarrow \boldsymbol{q}^{\otimes Rn}$ and $\boldsymbol{f} \leftarrow \boldsymbol{\eta}^{\otimes n}$, recalling that R is bounded away from zero, we arrive at

$$\lim_{n\to\infty} \sum_{i=1}^{K_n(x)} Q_i^{n\downarrow} = \Phi(x). \tag{6.110}$$

Applying the same argument to $\boldsymbol{P}^n$ gives

$$\lim_{n\to\infty} \sum_{i=1}^{K_n^P(y)} P_i^{n\downarrow} = \Phi\left(\frac{y}{\sqrt{V(\boldsymbol{p})}} \right), \tag{6.111}$$

where $K_n^P(y) := \lfloor \exp\left(H(\boldsymbol{P}^n) + y\sqrt{n}\right) \rfloor$. Note that we did not include the variance in the definition of K_n^P (as we did in K_n), because we have not assumed $V(\boldsymbol{p}) > 0$. Indeed, if we interpret $\Phi(y/V(\boldsymbol{p}))$ as cumulative of the zero-mean Dirac distribution for $V(\boldsymbol{p}) = 0$, then this also holds for $V(\boldsymbol{p}) = 0$.

We now want to express Eq. (6.111) as a summation up to $K_n(x)$ for some x. Noticing that $R_\mu = D(\boldsymbol{p})/D(\boldsymbol{q}) + o(1)$, our choice of rate R_μ can be rearranged to give

$$H(\boldsymbol{Q}^n) \simeq H(\boldsymbol{P}^n) - \sqrt{n\frac{D(\boldsymbol{p})}{D(\boldsymbol{q})}V(\boldsymbol{q})\mu}. \tag{6.112}$$

Therefore, $K_n^P(y) = K_n(x)$ is equivalent to

$$\frac{y}{\sqrt{V(\boldsymbol{p})}} = x\sqrt{\frac{V(\boldsymbol{Q}^n)}{V(\boldsymbol{P}^n)}} - \frac{\mu}{\sqrt{\nu}} = \frac{x-\mu}{\sqrt{\nu}} + o(1). \tag{6.113}$$

Finally, using the continuity of Φ gives the desired tail bound

$$\lim_{n\to\infty} \sum_{i=1}^{K_n(x)} P_i^{n\downarrow} = \lim_{n\to\infty} \sum_{i=1}^{K_n^P(y)} P_i^{n\downarrow} = \lim_{n\to\infty} \sum_{i=1}^{K_n^P\left(\frac{x-\mu}{\sqrt{\nu}}\right)} P_i^{n\downarrow} = \Phi\left(\frac{x-\mu}{\sqrt{\nu}}\right) = \Phi_{\mu,\nu}(x). \tag{6.114}$$

Recalling that ν is proportional to $V(\boldsymbol{p})$, we note that all of the above expressions are still well-defined if $V(\boldsymbol{p}) = 0$, where we understand $\Phi_{\mu,0}$ to be the cumulative of the Dirac distribution with mean μ. □

6.6.2 Bistochastic flat-to-flat

We start with the case of flat-to-flat conversion, where both the initial and target states are flat. In this boundary case, which was not considered in Ref. [6.20], we can provide an exact single-shot expression for the optimal rate.

Proposition 6.11 (Bistochastic flat-to-flat)**.** *For any initial state $\boldsymbol{p}$ and target state $\boldsymbol{q}$ such that $V(\boldsymbol{p}) = V(\boldsymbol{q}) = 0$, and infidelity $\epsilon \in [0,1)$, the optimal interconversion rate is given by*

$$R_0^*(n,\epsilon) = \frac{1}{n}\left\lfloor \frac{nD(\boldsymbol{p}) - \log(1-\epsilon)}{D(\boldsymbol{q})} \right\rfloor \simeq \frac{D(\boldsymbol{p})}{D(\boldsymbol{q})}. \tag{6.115}$$

Before proving this, we first consider the optimal majorising distribution in the case where all distributions involved are flat.

Lemma 6.12 (Single-shot flat-to-flat)**.** *If we let $\boldsymbol{a}$ and $\boldsymbol{b}$ be distributions such that $V(\boldsymbol{a}) = V(\boldsymbol{b}) = 0$, then*

$$\min\left\{\epsilon | \boldsymbol{a} \succ_\epsilon \boldsymbol{b}\right\} = \max\left\{1 - \frac{\exp H(\boldsymbol{b})}{\exp H(\boldsymbol{a})}, 0\right\}. \tag{6.116}$$

Proof. First, from Lemma 6.5 we know that

$$\min\left\{\epsilon | \boldsymbol{a} \succ_\epsilon \boldsymbol{b}\right\} = \min\left\{\epsilon | \boldsymbol{a} \,{}_\epsilon\!\succ \boldsymbol{b}\right\}. \tag{6.117}$$

Now, the right hand side of the above equation is minimised by a state $\tilde{\boldsymbol{a}}^\star$, whose explicit construction (found in Ref. [6.48]) we present in Appendix 6.B while proving Lemma 6.4. Using it one finds that

$$\tilde{\boldsymbol{a}}^{\star\downarrow} = \begin{cases} \boldsymbol{a}^\downarrow & \text{if } \boldsymbol{a} \succ \boldsymbol{b}, \\ \boldsymbol{b}^\downarrow & \text{if } \boldsymbol{a} \not\succ \boldsymbol{b}. \end{cases} \tag{6.118}$$

Since $\boldsymbol{a}$ and $\boldsymbol{b}$ are flat, in the first case we have

$$F(\tilde{\boldsymbol{a}}^{\star\downarrow}, \boldsymbol{b}^\downarrow) = \frac{\exp H(\boldsymbol{b})}{\exp H(\boldsymbol{a})}, \tag{6.119}$$

which leads to Eq. (6.116). □

We can now apply this to give an exact expression for the optimal rate.

Proof of Proposition 6.11. Applying Lemma 6.12 to $\boldsymbol{a} = \boldsymbol{P}^n$ and $\boldsymbol{b} = \boldsymbol{Q}^n$ we find that $\epsilon_0^*(n,R)$ vanishes for any $R \leq D(\boldsymbol{p})/D(\boldsymbol{q})$, and

$$\frac{1}{n}\log\left(1 - \epsilon_0^*(n,R)\right) = D(\boldsymbol{p}) - RD(\boldsymbol{q}) \tag{6.120}$$

for any $R \geq D(\boldsymbol{p})/D(\boldsymbol{q})$. Converting the expression for the optimal infidelity into an expression for the optimal rate, and recalling that $nR^*(n,\epsilon)$ must be an integer, gives Proposition 6.11 as required. □

6.6.3 Bistochastic distillation

We now consider distillation, in which the target state is flat.

Proposition 6.13 (Bistochastic distillation)**.** *For any initial state $\boldsymbol{p}$ and target state $\boldsymbol{q}$ such that $V(\boldsymbol{q}) = 0$, and infidelity $\epsilon \in (0,1)$, the optimal interconversion rate has the second-order expansion*

$$R_0^*(n,\epsilon) \simeq \frac{1}{D(\boldsymbol{q})}\left[D(\boldsymbol{p}) + \sqrt{\frac{V(\boldsymbol{p})}{n}}\Phi^{-1}(\epsilon)\right]. \tag{6.121}$$

Instead of utilising the techniques of Ref. [6.20], we will, similar to the flat-to-flat case, prove Proposition 6.13 by first considering a single-shot expression for the optimal error.

Lemma 6.14 (Single-shot distillation)**.** *Let* $\boldsymbol{a}$ *and* $\boldsymbol{b}$ *be distributions with* $V(\boldsymbol{b}) = 0$*. Then*

$$\min\left\{\epsilon | \boldsymbol{a} \succ_\epsilon \boldsymbol{b}\right\} = \sum_{i > \exp H(\boldsymbol{b})} a_i^{\downarrow}. \tag{6.122}$$

Proof. First, from Lemma 6.5 we know that

$$\min\left\{\epsilon | \boldsymbol{a} \succ_\epsilon \boldsymbol{b}\right\} = \min\left\{\epsilon | \boldsymbol{a} \,{}_\epsilon\!\succ \boldsymbol{b}\right\}. \tag{6.123}$$

Now, to find a state minimising the right hand side of the above equation, consider a distribution $\tilde{\boldsymbol{a}}$ such that $\tilde{\boldsymbol{a}} \succ \boldsymbol{b}$. Since $\boldsymbol{b}$ is flat, this is equivalent to the statement that $\tilde{\boldsymbol{a}}$ has a support which is no larger than that of $\boldsymbol{b}$. This condition is clearly necessary; it is sufficient as any distribution with d or fewer non-zero entries majorises the flat distribution over d entries. Using the Schwarz inequality one can then show that the distribution $\tilde{\boldsymbol{a}}$ which contains at most $\exp H(\boldsymbol{b})$ non-zero elements, and is closest to $\boldsymbol{a}$, is simply the truncated-and-rescaled distribution,

$$\tilde{a}_i^{\downarrow} := \frac{1}{\sum_{j=1}^{\exp(H(\boldsymbol{b}))} a_j^{\downarrow}} \begin{cases} a_i^{\downarrow} & \text{if } \; i \leq \exp H(\boldsymbol{b}), \\ 0 & \text{if } \; i > \exp H(\boldsymbol{b}). \end{cases} \tag{6.124}$$

It is a straightforward calculation to show that the infidelity of such a smoothed state is given by mass of the truncated tail

$$\delta(\boldsymbol{a}, \tilde{\boldsymbol{a}}) = \sum_{i > \exp H(\boldsymbol{b})} a_i^{\downarrow}. \tag{6.125}$$

□

We can now use Lemma 6.10 to bound this tail, giving a second-order expansion for asymptotic case.

Proof of Proposition 6.13. Applying Lemma 6.14 to $\boldsymbol{a} = \boldsymbol{P}^n$ and $\boldsymbol{b} = \boldsymbol{Q}^n$ we find that

$$\epsilon_0^*(n, R) = \sum_{i > \exp H(\boldsymbol{Q}^n)} P_i^{n\downarrow}. \tag{6.126}$$

Now consider a rate of the form

$$r_\mu(n) := \frac{1}{D(\boldsymbol{q})}\left[D(\boldsymbol{p}) + \sqrt{\frac{V(\boldsymbol{p})}{n}}\mu\right], \tag{6.127}$$

for some $\mu \in \mathbb{R}$. We then have

$$H(\boldsymbol{Q}^n) = H(\boldsymbol{P}^n) - \mu\sqrt{V(\boldsymbol{P}^n)}, \tag{6.128}$$

and so if we apply the first bound of Lemma 6.10 (with $\boldsymbol{P}^n$ in place of $\boldsymbol{Q}^n$), we arrive at

$$\lim_{n\to\infty} \epsilon_0^*(n, r_\mu) = 1 - \Phi(-\mu) = \Phi(\mu). \tag{6.129}$$

Reversing the relationship between infidelity and rate, this implies

$$R_0^*(n,\epsilon) \simeq \frac{1}{D(\boldsymbol{q})}\left[D(\boldsymbol{p}) + \sqrt{\frac{V(\boldsymbol{p})}{n}}\Phi^{-1}(\epsilon)\right], \tag{6.130}$$

as required. □

6.6.4 Bistochastic formation

For our final special case, we consider flat initial states.

Proposition 6.15 (Bistochastic formation)**.** *For any initial state $\boldsymbol{p}$ and target state $\boldsymbol{q}$ such that $V(\boldsymbol{p}) = 0$, and infidelity $\epsilon \in (0,1)$, the optimal interconversion rate has the second-order expansion*

$$R_0^*(n,\epsilon) \simeq \frac{1}{D(\boldsymbol{q})}\left[D(\boldsymbol{p}) + \sqrt{\frac{D(\boldsymbol{p})}{D(\boldsymbol{q})}\frac{V(\boldsymbol{q})}{n}}\Phi^{-1}(\epsilon)\right]. \tag{6.131}$$

This proof is more involved than distillation, and will involve some of the techniques used in the proof of the general interconversion problem, as first developed in Ref. [6.20]. As with distillation, we will attempt to bound the rate by bounding the optimal infidelity between the total initial state $\boldsymbol{P}^n$ and a state $\tilde{\boldsymbol{P}}^n$ that majorises the total target state $\boldsymbol{Q}^n$. We will thus fix our rate $R_\mu(n)$ to be given by Eq. (6.106) from Lemma 6.10, and look for bounds on infidelity between $\tilde{\boldsymbol{P}}^n$ and $\boldsymbol{P}^n$. More precisely, our proof is split into two parts: achieveability (upper bound on optimal error/lower bound on optimal rate) and optimality (lower bound on optimal error/upper bound on optimal rate).

Achieveability

Sketch of construction. The general idea here is to construct a distribution $\tilde{\boldsymbol{P}}^n$ which is close to the total initial state $\boldsymbol{P}^n$ *and* majorises total target state $\boldsymbol{Q}^n$. By the equivalence of pre- and post-majorisation, Lemma 6.5, this will prove that there exists a distribution $\tilde{\boldsymbol{Q}}^n$ that is majorised by the total initial state $\boldsymbol{P}^n$ and is close to the total target state $\boldsymbol{Q}^n$. We will start by defining two bins (sets) of indices, B and B'. We will then construct a *scaled distribution* $\boldsymbol{S}^n$ such that for indices belonging to B it has the same shape as $\boldsymbol{P}^n$ (i.e., it is flat), but has as much mass as $\boldsymbol{Q}^n$ has over the indices belonging to B'. The first property will guarantee that $\boldsymbol{S}^n$ lies close to $\boldsymbol{P}^n$, and the second that it lies close to a *majorising distribution* $\tilde{\boldsymbol{P}}^n \succ \boldsymbol{Q}^n$. We will then analyse $\delta(\boldsymbol{P}^n, \tilde{\boldsymbol{P}}^n)$, giving an upper bound on the optimal error.

Binning. For some small $\zeta > 0$, define two bins of indices

$$B := \{1, \ldots, K_n(\mu + \zeta/2)\}, \tag{6.132a}$$

$$B' := \{K_n(\mu + \zeta), \ldots, \infty\}, \tag{6.132b}$$

where ∞ is shorthand for the largest index, and μ fixes the value of rate $R_\mu(n)$. We will consider B as a bin on the indices of $\boldsymbol{P}^n$ and B' on those of $\boldsymbol{Q}^n$. We denote the complements of these bins as $\bar{B}$ and $\bar{B}'$, respectively.

Scaled distribution $\boldsymbol{S}^n$. For any $j \in B$ define

$$S_j^n := \frac{\sum_{k\in B'} Q_k^{n\downarrow}}{\sum_{k\in B} P_k^{n\downarrow}} \cdot P_j^{n\downarrow} = \frac{1}{|B|} \sum_{k\in B'} Q_k^{n\downarrow}, \tag{6.133}$$

and $S_l^n := 1 - \sum_{k\in B'} Q_k^{n\downarrow}$ for some arbitrary $l \notin B$ such that $\boldsymbol{S}^n$ is normalised.

By construction the mass of $\boldsymbol{S}^n$ on B is equal to that of $\boldsymbol{Q}^n$ on B', i.e.

$$\sum_{j\in B} S_j^n := \sum_{j\in B'} Q_k^{n\downarrow}. \tag{6.134}$$

We now want to show that this implies the existence of a nearby majorising distribution $\tilde{\boldsymbol{P}}^n \succ \boldsymbol{Q}^n$.

Majorising distribution $\tilde{\boldsymbol{P}}^n$. Instead of giving an explicit construction of $\tilde{\boldsymbol{P}}^n$, we instead present an existence proof. Specifically we will leverage the following lemma:

Lemma 6.16. *For non-negative vectors $\boldsymbol{a}$ and $\boldsymbol{b}$ such that $\sum_i a_i = \sum_i b_i$, there exists a vector $\tilde{\boldsymbol{a}}$ such that $\sum_k \tilde{a}_k = \sum_k a_k$, $\tilde{\boldsymbol{a}} \succ \boldsymbol{b}$ and $\|\tilde{\boldsymbol{a}} - \boldsymbol{a}\|_\infty \leq \|\boldsymbol{b}\|_\infty$.*

Proof. Consider a function over indices, $f : \mathbb{N} \to \mathbb{N}$, and the vector $\tilde{\boldsymbol{a}}$ given by the follow action of f on $\boldsymbol{b}$,

$$\tilde{a}_i := \sum_j \{b_j | f(j) = i\}. \tag{6.135}$$

Clearly such a mapping can only concentrate a distribution, and so $\tilde{\boldsymbol{a}} \succ \boldsymbol{b}$. Now, among $\tilde{\boldsymbol{a}}$ of the above form we choose that which is closest to $\boldsymbol{a}$ in l_∞-norm. Let i be an index at which the l_∞-norm of $\tilde{\boldsymbol{a}} - \boldsymbol{a}$, denoted by Δ, is achieved,

$$i \in \arg\max_j |\tilde{a}_j - a_j|. \tag{6.136}$$

We are going to assume that $\Delta > \|\boldsymbol{b}\|_\infty$ and show that this would imply that $\tilde{\boldsymbol{a}}$ cannot be optimal, proving $\|\tilde{\boldsymbol{a}} - \boldsymbol{a}\|_\infty \leq \|\boldsymbol{b}\|_\infty$ by way of contradiction.

There are two cases to consider: $\tilde{a}_i > a_i$ and $\tilde{a}_i < a_i$. We start with $\tilde{a}_i > a_i$. As $\tilde{a}_i > 0$, there must exist some α such that $f(\alpha) = i$ and $b_\alpha > 0$. As $\boldsymbol{a} \neq \tilde{\boldsymbol{a}}$ and $\sum_k a_k = \sum_k \tilde{a}_k$, there must exist a k such that $\tilde{a}_k < a_k$. Consider changing the map to from $f(\alpha) = i$ to $f(\alpha) = k \neq i$. This has the effect of lowering $\tilde{a}_i$ by b_α and raising $\tilde{a}_k$ by the same amount. Given that $b_\alpha < \Delta$ by assumption, this means that $\tilde{a}_i - a_i$ changes from Δ to $(0, \Delta)$, and $\tilde{a}_k - a_k$ changes from $[-\Delta, 0)$ to $(-\Delta, \Delta)$. As such, $|\tilde{a}_i - a_i|$ and $|\tilde{a}_j - a_j|$ are now both strictly smaller than Δ. Since all other entries of $\tilde{\boldsymbol{a}}$ are unchanged, we have reduced the number of indices j such that $|\tilde{a}_j - a_j| = \Delta$ by at least one. Similarly for the case $\tilde{a}_i < a_i$ one can make an analogous argument by changing $f(\alpha) = k \neq i$ to $f(\alpha) = k$ for some k such that $\tilde{a}_k > a_k$. Iterating this we can keep decreasing the number of indices at which the norm is achieved, eventually giving us $|\tilde{a}_j - a_j| < \Delta$ for all j, i.e. $|\tilde{\boldsymbol{a}} - \boldsymbol{a}| < \Delta$. This shows that the original choice of $\tilde{\boldsymbol{a}}$ was not optimal as assumed, proving $\Delta \leq \|\boldsymbol{b}\|_\infty$ by contradiction. □

With the use of the above lemma we can now get the desired majorising distribution $\tilde{\boldsymbol{P}}^n \succ \boldsymbol{Q}^n$.

Lemma 6.17. *There exists a distribution $\tilde{\boldsymbol{P}}^n$ such that $\tilde{\boldsymbol{P}}^n \succ \boldsymbol{Q}^n$ and*

$$\left|\tilde{P}_j^n - S_j^n\right| \leq 1/K_n(\mu+\zeta) \tag{6.137}$$

for all $j \in B$.

Proof. The idea here is to apply Lemma 6.16 to the restriction of each distribution to its corresponding bin. Specifically if we take $\boldsymbol{a} := \boldsymbol{S}^n|_B$ and $\boldsymbol{b} := \boldsymbol{Q}^{n\downarrow}\big|_{B'}$, then Lemma 6.16 gives us a vector $\tilde{\boldsymbol{a}}$ such that $\tilde{\boldsymbol{a}} \succ \boldsymbol{Q}^{n\downarrow}\big|_{B'}$ and

$$\|\tilde{\boldsymbol{a}} - \boldsymbol{a}\|_\infty \leq \|\boldsymbol{b}\|_\infty = \left\|\boldsymbol{Q}^{n\downarrow}\Big|_{B'}\right\|_\infty = \max_{j\in B'} Q_j^{n\downarrow} = Q^{n\downarrow}_{K_n(\mu+\zeta)} \leq 1/K_n(\mu+\zeta), \tag{6.138}$$

where the final inequality follows from normalisation of $\boldsymbol{Q}^n$. We now define our majorising distribution within bin B as $\tilde{\boldsymbol{P}}^n\big|_B := \tilde{\boldsymbol{a}}$, so that $|\tilde{P}_j^n - S_j^n| \leq 1/K_n(\mu+\zeta)$ for any $j \in B$ as desired; and, once again, in order to normalise the distribution we also define $\tilde{P}_l^n = 1 - \sum_i \tilde{a}_i$ for some arbitrary $l \notin B$.

The fact that $\tilde{\boldsymbol{a}}$ majorises the restriction of $\boldsymbol{Q}^{n\downarrow}$ to B', together with the sharpness of $\tilde{\boldsymbol{P}}^n$ outside of B, gives

$$\tilde{\boldsymbol{P}}^n\Big|_B \succ \boldsymbol{Q}^{n\downarrow}\Big|_{B'} \quad \text{and} \quad \tilde{\boldsymbol{P}}^n\Big|_{\bar{B}} \succ \boldsymbol{Q}^{n\downarrow}\Big|_{\bar{B}'}. \tag{6.139}$$

Next, we note that majorisation spreads over direct sum, i.e. $\boldsymbol{\alpha}_1 \succ \boldsymbol{\beta}_1$ and $\boldsymbol{\alpha}_2 \succ \boldsymbol{\beta}_2$ implies $\boldsymbol{\alpha}_1 \oplus \boldsymbol{\alpha}_2 \succ \boldsymbol{\beta}_1 \oplus \boldsymbol{\beta}_2$. This can be seen by using Theorem 6.2 (i.e., the equivalence of majorisation relation between two distribution with the existence of a bistochastic map between them), and noticing that bistochasticity is preserved under direct sum. Applying this to $\tilde{\boldsymbol{P}}^n$ gives the desired majorisation property: $\tilde{\boldsymbol{P}}^n \succ \tilde{\boldsymbol{P}}^n|_B \oplus \tilde{\boldsymbol{P}}^n|_{\overline{B}} \succ \boldsymbol{Q}^{n\downarrow}|_{B'} \oplus \boldsymbol{Q}^{n\downarrow}|_{\overline{B'}} \succ \boldsymbol{Q}^n$. □

Infidelity. By now we have proven the existence of a majorising distribution $\tilde{\boldsymbol{P}}^n \succ \boldsymbol{Q}^n$ and bounded its distance from the scaled distribution $\boldsymbol{S}^n$ on the restriction to B. The final step involves bounding the infidelity $\delta(\tilde{\boldsymbol{P}}^n, \boldsymbol{P}^n)$. To achieve this we will first show that the closeness of $\boldsymbol{S}^n$ and $\tilde{\boldsymbol{P}}^n$ on B, as given in Lemma 6.17, allows us to bound $F(\tilde{\boldsymbol{P}}^n, \boldsymbol{P}^n)$ in terms of $\boldsymbol{S}^n$.

Lemma 6.18. *Asymptotically, the fidelity between the majorising distribution $\tilde{\boldsymbol{P}}^n$ and the total initial distribution $\boldsymbol{P}^n$ can be bounded as follows*

$$\liminf_{n\to\infty} F\left(\tilde{\boldsymbol{P}}^n, \boldsymbol{P}^n\right) \geq \liminf_{n\to\infty} \left(\sum_{j\in B} \sqrt{S_j^n P_j^n}\right)^2. \tag{6.140}$$

Proof. First, we apply Lemma 6.17 and the fact that $\sqrt{x-y} \geq \sqrt{x} - \sqrt{y}$ for all $x \geq y \geq 0$ to break the fidelity into the desired expression and an error term

$$\begin{aligned}\sqrt{F(\tilde{\boldsymbol{P}}^n, \boldsymbol{P}^{n\downarrow})} &\geq \sum_{j\in B} \sqrt{\tilde{P}_j^n P_j^{n\downarrow}} \geq \sum_{j\in B} \sqrt{\max\{S_j^n - 1/K_n(\mu+\zeta), 0\} P_j^{n\downarrow}} \\ &\geq \sum_{j\in B} \left[\sqrt{S_j^n P_j^{n\downarrow}} - \sqrt{P_j^{n\downarrow}/K_n(\mu+\zeta)}\right].\end{aligned} \tag{6.141}$$

We can express the second error term as

$$\sum_{j\in B}\sqrt{\frac{P_j^{n\downarrow}}{K_n(\mu+\zeta)}}=\sqrt{\frac{|B|}{K_n(\mu+\zeta)}}=\sqrt{\frac{K_n(\mu+\zeta/2)}{K_n(\mu+\zeta)}}. \tag{6.142}$$

Given that $\zeta>0$ is a constant and $V(\boldsymbol{q})>0$, we have that $K_n(\mu+\zeta/2)/K_n(\mu+\zeta)$ is decaying exponentially as $n\to\infty$. Taking the limit inferior of Eq. (6.141) therefore gives the required bound. □

Using the above result on fidelity, we can now prove achieveability.

Proof of Proposition 6.15 (achieveability). Substituting the definition of $\boldsymbol{S}^n$ into Lemma 6.18 gives

$$\liminf_{n\to\infty} F\left(\tilde{\boldsymbol{P}}^n,\boldsymbol{P}^n\right)\geq\liminf_{n\to\infty}\sum_{i\in B'}Q_i^{n\downarrow}\sum_{j\in B}P_j^{n\downarrow}. \tag{6.143}$$

By applying Lemma 6.10 we then obtain

$$\sum_{i\in B}P_i^{n\downarrow}\xrightarrow{n\to\infty}\Phi_{\mu,0}(\mu+\zeta/2)=1, \tag{6.144a}$$

$$\sum_{i\in B'}Q_i^{n\downarrow}\xrightarrow{n\to\infty}1-\Phi(\mu+\zeta), \tag{6.144b}$$

and therefore

$$\limsup_{n\to\infty}\delta\left(\tilde{\boldsymbol{P}}^n,\boldsymbol{P}^n\right)\leq\Phi(\mu+\zeta). \tag{6.145}$$

Due to the equivalence between pre- and post-majorisation, Lemma 6.5, the above means that there exists a distribution $\tilde{\boldsymbol{Q}}^n$ that is majorised by the total initial state $\boldsymbol{P}^n$ and such that

$$\limsup_{n\to\infty}\delta\left(\tilde{\boldsymbol{Q}}^n,\boldsymbol{Q}^n\right)\leq\Phi(\mu+\zeta). \tag{6.146}$$

As this is true for any $\zeta>0$ we can take $\zeta\searrow 0$ and conclude that the optimal infidelity is upper bounded

$$\limsup_{n\to\infty}\epsilon_0^*(n,R_\mu)\leq\Phi(\mu), \tag{6.147}$$

which implies a corresponding lower bound on the optimal rate

$$R_0^*(n,\epsilon)\gtrsim\frac{1}{D(\boldsymbol{q})}\left[D(\boldsymbol{p})+\sqrt{\frac{D(\boldsymbol{p})}{D(\boldsymbol{q})}\frac{V(\boldsymbol{q})}{n}}\Phi^{-1}(\epsilon)\right], \tag{6.148}$$

as required. □

Optimality

We now turn our attention to a corresponding second-order upper bound on the optimal rate. We will make use of the following lemma from Ref. [6.20], which upper bounds the fidelity between a flat state and a majorising distribution.

Lemma 6.19 (Lemma 6 of Ref. [6.20]). *Let $\boldsymbol{a}$, $\tilde{\boldsymbol{a}}$ and $\boldsymbol{b}$ be any distributions such that $V(\boldsymbol{a}) = 0$ and $\tilde{\boldsymbol{a}} \succ \boldsymbol{b}$. Also, let $M \leq \exp H(\boldsymbol{a})$. Then*

$$\sqrt{F(\tilde{\boldsymbol{a}}, \boldsymbol{a})} \leq \sqrt{\sum_{i \leq N} a_i^{\downarrow} \sum_{j \leq N} b_j^{\downarrow}} + \sqrt{\sum_{i > N} a_i^{\downarrow} \sum_{j > N} b_j^{\downarrow}}, \tag{6.149}$$

where $N := |\{i | b_i \geq 1/M\}|$.

We can now use the above lemma to obtain an optimality bound which matches that given for achieveability.

Proof of Proposition 6.15 (optimality). Consider any distribution $\tilde{\boldsymbol{P}}^n \succ \boldsymbol{Q}^n$. Now choose $\tilde{\boldsymbol{a}} = \tilde{\boldsymbol{P}}^n$, $\boldsymbol{a} = \boldsymbol{P}^n$ and $\boldsymbol{b} = \boldsymbol{Q}^n$. Also, notice that $M := K_n(\mu - \zeta)$ satisfies $M \leq \exp H(\boldsymbol{a})$. Hence, we can apply Lemma 6.19 to upper bound the fidelity,

$$\sqrt{F(\tilde{\boldsymbol{P}}^n, \boldsymbol{P}^n)} \leq \sqrt{\sum_{i \leq N} P_i^{n\downarrow} \sum_{j \leq N} Q_j^{n\downarrow}} + \sqrt{\sum_{i > N} P_i^{n\downarrow} \sum_{j > N} Q_j^{n\downarrow}}, \tag{6.150}$$

where $N := |\{i | Q_i^n \geq 1/K_n(\mu - \zeta)\}|$. By the standard central limit theorem Lemma 6.8, we have

$$\sum_{i=1}^{N} Q_i^{n\downarrow} = \sum_i \left\{ Q_i^n \middle| Q_i^n \geq \frac{1}{K_n(\mu - \zeta)} \right\} \to \Phi(\mu - \zeta). \tag{6.151}$$

The normalisation of $\boldsymbol{Q}^n$ gives us that $N \leq K_n(\mu - \zeta)$, and so we can apply Lemma 6.10 to obtain

$$\sum_{i=1}^{N} P_i^{n\downarrow} \leq \sum_{i=1}^{K_n(\mu-\zeta)} P_i^{n\downarrow} \to \Phi_{\mu,0}(\mu - \zeta) = 0. \tag{6.152}$$

Applying these limits to Eq. (6.150) yields

$$\liminf_{n\to\infty} \delta\left(\tilde{\boldsymbol{P}}^n, \boldsymbol{P}^n\right) \geq \Phi(\mu - \zeta) \tag{6.153}$$

for any $\tilde{\boldsymbol{P}}^n \succ \boldsymbol{Q}^n$. Due to the equivalence between pre- and post-majorisation (c.f. Lemma 6.5) the above means that for any distribution $\tilde{\boldsymbol{Q}}^n$ that is majorised by the total initial state $\boldsymbol{P}^n$ we have

$$\liminf_{n\to\infty} \delta\left(\tilde{\boldsymbol{Q}}^n, \boldsymbol{Q}^n\right) \geq \Phi(\mu - \zeta) \tag{6.154}$$

Taking $\zeta \searrow 0$ this gives a lower bound on the optimal infidelity

$$\liminf_{n\to\infty} \epsilon_0^*(n, R_\mu) \geq \Phi(\mu), \tag{6.155}$$

which implies a corresponding upper bound on the optimal rate

$$R_0^*(n, \epsilon) \lesssim \frac{1}{D(\boldsymbol{q})} \left[D(\boldsymbol{p}) + \sqrt{\frac{D(\boldsymbol{p})}{D(\boldsymbol{q})} \frac{V(\boldsymbol{q})}{n}} \Phi^{-1}(\epsilon) \right]. \tag{6.156}$$

□

6.6.5 Bistochastic interconversion

Finally we turn to the general case in which neither relative entropy variance is vanishing.

Proposition 6.20 (Bistochastic interconversion)**.** *For any initial state $\boldsymbol{p}$ and target state $\boldsymbol{q}$ such that $V(\boldsymbol{p}), V(\boldsymbol{q}) > 0$, and infidelity $\epsilon \in (0,1)$, the optimal interconversion rate has the second-order expansion*

$$R_0^*(n,\epsilon) \simeq \frac{1}{D(\boldsymbol{q})}\left[D(\boldsymbol{p}) + \sqrt{\frac{V(\boldsymbol{p})}{n}}Z_{1/\nu}^{-1}(\epsilon)\right], \tag{6.157}$$

where ν is given in 6.100.

The proof is similar to that of formation, so it will also utilise many of the ideas inspired by Ref. [6.20]. The main complication is that for the general interconversion problem the binning of indices is more elaborate: we now have two *sets* of bins instead of two *individual* bins, and we need to introduce a function A which controls the relative placement of these bins. Once again we will break the proof into both achieveability and optimality bounds.

Achieveability

Sketch of proof. As with formation, the idea here will be to give an explicit construction of $\tilde{\boldsymbol{P}}^n \succ \boldsymbol{Q}^n$ which is close to $\boldsymbol{P}^n$. Again, due to the equivalence of pre- and post-majorisation, this will prove that there exists a distribution $\tilde{\boldsymbol{Q}}^n$ that is majorised by the total initial state $\boldsymbol{P}^n$ and is close to the total target state $\boldsymbol{Q}^n$. We will start by introducing two sets of bins for each distribution. Using these bins, we will once again construct a *scaled distribution* $\boldsymbol{S}^n$ which reflects the fine-grained features of $\boldsymbol{P}^n$ (same shape within corresponding bins) and coarse-grained features of $\boldsymbol{Q}^n$ (same mass within corresponding bins). We will then show that $\boldsymbol{S}^n$ necessarily lies close to a *majorising distribution* $\tilde{\boldsymbol{P}}^n$. Finally we will analyse the infidelity of this distribution with respect to the total initial distribution $\boldsymbol{P}^n$ and, by taking the appropriate limits of the parameters in our construction, prove the desired achieveabilty bound of Proposition 6.20.

In Section 6.6.4 our construction was parameterised by a single slack parameter $\zeta > 0$. Here, we will have three parameters: $\lambda > 0$, $I \in \mathbb{N}$, and a monotone continuously differentiable function $1 \geq A \geq \Phi$ pointwise. The parameter λ will control the width of our bins, I the number of bins, and A the relative placements of the two sets of bins.

Binning. For $-I \leq i < I$ we define our two sets of bins as

$$B_i := \{K_n(x_i), \ldots, K_n(x_{i+1}) - 1\}, \tag{6.158a}$$
$$B'_i := \{K_n(y_i), \ldots, K_n(y_{i+1}) - 1\}, \tag{6.158b}$$

where the two sequences are defined by

$$x_i := \lambda\frac{i-1}{I} \quad \text{and} \quad y_i := \Phi^{-1}\left(A\left(\lambda\frac{i+1}{I}\right)\right) \tag{6.159}$$

for $-I \leq i \leq I$. We will consider B_i as bins on the indices of $\boldsymbol{P}^n$ and B'_i on those of $\boldsymbol{Q}^n$. We note that $A \geq \Phi$ implies $y_i \geq x_{i+1} + \lambda/I$, resulting in B_i being gapped away from B'_i, i.e., all indices belonging to B'_i are much larger than those belonging to B_i. This choice plays a role analogous to that of the slack parameter ζ for the bins in Section 6.6.4. For convenience we also define

$$B := \bigcup_{i=-I}^{I-1} B_i \quad \text{and} \quad B' := \bigcup_{i=-I}^{I-1} B'_i \tag{6.160}$$

to be the union of bins, and $\overline{B}$ and $\overline{B'}$ to be the corresponding complements.

Scaled distribution $\boldsymbol{S}^n$. As in Section 6.6.4, we now define $\boldsymbol{S}^n$ in each bin B_i to have the shape of $\boldsymbol{P}^n$ within B_i, but the mass of $\boldsymbol{Q}^n$ within B'_i. As the bins B_i are all disjoint, for any $j \in B$ there exists a unique $-I \leq i < I$ such that $j \in B_i$. For such indices we define $\boldsymbol{S}^n$ as

$$S^n_j := \frac{\sum_{k \in B'_i} Q^{n\downarrow}_k}{\sum_{k \in B_i} P^{n\downarrow}_k} \cdot P^{n\downarrow}_j. \tag{6.161}$$

We normalise $\boldsymbol{S}^n$ by taking $S^n_l := 1 - \sum_{j \in B} S^n_j$ for some arbitrary $l \notin B$.

Majorising distribution $\tilde{\boldsymbol{P}}^n$. We now want to prove a result analogous to Lemma 6.17: the existence of a distribution that simultaneously majorises the total target distribution $\boldsymbol{Q}^n$ and is close to $\boldsymbol{S}^n$ (within each bin).

Lemma 6.21 (Existence of a majorising distribution)**.** *There exists a distribution $\tilde{\boldsymbol{P}}^n$ such that $\tilde{\boldsymbol{P}}^n \succ \boldsymbol{Q}^n$ and*

$$\left| \tilde{P}^n_j - S^n_j \right| \leq 1/K_n(y_i) \tag{6.162}$$

for all $j \in B_i$ and $-I \leq i < I$.

Proof. The proof is analogous to that of Lemma 6.17, with an application of Lemma 6.16 for each pair $\boldsymbol{S}^n|_{B_i}$ and $\boldsymbol{Q}^n|_{B'_i}$, with $-I \leq i < I$. This gives us $\tilde{\boldsymbol{P}}^n$ such that for all $-I \leq i < I$ and $j \in B_i$ it is close to $\boldsymbol{S}^n$

$$\left| \tilde{P}^n_j - S^n_j \right| \leq \max_{j \in B'_i} Q^{n\downarrow}_j \leq 1/K_n(y_i), \tag{6.163}$$

and possesses the majorisation properties

$$\tilde{\boldsymbol{P}}^n|_{B_i} \succ \boldsymbol{Q}^{n\downarrow}|_{B'_i} \quad \text{and} \quad \tilde{\boldsymbol{P}}^n|_{\bar{B}} \succ \boldsymbol{Q}^{n\downarrow}|_{\bar{B}'}. \tag{6.164}$$

Splitting the majorisation across the direct sum, as explained in the proof of Lemma 6.17, gives us the desired overall majorisation

$$\begin{aligned} \tilde{\boldsymbol{P}}^n &\succ \bigoplus_{i=-I}^{I-1} \tilde{\boldsymbol{P}}^n|_{B_i} \oplus \tilde{\boldsymbol{P}}^n|_{\overline{B}} \\ &\succ \bigoplus_{i=-I}^{I-1} \boldsymbol{Q}^{n\downarrow}|_{B'_i} \oplus \boldsymbol{Q}^{n\downarrow}|_{\overline{B'}} \succ \boldsymbol{Q}^n. \end{aligned} \tag{6.165}$$

□

Infidelity. The next step involves bounding the fidelity between the total initial state $\boldsymbol{P}^n$ and majorising distribution $\tilde{\boldsymbol{P}}^n$ given by the above construction. We will start by bounding the fidelity for a fixed set of parameters A, λ and I.

Lemma 6.22. *For any monotone continuously differentiable function $1 \geq A \geq \Phi$, $\lambda \geq 0$, and $I \in \mathbb{N}$ there exists a sequence of distributions $\tilde{\boldsymbol{P}}^n \succ \boldsymbol{Q}^n$ such that*

$$\liminf_{n\to\infty} F\left(\tilde{\boldsymbol{P}}^n, \boldsymbol{P}^{n\downarrow}\right) \geq \left(\int_{-\lambda}^{\lambda} \sqrt{A'\left(x+\frac{\lambda}{I}\right)\Phi'_{\mu,\nu}\left(x-\frac{\lambda}{I}\right)}\,\mathrm{d}x\right)^2, \tag{6.166}$$

with the prime superscript in A' and $\Phi'_{\mu,\nu}$ denoting a derivative.

Proof. The first part of the proof is analogous to the proof of Lemma 6.18. More precisely using Lemma 6.21 (in place of Lemma 6.17) and employing the fact that B_i is gapped away from B_i', we can apply the argument presented there to obtain

$$\liminf_{n\to\infty} F\left(\tilde{\boldsymbol{P}}^n, \boldsymbol{P}^n\right) \geq \liminf_{n\to\infty} \left(\sum_{j\in B} \sqrt{S_j^n P_j^n}\right)^2. \tag{6.167}$$

Inserting the definition of the scaled distribution $\boldsymbol{S}^n$ yields

$$\begin{aligned}
\liminf_{n\to\infty} \sqrt{F\left(\tilde{\boldsymbol{P}}^n, \boldsymbol{P}^{n\downarrow}\right)} &\geq \liminf_{n\to\infty} \sum_{i=-I}^{I-1} \sum_{j\in B_i} \sqrt{S_j^n P_j^{n\downarrow}} \\
&= \liminf_{n\to\infty} \sum_{i=-I}^{I-1} \sum_{j\in B_i} \sqrt{\frac{\sum_{k\in B_i'} Q_k^{n\downarrow}}{\sum_{k\in B_i} P_k^{n\downarrow}} \cdot P_j^{n\downarrow}} \sqrt{P_j^{n\downarrow}} \\
&= \liminf_{n\to\infty} \sum_{i=-I}^{I-1} \sqrt{\sum_{j\in B_i'} Q_j^{n\downarrow} \sum_{k\in B_i} P_k^{n\downarrow}}.
\end{aligned} \tag{6.168}$$

Recalling that $\Phi(y_i) = A(x_{i+2})$ and applying Lemma 6.10 gives

$$\lim_{n\to\infty} \sum_{j\in B_i} P_j^{n\downarrow} = \Phi_{\mu,\nu}(x_{i+1}) - \Phi_{\mu,\nu}(x_i), \tag{6.169a}$$

$$\lim_{n\to\infty} \sum_{j\in B_i'} Q_j^{n\downarrow} = A(x_{i+3}) - A(x_{i+2}). \tag{6.169b}$$

Substituting these into our lower bound on fidelity yields

$$\begin{aligned}
\liminf_{n\to\infty} \sqrt{F\left(\tilde{\boldsymbol{P}}^n, \boldsymbol{P}^{n\downarrow}\right)} &\geq \sum_{i=-I}^{I-1} \sqrt{A(x_{i+3}) - A(x_{i+2})} \times \sqrt{\Phi_{\mu,\nu}(x_{i+1}) - \Phi_{\mu,\nu}(x_i)} \\
&\geq \sum_{i=-I}^{I-1} \sqrt{A\left(\frac{\lambda(i+2)}{I}\right) - A\left(\frac{\lambda(i+1)}{I}\right)} \times \sqrt{\Phi_{\mu,\nu}\left(\frac{\lambda i}{I}\right) - \Phi_{\mu,\nu}\left(\frac{\lambda(i-1)}{I}\right)}.
\end{aligned} \tag{6.170}$$

Using the differentiability of A and $\Phi_{\mu,\nu}$ we can express these finite differences as integrals

$$\liminf_{n\to\infty} \sqrt{F\left(\tilde{\boldsymbol{P}}^n, \boldsymbol{P}^{n\downarrow}\right)} \geq \sum_{i=-I}^{I-1} \sqrt{\int_{\lambda\frac{i}{I}}^{\lambda\frac{i+1}{I}} A'\left(x+\frac{\lambda}{I}\right) \mathrm{d}x} \times \sqrt{\int_{\lambda\frac{i}{I}}^{\lambda\frac{i+1}{I}} \Phi'_{\mu,\nu}\left(x-\frac{\lambda}{I}\right) \mathrm{d}x}. \tag{6.171}$$

Finally, we apply the Schwarz inequality to arrive at the desired bound

$$\begin{aligned}\liminf_{n\to\infty} \sqrt{F\left(\tilde{\boldsymbol{P}}^n, \boldsymbol{P}^{n\downarrow}\right)} &\geq \sum_{i=-I}^{I-1} \int_{\lambda\frac{i}{I}}^{\lambda\frac{i+1}{I}} \sqrt{A'\left(x+\frac{\lambda}{I}\right)\Phi'_{\mu,\nu}\left(x-\frac{\lambda}{I}\right)} \mathrm{d}x \\ &= \int_{-\lambda}^{\lambda} \sqrt{A'\left(x+\frac{\lambda}{I}\right)\Phi'_{\mu,\nu}\left(x-\frac{\lambda}{I}\right)} \mathrm{d}x. \end{aligned} \tag{6.172}$$

□

Now, by taking the appropriate limits of our parameters A, λ and I, we will get the desired achieveability bound on the optimal infidelity, and therefore also on the optimal rate.

Proof of Proposition 6.20 (achieveability). By Lemma 6.22 we know that there exists a family of distributions $\tilde{\boldsymbol{P}}^n$ majorising $\boldsymbol{Q}^n$ and such that $\liminf_{n\to\infty} F\left(\tilde{\boldsymbol{P}}^n, \boldsymbol{P}^{n\downarrow}\right)$ is lower-bounded by

$$\left(\int_{-\lambda}^{\lambda} \sqrt{A'\left(x+\frac{\lambda}{I}\right)\Phi'_{\mu,\nu}\left(x-\frac{\lambda}{I}\right)} \mathrm{d}x\right)^2. \tag{6.173}$$

Due to the equivalence between pre- and post-majorisation, Lemma 6.5, this means that there exists a family of distributions $\tilde{\boldsymbol{Q}}^n$ that is majorised by the total initial state $\boldsymbol{P}^n$ and such that their fidelity with the total target state $\boldsymbol{Q}^n$ is also lower bounded by the above expression, which implies a lower bound on the asymptotic ideal fidelity

$$\begin{aligned}\liminf_{n\to\infty} \sqrt{1-\epsilon_0^*(n, R_\mu)} &\geq \liminf_{n\to\infty} \sqrt{F\left(\tilde{\boldsymbol{Q}}^n, \boldsymbol{Q}^{n\downarrow}\right)} \\ &\geq \int_{-\lambda}^{\lambda} \sqrt{A'\left(x+\frac{\lambda}{I}\right)\Phi'_{\mu,\nu}\left(x-\frac{\lambda}{I}\right)} \mathrm{d}x. \end{aligned} \tag{6.174}$$

As the left hand side is independent of I, λ and A, we can now take the desired limits. Note that the order of limits will be important: first we will take $I \to \infty$, then $\lambda \to \infty$, followed by a supremum over A.

Firstly, we take the limit inferior $I \to \infty$. As a consequence of the fact that λ is still finite, together with the continuous differentiability of A and $\Phi_{\mu,\nu}$, we have the point-wise limit

$$\sqrt{A'\left(x+\frac{\lambda}{I}\right)}\sqrt{\Phi'_{\mu,\nu}\left(x-\frac{\lambda}{I}\right)} \quad \xrightarrow{I\to\infty} \quad \sqrt{A'(x)}\sqrt{\Phi'_{\mu,\nu}(x)}. \tag{6.175}$$

Using the compactness of $[-2\lambda, 2\lambda]$ we can apply the dominated convergence theorem to move this limit inside the integral, which gives

$$\begin{aligned}\liminf_{n\to\infty} \sqrt{1-\epsilon_0^*(n,R_\mu)} &\geq \liminf_{I\to\infty} \int_{-\lambda}^{\lambda} \sqrt{A'\left(x+\frac{\lambda}{I}\right)\Phi'_{\mu,\nu}\left(x-\frac{\lambda}{I}\right)}\,\mathrm{d}x \\ &= \int_{-\lambda}^{\lambda} \lim_{I\to\infty} \sqrt{A'\left(x+\frac{\lambda}{I}\right)\Phi'_{\mu,\nu}\left(x-\frac{\lambda}{I}\right)}\,\mathrm{d}x \\ &= \int_{-\lambda}^{\lambda} \sqrt{A'(x)\,\Phi'_{\mu,\nu}(x)}\,\mathrm{d}x. \end{aligned} \tag{6.176}$$

Secondly, we want to take $\lambda \to \infty$. The existence of this limit follows from monotone convergence theorem, which we can apply due to the monotonicity of A and $\Phi_{\mu,\nu}$, together with the boundedness of the continuous fidelity. Taking the limit gives us a bound in terms of the continuous fidelity

$$\begin{aligned}\liminf_{n\to\infty} \sqrt{1-\epsilon_0^*(n,R_\mu)} &\geq \lim_{\lambda\to\infty} \int_{-\lambda}^{\lambda} \sqrt{A'(x)\,\Phi'_{\mu,\nu}(x)}\,\mathrm{d}x \\ &\geq \int_{-\infty}^{\infty} \sqrt{A'(x)\,\Phi'_{\mu,\nu}(x)}\,\mathrm{d}x = \sqrt{\mathcal{F}\left(A',\Phi'_{\mu,\nu}\right)}. \end{aligned} \tag{6.177}$$

Lastly, we want to take a supremum over all continuously differentiable monotone functions $1 \geq A \geq \Phi$, which gives us the Rayleigh-normal distribution

$$\liminf_{n\to\infty} \epsilon_0^*(n,R_\mu) \leq 1 - \sup_{A\geq\Phi} \mathcal{F}\left(A',\Phi'_{\mu,\nu}\right) =: Z_\nu(\mu). \tag{6.178}$$

Using the above and the duality property of Rayleigh-normal distributions, Eq. (6.45), we obtain the lower bound on the optimal rate

$$R_0^*(n,\epsilon) \gtrsim \frac{1}{D(\boldsymbol{q})}\left[D(\boldsymbol{p}) + \sqrt{\frac{V(\boldsymbol{p})}{n}}\,Z_{1/\nu}^{-1}(\epsilon)\right]. \tag{6.179}$$

□

Optimality

We now proceed to the proof of the optimality of Proposition 6.20. To this end, we will employ two lemmas originally proved in Ref. [6.20]. The idea is to start by showing that, after a particular coarse-graining, the fidelity between Φ and $\Phi_{\mu,\nu}$ is close to the optimal fidelity $\sup_{A\geq\Phi} \mathcal{F}(A',\Phi'_{\mu,\nu}) = 1 - Z_\nu(\mu)$.

Lemma 6.23 (Lemma 17 of Ref. [6.20]). *For any $\zeta > 0$, there exist real numbers $s \leq t \leq t' \leq s'$ such that $\Phi'(x)/\Phi'_{\mu,\nu}(x)$ is strictly monotone decreasing for $x \in (s,s')$ and*

$$\frac{\Phi(t)}{\Phi_{\mu,\nu}(t)} = \frac{\Phi'(s)}{\Phi'_{\mu,\nu}(s)}, \tag{6.180a}$$

$$\frac{1-\Phi(t')}{1-\Phi_{\mu,\nu}(t')} = \frac{\Phi'(s')}{\Phi'_{\mu,\nu}(s')}. \tag{6.180b}$$

Moreover, if we define $\mathcal{F}_{t,t'}(\cdot,\cdot)$ to be the fidelity of distributions which have been coarse-grained on $x \leq t$ and $x \geq t'$, specifically

$$\mathcal{F}_{t,t'}(p,q) := \left(\sqrt{\int_{-\infty}^{t} p(x)\,\mathrm{d}x \int_{-\infty}^{t} q(x)\,\mathrm{d}x} + \int_{t}^{t'} \sqrt{p(x)q(x)}\,\mathrm{d}x + \sqrt{\int_{t'}^{\infty} p(x)\,\mathrm{d}x \int_{t}^{\infty} q(x)\,\mathrm{d}x} \right)^2, \tag{6.181}$$

then this coarse-grained fidelity has an upper bound

$$\mathcal{F}_{t,t'}(\Phi', \Phi'_{\mu,\nu}) - \zeta \leq \sup_{A \geq \Phi} \mathcal{F}(A', \Phi'_{\mu,\nu}). \tag{6.182}$$

Notice that Φ and $\Phi_{\mu,\nu}$ are exactly the distributions that appear in central limit theorem, Lemma 6.10. Thus, we would like to relate the fidelity $F(\boldsymbol{P}^n, \tilde{\boldsymbol{P}}^n)$ back to the Rayleigh-normal distribution via this coarse-grained fidelity between Gaussians. To be able to argue this for *any* $\tilde{\boldsymbol{P}}^n \succ \boldsymbol{Q}^n$, we will first give a sufficient condition for a distribution $\boldsymbol{a}$ to have the highest possible fidelity with respect to a second distribution $\boldsymbol{b}$ among all distributions satisfying a majorisation-like condition.

Lemma 6.24 (Lemma 15 of Ref. [6.20])**.** *Let $\boldsymbol{a}$ and $\boldsymbol{b}$ be probability distributions such that a_i/b_i is strictly decreasing for all i. Then, for any distribution $\boldsymbol{c}$ such that*

$$\sum_{i=1}^{k} a_i \leq \sum_{i=1}^{k} c_i \quad \forall k, \tag{6.183}$$

we have

$$F(\boldsymbol{c}, \boldsymbol{b}) \leq F(\boldsymbol{a}, \boldsymbol{b}), \tag{6.184}$$

with equality if and only if $\boldsymbol{c} = \boldsymbol{a}$.

We are now ready for the optimality proof.

Proof of Proposition 6.20 (optimality). To prove optimality we need to show that for any $\tilde{\boldsymbol{P}}^n \succ \boldsymbol{Q}^n$, the infidelity between $\tilde{\boldsymbol{P}}^n$ and $\boldsymbol{P}^n$ can be lower bounded by the Rayleigh-normal distribution. This, through Lemma 6.5, will yield a lower bound on the infidelity between any final state $\tilde{\boldsymbol{Q}}^n$ (i.e., any distribution majorised by the total initial state $\boldsymbol{P}^n$) and the total target state $\boldsymbol{Q}^n$. We will start by using the monotonicity of fidelity under coarse-graining to bound the fidelity between $\tilde{\boldsymbol{P}}^n$ and $\boldsymbol{P}^n$ by the fidelity between their coarse-grained versions (with coarse-graining over particularly chosen bins of indices). We will then use Lemma 6.24 (along with the monotonicity properties of Lemma 6.23) to bound the fidelity between coarse-grained versions of $\boldsymbol{P}^n$ and $\tilde{\boldsymbol{P}}^n$ by the fidelity between coarse-grained versions of $\boldsymbol{P}^n$ and $\boldsymbol{Q}^n$. Next, by applying Lemma 6.10, we will show that this coarse-grained fidelity asymptotes to a fidelity between Gaussians. We will conclude by returning to the last part of Lemma 6.23, which will allow us to give a final bound in terms of the Rayleigh-normal distribution.

Fix $\zeta > 0$ and $I \in \mathbb{N}$. Let $t, t' \in \mathbb{R}$ be those given by Lemma 6.23 and introduce

$$z_i := t(1 - i/I) + t'(i/I) \tag{6.185}$$

for $0 \leq i \leq I$. Define a set of bins

$$B_i := \{K_n(z_i), \ldots, K_n(z_{i+1}) - 1\} \tag{6.186}$$

for $0 \leq i < I$, as well as two end bins

$$B_{-1} := \{1, \ldots, K_n(t) - 1\}, \tag{6.187a}$$
$$B_I := \{K_n(t'), \ldots, \infty\}, \tag{6.187b}$$

so that $\{B_i\}$ now gives a partition of all indices. Using the monotonicity of fidelity under coarse-graining we get

$$\begin{aligned}\sqrt{F(\tilde{\boldsymbol{P}}^n, \boldsymbol{P}^n)} &= \sum_{i=-1}^{I} \sum_{j \in B_i} \sqrt{\tilde{P}_j^{n\downarrow} P_j^{n\downarrow}} \\ &\leq \sum_{i=-1}^{I} \sqrt{\sum_{j \in B_i} \tilde{P}_j^{n\downarrow} \sum_{j \in B_i} P_j^{n\downarrow}} =: r_n. \end{aligned} \tag{6.188}$$

We now define the limiting coarse-grained versions of total initial and target distributions

$$a_i := \lim_{n\to\infty} \sum_{j \in B_i} Q_j^{n\downarrow}, \tag{6.189a}$$
$$b_i := \lim_{n\to\infty} \sum_{j \in B_i} P_j^{n\downarrow}, \tag{6.189b}$$

for $-1 \leq i \leq I$. We note that the above limits all exist by Lemma 6.10, specifically

$$a_{-1} = \Phi(t), \tag{6.190a}$$
$$a_i = \Phi(z_{i+1}) - \Phi(z_i), \tag{6.190b}$$
$$a_I = 1 - \Phi(t'), \tag{6.190c}$$
$$b_{-1} = \Phi_{\mu,\nu}(t), \tag{6.190d}$$
$$b_i = \Phi_{\mu,\nu}(z_{i+1}) - \Phi_{\mu,\nu}(z_i), \tag{6.190e}$$
$$b_I = 1 - \Phi_{\mu,\nu}(t'). \tag{6.190f}$$

We would also like to analogously define a distribution $\boldsymbol{c}$ that would be a coarse-grained version of $\tilde{\boldsymbol{P}}^n$, but we have no guarantees that the corresponding limits exist. In lieu of this, we will use the vectorial Bolzano-Weirestrass Theorem[5], which gives that there exists a strictly increasing set of indices $\{m_l\}_l \subset \mathbb{N}$ such that $\{r_n\}$ limits to its limit superior

$$\lim_{l\to\infty} r_{m_l} = \limsup_{n\to\infty} r_n, \tag{6.191}$$

[5]Start with a subset of indices $\{m_l^{(-2)}\}_l \subseteq \mathbb{N}$ on which $\{r_n\}_n$ converges to its limit superior. The scalar version of Bolzano-Weirestrass gives a subset $\{m_l^{(i)}\}_l \subseteq \{m_l^{(i-1)}\}_l$ on which c_i exists. Applying this for each $-1 \leq i \leq I$, one obtains a set of indices $\{m_l\}_l := \{m_l^{(I)}\}_l$ with the desired property.

and that all the limits

$$c_i := \lim_{l\to\infty} \sum_{j\in B_i} \tilde{P}_j^{m_l\downarrow} \tag{6.192}$$

exist, for all $-1 \le i \le I$.

We now want to apply Lemma 6.24 to bound the fidelity $F(\boldsymbol{c},\boldsymbol{b})$ with $F(\boldsymbol{a},\boldsymbol{b})$, but first we must show that a_i/b_i is strictly decreasing. Lemma 6.23 allows us to relate this ratio at the ends, $i=-1$ and $i=I$, to the ratio of Gaussian derivatives,

$$\frac{a_{-1}}{b_{-1}} = \frac{\Phi(t)}{\Phi_{\mu,\nu}(t)} = \frac{\Phi'(s)}{\Phi'_{\mu,\nu}(s)}, \tag{6.193a}$$

$$\frac{a_I}{b_I} = \frac{1-\Phi(t')}{1-\Phi_{\mu,\nu}(t')} = \frac{\Phi'(s')}{\Phi'_{\mu,\nu}(s')}. \tag{6.193b}$$

For $0 \le i < I$ we can apply Cauchy's mean value theorem, which gives that there exists some $s_i \in (z_i, z_{i+1})$ the ratio of finite differences is given by a ratio of derivatives

$$\frac{a_i}{b_i} = \frac{\Phi(z_{i+1})-\Phi(z_i)}{\Phi_{\mu,\nu}(z_{i+1})-\Phi_{\mu,\nu}(z_i)} = \frac{\Phi'(s_i)}{\Phi'_{\mu,\nu}(s_i)}. \tag{6.194}$$

Given that $\{s, s_0, \dots, s_{I-1}, s'\}$ is a strictly increasing sequence, and that $\Phi'(x)/\Phi'_{\mu,\nu}(x)$ is strictly decreasing on (s,s') by Lemma 6.23, we therefore have that a_i/b_i is strictly decreasing as required.

Now that we have shown that a_i/b_i is strictly decreasing, we can apply Lemma 6.24. This gives us

$$\limsup_{n\to\infty} F(\tilde{\boldsymbol{P}}^n, \boldsymbol{P}^n) \le F(\boldsymbol{c},\boldsymbol{b}) \le F(\boldsymbol{a},\boldsymbol{b}). \tag{6.195}$$

Expanding this out, we have

$$\begin{aligned}\limsup_{n\to\infty} \sqrt{F(\tilde{\boldsymbol{P}}^n, \boldsymbol{P}^n)} \le{}& \sqrt{\Phi(z_0)\Phi_{\mu,\nu}(z_0)} \\ &+ \sum_{i=0}^{I-1} \sqrt{\Phi(z_{i+1})-\Phi(z_i)}\sqrt{\Phi_{\mu,\nu}(z_{i+1})-\Phi_{\mu,\nu}(z_i)} \\ &+ \sqrt{\Big(1-\Phi(z_I)\Big)\Big(1-\Phi_{\mu,\nu}(z_I)\Big)}.\end{aligned} \tag{6.196}$$

If we take $I\to\infty$, these finite differences above approach derivatives, and we get a bound in terms of a coarse-grained fidelity

$$\begin{aligned}\limsup_{n\to\infty} \sqrt{F(\tilde{\boldsymbol{P}}^n, \boldsymbol{P}^n)} \le{}& \sqrt{\Phi(t)\Phi_{\mu,\nu}(t)} \\ &+ \int_t^{t'} \sqrt{\Phi'(x)\Phi'_{\mu,\nu}(x)}\,\mathrm{d}x \\ &+ \sqrt{\Big(1-\Phi(t')\Big)\Big(1-\Phi_{\mu,\nu}(t')\Big)}.\end{aligned} \tag{6.197}$$

Finally, we apply the last part of Lemma 6.23, which allows us to bound this in terms of the Rayleigh-normal distribution, giving

$$\liminf_{n\to\infty} \delta(\tilde{\boldsymbol{P}}^n, \boldsymbol{P}^n) \geq Z_\nu(\mu) - \zeta. \tag{6.198}$$

Taking $\zeta \searrow 0$, we can once more use the equivalence between pre- and post-majorisation, Lemma 6.5, to conclude that

$$\liminf_{n\to\infty} \epsilon_0^*(n, R_\mu) \geq Z_\nu(\mu). \tag{6.199}$$

Using the above together with the duality property of Rayleigh-normal distributions, Eq.17 (6.45), we obtain the upper bound on the optimal rate

$$R_0^*(n,\epsilon) \lesssim \frac{1}{D(\boldsymbol{q})}\left[D(\boldsymbol{p}) + \sqrt{\frac{V(\boldsymbol{p})}{n}} Z_{1/\nu}^{-1}(\epsilon)\right]. \tag{6.200}$$

□

6.7 Outlook

In this paper we have derived the exact second-order asymptotics of state interconversion under thermal operations between any two energy-incoherent states. It is then natural to ask whether such a characterisation is also possible for general, not necessarily energy-incoherent, states. Due to the fact that thermal operations are time-translation covariant, such that coherence and athermality form independent resources [6.13, 6.31], it seems unlikely that the current approach can be easily generalised. Instead, one would need to rely on the full power of Gibbs-preserving maps [6.60, 6.61] that form a superset of the thermal operations. For such maps, we believe that a reasonable conjecture is in fact given by Eq. (6.1a), with the relative entropy and relative entropy variance replaced by their fully quantum analogues given in Refs. [6.37] and [6.21, 6.22], respectively.

We also provided a physical interpretation of our main result by considering several thermodynamic scenarios and explaining how our work can be employed to rigorously address the problem of thermodynamic irreversibility. We derived optimal values of distillable work and work of formation, and related them to the infidelity of these processes. This could potentially be used to clarify the notion of imperfect work [6.40, 6.50, 6.59], and to construct a comparison platform allowing one to continuously distinguish between work-like and heat-like forms of energy. We also discussed thermodynamic processes with finite-size working bodies, focusing particularly on the optimal performance of heat engines. We have shown that there are non-trivial conditions under which an engine can operate at Carnot efficiency and extract perfect work. This opens the possibility of engineering finite heat-baths and working bodies in order to minimise undesirable dissipation of free energy. Moreover, our formalism is general enough to address other interesting problems involving finite-size baths, like fluctuation theorems, Landauers' principle or the third law of thermodynamics [6.53, 6.54, 6.56].

A number of natural technical extensions to our result suggest themselves. We have used the infidelity as our error measure, and conjecture that Theorem 6.3 will also hold when ϵ is a bound on the total variational distance. Our second-order expansion falls into a larger class of results known as *small*

deviation bounds, in which we consider a fixed error threshold ϵ. Two natural extensions are to the regime of *large deviations* [6.62], in which a fixed rate is considered, and *moderate deviations* [6.63,6.64], in which the rate approaches its optimum *and* the error vanishes. Last, but not least, we expect that our treatment of approximate majorisation can be extended to cover other distance measures.

Acknowledgements

We want to thank Nelly Ng for insightful discussions on the interplay between smoothing and thermomajorisation, as well as for the comments on the manuscript; Masahito Hayashi for discussions and feedback on early technical notes; Matteo Lostaglio for useful discussions on thermodynamic interpretations of second-order rates; and Joe Renes for helpful comments on the manuscript. CC and KK acknowledge support from the ARC via the Centre of Excellence in Engineered Quantum Systems (EQuS), project number CE110001013. CC also acknowledges support from the Australian Institute for Nanoscale Science and Technology Postgraduate Scholarship (John Makepeace Bennett Gift). MT is funded by an ARC Discovery Early Career Researcher Award, project number DE160100821.

6.A Work extraction with Carnot efficiency

Here we consider a thermodynamic process involving a heat engine with a finite working body. Suppose the working body starts at the cold temperature, T_c. Then, the heat flowing from the hot bath will steadily increase the temperature of the working body, from T_c to $T_{\mathrm{c}'}$. We assume that thermalisation happens on a much shorter time-scale than the heat flow, so that the working body is at all times in thermal equilibrium. We are now interested in the amount of work that can be extracted if the engine at each time step operates with the maximal allowed Carnot efficiency (with respect to the background hot temperature and the instantaneous temperature of the working body).

By definition, the efficiency of an infinitesimal process involving the flow of heat dQ_in from the hot bath to the engine, dQ_out of which flows out to the cold bath, while the remaining energy is converted into work dW, is given by

$$\eta := \frac{\mathrm{d}W}{\mathrm{d}Q_\mathrm{in}} = \frac{\mathrm{d}W}{\mathrm{d}W + \mathrm{d}Q_\mathrm{out}}, \tag{6.201}$$

where the equality comes from the conservation of energy. Hence, the work extracted during the infinitesimal flow of heat dQ_out into the cold bath at temperature T_x that heats it up by dT_x,

$$\mathrm{d}Q_\mathrm{out} = \frac{\mathrm{d}\langle E\rangle_{\boldsymbol{\gamma}_\mathrm{x}}}{\mathrm{d}T_\mathrm{x}}\,\mathrm{d}T_\mathrm{x}, \tag{6.202}$$

is given by

$$\mathrm{d}W = \frac{\eta}{1-\eta}\,\mathrm{d}Q_\mathrm{out}. \tag{6.203}$$

The Carnot efficiency of an engine acting between the hot bath at fixed temperature T_h and the working body at varying temperature T_x is given by

$$\eta_\mathrm{c}(T_\mathrm{x}) = 1 - \frac{T_\mathrm{x}}{T_\mathrm{h}}. \tag{6.204}$$

Thus, the work produced by an engine working at maximum allowed efficiency is given by

$$\mathrm{d}W = \left(\frac{T_\mathrm{h}}{T_\mathrm{x}} - 1\right) \frac{\mathrm{d}\langle E\rangle_{\boldsymbol{\gamma}_\mathrm{x}}}{\mathrm{d}T_\mathrm{x}}\, \mathrm{d}T_\mathrm{x}. \tag{6.205}$$

We now want to calculate the total work W extracted by such an optimal engine while the temperature of the working body changes from T_c to $T_{\mathrm{c}'}$,

$$W = \int_{T_\mathrm{c}}^{T_{\mathrm{c}'}} \mathrm{d}W. \tag{6.206}$$

By simply integrating by parts we get

$$W = \left(\frac{T_\mathrm{h}}{T_\mathrm{x}} - 1\right) \langle E\rangle_{\boldsymbol{\gamma}_\mathrm{x}} \Bigg|_{T_\mathrm{x}=T_\mathrm{c}}^{T_\mathrm{x}=T_{\mathrm{c}'}} + T_\mathrm{h} \int_{T_\mathrm{c}}^{T_{\mathrm{c}'}} \frac{\langle E\rangle_{\boldsymbol{\gamma}_\mathrm{x}}}{T_\mathrm{x}^2}\, \mathrm{d}T_\mathrm{x}. \tag{6.207}$$

The second term can be calculated by switching from temperature T_x to inverse temperature β_x and recalling that the average energy is a negative derivative of $\log \mathcal{Z}_\mathrm{x}$ over β_x:

$$T_\mathrm{h} \int_{T_\mathrm{c}}^{T_{\mathrm{c}'}} \frac{\langle E\rangle_{\boldsymbol{\gamma}_\mathrm{x}}}{T_\mathrm{x}^2}\, \mathrm{d}T_\mathrm{x} = \frac{1}{\beta_\mathrm{h}} \int_{\beta_\mathrm{c}}^{\beta_{\mathrm{c}'}} \frac{\mathrm{d}\log \mathcal{Z}_\mathrm{x}}{\mathrm{d}\beta_\mathrm{x}}\, \mathrm{d}\beta_\mathrm{x} = \frac{1}{\beta_\mathrm{h}} \left(\log \mathcal{Z}_{\mathrm{c}'} - \log \mathcal{Z}_\mathrm{c}\right). \tag{6.208}$$

By noting that the entropy of a thermal equilibrium state is given by

$$H(\boldsymbol{\gamma}_\mathrm{x}) = \beta_\mathrm{x} \langle E\rangle_{\boldsymbol{\gamma}_\mathrm{x}} + \log \mathcal{Z}_\mathrm{x}, \tag{6.209}$$

we thus have

$$W = \left(\langle E\rangle_{\boldsymbol{\gamma}_\mathrm{c}} - \frac{H(\boldsymbol{\gamma}_\mathrm{c})}{\beta_\mathrm{h}}\right) - \left(\langle E\rangle_{\boldsymbol{\gamma}_{\mathrm{c}'}} - \frac{H(\boldsymbol{\gamma}_{\mathrm{c}'})}{\beta_\mathrm{h}}\right). \tag{6.210}$$

Finally, comparing the above with Eq. (6.25) we arrive at

$$W = k_B T_\mathrm{h} \big(D(\boldsymbol{\gamma}_\mathrm{c} || \boldsymbol{\gamma}_\mathrm{h}) - D(\boldsymbol{\gamma}_{\mathrm{c}'} || \boldsymbol{\gamma}_\mathrm{h})\big), \tag{6.211}$$

which is equal to the change of free energy of the finite bath.

6.B Proof of Lemma 6.4

6.B.1 Preliminaries

In Ref. [6.48] an explicit construction the solution $\tilde{\boldsymbol{p}}^\star$ to the following maximisation problem,

$$\tilde{\boldsymbol{p}}^\star = \operatorname*{arg\,max}_{\tilde{\boldsymbol{p}}:\ \tilde{\boldsymbol{p}} \succ \boldsymbol{q}} F(\boldsymbol{p}, \tilde{\boldsymbol{p}}), \tag{6.212}$$

was given. We will now describe the construction of this optimal distribution, as it is crucial for our proof, the second part which will very closely follow the reasoning presented in Ref. [6.48]. As explained

in Section 6.3.3, without loss of generality we will assume that all the distributions are non-increasingly ordered.

First, for any distribution $\boldsymbol{a}$ define

$$E_k^{\boldsymbol{a}} := \sum_{i=k}^{d} a_i, \quad \Delta_{k_1}^{k_2}(\boldsymbol{a}) := \sum_{i=k_1}^{k_2-1} a_i = E_{k_1}^{\boldsymbol{a}} - E_{k_2}^{\boldsymbol{a}}. \tag{6.213}$$

Note that $\boldsymbol{p} \succ \boldsymbol{q}$ is equivalent to $E_k^{\boldsymbol{p}} \leq E_k^{\boldsymbol{q}}$ for all k. Now, for a given $\boldsymbol{p}$ and $\boldsymbol{q}$ the construction of $\tilde{\boldsymbol{p}}^\star$ is given by the following iterative procedure. Set $l_0 = d+1$ and define

$$l_j := \underset{k<l_{j-1}}{\arg\min} \frac{\Delta_k^{l_{j-1}}(\boldsymbol{q})}{\Delta_k^{l_{j-1}}(\boldsymbol{p})}, \quad r_j := \frac{\Delta_{l_j}^{l_{j-1}}(\boldsymbol{q})}{\Delta_{l_j}^{l_{j-1}}(\boldsymbol{p})}. \tag{6.214}$$

If the minimisation defining l_j does not have a unique solution then l_j is chosen to be the smallest possible. We will also denote by N an index for which $l_N = 1$. The i-th entry of the optimal vector for $i \in \{l_j, \ldots, l_{j-1} - 1\}$ is then given by

$$\tilde{p}_i^\star = r_j p_i. \tag{6.215}$$

It is straightforward to verify that $\tilde{\boldsymbol{p}}^\star$ is normalised, and $\tilde{\boldsymbol{p}}^\star \succ \boldsymbol{q}$ as the construction guarantees that $E_k^{\boldsymbol{p}} \leq E_k^{\boldsymbol{q}}$ for all k. Moreover, the optimal fidelity between $\boldsymbol{p}$ and a distribution that majorises $\boldsymbol{q}$ is given by

$$\begin{aligned}
\sqrt{F(\boldsymbol{p}, \tilde{\boldsymbol{p}}^\star)} &= \sum_{i=1}^{d} \sqrt{p_i \tilde{p}_i^\star} = \sum_{j=1}^{N} \sum_{i=l_j}^{l_{j-1}-1} \sqrt{p_i \tilde{p}_i^\star} \\
&= \sum_{j=1}^{N} \left(\frac{\Delta_{l_j}^{l_{j-1}}(\boldsymbol{q})}{\Delta_{l_j}^{l_{j-1}}(\boldsymbol{p})} \right)^{\frac{1}{2}} \sum_{i=l_j}^{l_{j-1}-1} p_i \\
&= \sum_{j=1}^{N} \left(\Delta_{l_j}^{l_{j-1}}(\boldsymbol{q}) \Delta_{l_j}^{l_{j-1}}(\boldsymbol{p}) \right)^{\frac{1}{2}}.
\end{aligned} \tag{6.216}$$

The crucial observation in proving the optimality of the above, which we will also need in our proof, is that for all j we have

$$r_j < r_{j+1}. \tag{6.217}$$

This follows from the definition of l_j, r_j and the fact that for $a, b, c, d > 0$ one has (see Ref. [6.48] for details)

$$\frac{a}{b} \leq \frac{a+c}{b+d} \quad \Longleftrightarrow \quad \frac{a}{b} < \frac{c}{d}. \tag{6.218}$$

6.B.2 Proper proof

Proof. We will prove the equality in Eq. (6.86) by showing that the following two inequalities hold

$$\max_{\tilde{\boldsymbol{p}}:\ \tilde{\boldsymbol{p}} \succ \boldsymbol{q}} F(\boldsymbol{p}, \tilde{\boldsymbol{p}}) \leq \max_{\tilde{\boldsymbol{q}}:\ \boldsymbol{p} \succ \tilde{\boldsymbol{q}}} F(\boldsymbol{q}, \tilde{\boldsymbol{q}}), \tag{6.219a}$$

$$\max_{\tilde{\boldsymbol{p}}:\ \tilde{\boldsymbol{p}} \succ \boldsymbol{q}} F(\boldsymbol{p}, \tilde{\boldsymbol{p}}) \geq \max_{\tilde{\boldsymbol{q}}:\ \boldsymbol{p} \succ \tilde{\boldsymbol{q}}} F(\boldsymbol{q}, \tilde{\boldsymbol{q}}). \tag{6.219b}$$

We start with the easier part, Eq. (6.219a). It is enough to show that the inequality holds for any $\tilde{\boldsymbol{p}}$ within the constraints. Let us then take any $\tilde{\boldsymbol{p}}$ such that it majorises $\boldsymbol{q}$. Due to Theorem 6.2 this is equivalent to the existence of a bistochastic matrix B_0 such that $B_0\tilde{\boldsymbol{p}} = \boldsymbol{q}$. This implies that

$$\max_{\tilde{\boldsymbol{q}}:\ \boldsymbol{p}\succ\tilde{\boldsymbol{q}}} F(\boldsymbol{q},\tilde{\boldsymbol{q}}) = \max_{B} F(\boldsymbol{q}, B\boldsymbol{p}), \tag{6.220}$$

where the maximisation on the right hand side is over all bistochastic matrices B. We now observe that

$$\begin{aligned} \max_{B} F(\boldsymbol{q}, B\boldsymbol{p}) &= \max_{B} F(B_0\tilde{\boldsymbol{p}}, B\boldsymbol{p}) && \text{(6.221a)}\\ &\geq F(B_0\tilde{\boldsymbol{p}}, B_0\boldsymbol{p}) && \text{(6.221b)}\\ &\geq F(\tilde{\boldsymbol{p}}, \boldsymbol{p}), && \text{(6.221c)} \end{aligned}$$

where in the last step we used the fact that fidelity obeys data processing inequality. We thus have

$$\max_{\tilde{\boldsymbol{q}}:\ \boldsymbol{p}\succ\tilde{\boldsymbol{q}}} F(\boldsymbol{q},\tilde{\boldsymbol{q}}) \geq F(\tilde{\boldsymbol{p}}, \boldsymbol{p}), \tag{6.222}$$

for any $\tilde{\boldsymbol{p}}$ majorising $\boldsymbol{q}$, and so Eq. (6.219a) holds.

We now proceed to proving Eq. (6.219b). It is again enough to show that the inequality holds for any $\tilde{\boldsymbol{q}}$ within the constraints. Let us then take any $\tilde{\boldsymbol{q}}$ such that it is majorised by $\boldsymbol{p}$. We now have

$$\begin{aligned} \sqrt{F(\boldsymbol{q},\tilde{\boldsymbol{q}})} &= \sum_{i=1}^{d} \sqrt{q_i\tilde{q}_i} = \sum_{j=1}^{N}\sum_{i=l_j}^{l_{j-1}-1} \sqrt{q_i\tilde{q}_i}\\ &\leq \sum_{j=1}^{N} \left(\sum_{i=l_j}^{l_{j-1}-1} q_i\right)^{\frac{1}{2}} \left(\sum_{i=l_j}^{l_{j-1}-1} \tilde{q}_i\right)^{\frac{1}{2}}\\ &= \sum_{j=1}^{N} \left(\Delta_{l_j}^{l_{j-1}}(\boldsymbol{q})\Delta_{l_j}^{l_{j-1}}(\tilde{\boldsymbol{q}})\right)^{\frac{1}{2}}, \end{aligned} \tag{6.223}$$

where l_j and $\Delta_{k_1}^{k_2}$ are defined as in Eqs. (6.213)-(6.214). We now introduce

$$x_j := E_{l_j}^{\tilde{\boldsymbol{q}}} - E_{l_j}^{\boldsymbol{p}} \geq 0, \tag{6.224}$$

where the inequality holds for every j because $\boldsymbol{p}\succ\tilde{\boldsymbol{q}}$. Observing that

$$\Delta_{l_j}^{l_{j-1}}(\tilde{\boldsymbol{q}}) = \Delta_{l_j}^{l_{j-1}}(\boldsymbol{p}) + x_j - x_{j-1} \geq 0 \tag{6.225}$$

we arrive at $\sqrt{F(\boldsymbol{q},\tilde{\boldsymbol{q}})} \leq f(\boldsymbol{x})$ with

$$f(\boldsymbol{x}) := \sum_{j=1}^{N} \left(\Delta_{l_j}^{l_{j-1}}(\boldsymbol{q}) \left(\Delta_{l_j}^{l_{j-1}}(\boldsymbol{p}) + x_j - x_{j-1}\right)\right)^{\frac{1}{2}}. \tag{6.226}$$

We will now show that $f(\boldsymbol{x})$ achieves its maximum within the positive orthant $x_j \geq 0$ when $\boldsymbol{x} = \boldsymbol{0}$, which will finish the proof. This is because then

$$F(\boldsymbol{q}, \tilde{\boldsymbol{q}}) \leq F(\boldsymbol{p}, \tilde{\boldsymbol{p}}^\star) = \max_{\tilde{\boldsymbol{p}}:\ \tilde{\boldsymbol{p}} \succ \boldsymbol{q}} F(\boldsymbol{p}, \tilde{\boldsymbol{p}}) \tag{6.227}$$

for any $\tilde{\boldsymbol{q}}$ majorised by $\boldsymbol{p}$, which implies Eq. (6.219b). First, by direct calculation one can find a matrix M of second derivatives of $f(\boldsymbol{x})$, i.e., $M_{ij} = \frac{\partial^2 f(\boldsymbol{x})}{\partial x_i \partial x_j}$. Then, using Gershgorin circle theorem, one can verify that M is negative definite in the allowed region of $\boldsymbol{x}$, so that there are no local extrema and the maximal value must be obtained at the boundary. Finally,

$$\left.\frac{\partial f(\boldsymbol{x})}{\partial x_j}\right|_{x_j=0} = \frac{1}{2}\left(\sqrt{r_j} - \sqrt{r_{j+1}}\right) < 0, \tag{6.228}$$

which means that the maximal value is obtained for $\boldsymbol{x} = [0, \ldots, 0]$, which finishes the proof. □

6.C Efficient algorithm for calculating interconversion infidelities

The construction given in Section 6.B gives a natural algorithm for calculating $\max_{\tilde{\boldsymbol{p}}:\tilde{\boldsymbol{p}} \succ \boldsymbol{q}} F(\boldsymbol{p}, \tilde{\boldsymbol{p}})$. The run-time of this algorithm is $\mathcal{O}(d^2)$, where d is the size of the input distributions. We now want to argue that this algorithm can be adapted for states described by $\boldsymbol{p}^{\otimes n}$, such that the optimal interconversion rates can be numerically calculated in a time which is efficient in n, as is done in Figures 6.4 and 6.10.

The key property we will leverage is that whilst distributions such as $\boldsymbol{p}^{\otimes n}$ have an exponential number of entries, they only possess a polynomial number of *distinct* entries (in this case $\mathcal{O}(n^{d-1})$). As majorisation is invariant under permutations, it is only the distinct entries (and their degeneracies) that are relevant to our calculation.

Consider taking as input two distributions $\boldsymbol{P}$ and $\boldsymbol{Q}$ which possess $D = \exp(\mathcal{O}(n))$ total entries, but each only $\mathrm{poly}(n)$ distinct entries. This means there exists indices $1 = i_1 < i_2 < \cdots < i_t < i_{t+1} = D + 1$, where $t = \mathrm{poly}(n)$, such that $P_i^{\downarrow}$ and $Q_i^{\downarrow}$ are constant on each interval $i \in \{i_s, \ldots, i_{s+1} - 1\}$ for $s \in \{1, \ldots, t\}$.

The main step in the algorithm, and the bottleneck giving an exponential run-time, is calculating the pivot indices $\{l_j\}_j$ used to construct $\tilde{\boldsymbol{P}}$. Specifically these take the form

$$l_j := \operatorname*{arg\,max}_{k < l_{j-1}} \frac{\sum_{i=k}^{l_{j-1}-1} Q_i^{\downarrow}}{\sum_{i=k}^{l_{j-1}-1} P_i^{\downarrow}}. \tag{6.229}$$

Using the constancy of $P_i^{\downarrow}$ and $Q_i^{\downarrow}$ on each of the intervals we can see that, for a fixed s, the function being optimised takes the form

$$\frac{\sum_{i=k}^{l_{j-1}-1} Q_i^{\downarrow}}{\sum_{i=k}^{l_{j-1}-1} P_i^{\downarrow}} = \frac{\alpha k + \beta}{\gamma k + \delta}, \tag{6.230}$$

for any $k \in \{i_s, \ldots, i_{s+1} - 1\}$. As this function is monotonic as a function of i, we conclude that the indices l_j must lie on the edges of these intervals, i.e. $\{l_j\}_j \subseteq \{i_s\}_s$. This means that we can restrict our

attention only to the these 'edge indices' without loss of generality, lowering the algorithms run-time down from $\mathcal{O}(D^2)$ to $\mathcal{O}(t^2)$.

Using the above argument, we now have an efficient algorithm for computing $\epsilon_0^*(n, R; \boldsymbol{p}, \boldsymbol{q})$. Utilising the idea of embedding, this also allows us to calculate the thermal variant of this, $\epsilon_\beta^*(n, R; \boldsymbol{p}, \boldsymbol{q})$. Finally, by sweeping over R, we can use this to calculate $R_\beta^*(n, R; \boldsymbol{p}, \boldsymbol{q})$. Examples of this are shown in Figures 6.4 and 6.10.

6.D Approximate thermomajorisation based on total variation distance

Here we show how to modify the proofs of Lemmas 6.5, 6.6 and 6.7, so that they hold for approximate thermomajorisation defined with total variation distance $\delta(\boldsymbol{p}, \boldsymbol{q}) = \frac{1}{2}||\boldsymbol{p} - \boldsymbol{q}||$. In all three proofs one needs to replace the statement "embedding is fidelity-preserving" with "embedding preserves the total variation distance", and "fidelity is non-decreasing under stochastic maps" with "total variation distance is non-increasing under stochastic maps". This is already enough to prove modified Lemma 6.6. In Lemma 6.7 we additionally need to replace the expression $\epsilon_0 = (3 - 2\sqrt{2})/6$ with $\epsilon_0 = 1/6$. Finally, we replace the second part of Lemma 6.5 with the following reasoning.

Assume that $\boldsymbol{p} \succ_\epsilon \boldsymbol{q}$. This means that there exists a $\tilde{\boldsymbol{q}}$ such that $\boldsymbol{p} \succ \tilde{\boldsymbol{q}}$ and $\delta(\boldsymbol{q}, \tilde{\boldsymbol{q}}) \leq \epsilon$. Define M to be the largest integer such that

$$\sum_{i=1}^{M} p_i^{\downarrow} \leq 1 - \epsilon, \tag{6.231}$$

and then define $\tilde{\boldsymbol{p}}$ by cutting off the tail of $\boldsymbol{p}$, and placing all of its mass into the largest element

$$\tilde{p}_i := \begin{cases} p_i^{\downarrow} + \epsilon & \text{for } i = 1, \\ p_i^{\downarrow} & \text{for } 1 < i < M, \\ p_i^{\downarrow} + 1 - \epsilon - \sum_{i=1}^{M} p_i^{\downarrow} & \text{for } i = M, \\ 0 & \text{for } i > M. \end{cases} \tag{6.232}$$

By definition of M we can see that $\delta(\boldsymbol{p}, \tilde{\boldsymbol{p}}) = \epsilon$. We now need to show that $\tilde{\boldsymbol{p}} \succ \boldsymbol{q}$. For $k \geq M$ we have

$$\sum_{i=1}^{k} \tilde{p}_i^{\downarrow} = 1 \geq \sum_{i=1}^{k} q_i^{\downarrow}. \tag{6.233}$$

For $k < M$, we can use $\boldsymbol{p} \succ \tilde{\boldsymbol{q}}$ and $\delta(\boldsymbol{q}, \tilde{\boldsymbol{q}}) \leq \epsilon$ to give

$$\sum_{i=1}^{k} \tilde{p}_i^{\downarrow} = \sum_{i=1}^{k} p_i^{\downarrow} + \epsilon \geq \sum_{i=1}^{k} \tilde{q}_i^{\downarrow} + \epsilon \geq \sum_{i=1}^{k} q_i^{\downarrow}. \tag{6.234}$$

Thus $\boldsymbol{p} \succ_\epsilon \boldsymbol{q} \iff \boldsymbol{p} \succ_\epsilon \boldsymbol{q}$.

6.E Proof of Lemma 6.8

Proof. Consider a discrete random variable $L = -\log a$, such that $\langle L \rangle_{\boldsymbol{a}} = H(\boldsymbol{a})$ and $\mathrm{Var}_{\boldsymbol{a}}(L) = V(\boldsymbol{a})$, and therefore

$$\log k_n(x) = n\langle L \rangle_{\boldsymbol{a}} + x\sqrt{n\,\mathrm{Var}_{\boldsymbol{a}}(L)}. \tag{6.235}$$

If we let $\{L_j\}_{1\leq j\leq n}$ be i.i.d. copies of L, then we can write the tail bound of $\boldsymbol{a}^{\otimes n}$ in terms of tail bounds on the average of these variables as

$$\sum_i \left\{ \left(\boldsymbol{a}^{\otimes n}\right)_i \middle| \left(\boldsymbol{a}^{\otimes n}\right)_i \geq 1/k_n(x) \right\} \tag{6.236}$$

$$= \sum_{i_1,\ldots,i_n} \left\{ \prod_{j=1}^n a_{i_j} \middle| \prod_{j=1}^n a_{i_j} \geq 1/k_n(x) \right\} \tag{6.237}$$

$$= \sum_{i_1,\ldots,i_n} \left\{ \prod_{j=1}^n a_{i_j} \middle| -\sum_{j=1}^n \log a_{i_j} \geq \log k_n(x) \right\} \tag{6.238}$$

$$= \Pr\left[\sum_{j=1}^n L_j \leq n\langle L \rangle_{\boldsymbol{a}} + x\sqrt{n\,\mathrm{Var}_{\boldsymbol{a}}(L)} \right]. \tag{6.239}$$

By applying the standard central limit theorem

$$\lim_{n\to\infty} \Pr\left[\sum_{j=1}^n \left(\frac{L_j - \langle L \rangle_{\boldsymbol{a}}}{\sqrt{\mathrm{Var}_{\boldsymbol{a}}(L)}} \right) \leq x\sqrt{n} \right] = \Phi\left(x\right), \tag{6.240}$$

we get the desired bound. □

Bibliography

6.1 Robin Giles, *Mathematical Foundations of Thermodynamics* (Pergamon Press, 1964).

6.2 Bob Coecke, Tobias Fritz, and Robert W Spekkens, "A mathematical theory of resources," Inform. Comput. **250**, 59–86 (2016).

6.3 M. Horodecki and J. Oppenheim, "(Quantumness in the context of) resource theories," Int. J. Mod. Phys. B **27**, 1345019 (2013a).

6.4 Ryszard Horodecki, Paweł Horodecki, Michał Horodecki, and Karol Horodecki, "Quantum entanglement," Rev. Mod. Phys. **81**, 865–942 (2009).

6.5 T Baumgratz, M Cramer, and MB Plenio, "Quantifying coherence," Phys. Rev. Lett. **113**, 140401 (2014).

6.6 Iman Marvian and Robert W Spekkens, "How to quantify coherence: distinguishing speakable and unspeakable notions," Phys. Rev. A **94**, 052324 (2016).

6.7 Andreas Winter and Dong Yang, "Operational resource theory of coherence," Phys. Rev. Lett. **116**, 120404 (2016).

6.8 I. Marvian and R. W. Spekkens, "The theory of manipulations of pure state asymmetry: I. Basic tools, equivalence classes and single copy transformations," New J. Phys. **15**, 033001 (2013).

6.9 Iman Marvian, *Symmetry, Asymmetry and Quantum Information*, Ph.D. thesis, University of Waterloo (2012).

6.10 D. Janzing, P. Wocjan, R. Zeier, R. Geiss, and Th. Beth, "Thermodynamic cost of reliability and low temperatures: tightening Landauer's principle and the second law," Int. J. Theor. Phys. **39**, 2717–2753 (2000).

6.11 M. Horodecki and J. Oppenheim, "Fundamental limitations for quantum and nanoscale thermodynamics," Nat. Commun. **4**, 2059 (2013b).

6.12 Fernando G. S. L. Brandão, Michał Horodecki, Jonathan Oppenheim, Joseph M. Renes, and Robert W. Spekkens, "Resource theory of quantum states out of thermal equilibrium," Phys. Rev. Lett. **111**, 250404 (2013).

6.13 Matteo Lostaglio, Kamil Korzekwa, David Jennings, and Terry Rudolph, "Quantum coherence, time-translation symmetry, and thermodynamics," Phys. Rev. X **5**, 021001 (2015a).

6.14 Piotr Ćwikliński, Michał Studziński, Michał Horodecki, and Jonathan Oppenheim, "Limitations on the evolution of quantum coherences: towards fully quantum second laws of thermodynamics," Phys. Rev. Lett. **115**, 210403 (2015).

6.15 V. Narasimhachar and G. Gour, "Low-temperature thermodynamics with quantum coherence," Nat. Commun. **6**, 7689 (2015).

6.16 Gilad Gour, David Jennings, Francesco Buscemi, Runyao Duan, and Iman Marvian, "Quantum majorization and a complete set of entropic conditions for quantum thermodynamics," arXiv:1708.04302 (2017).

6.17 Michael A Nielsen, "Conditions for a class of entanglement transformations," Phys. Rev. Lett. **83**, 436 (1999).

6.18 Charles H Bennett, Herbert J Bernstein, Sandu Popescu, and Benjamin Schumacher, "Concentrating partial entanglement by local operations," Phys. Rev. A **53**, 2046 (1996).

6.19 Michał Horodecki, Jonathan Oppenheim, and Carlo Sparaciari, "Approximate majorization," arXiv:1706.05264 (2017).

6.20 Wataru Kumagai and Masahito Hayashi, "Second-order asymptotics of conversions of distributions and entangled states based on Rayleigh-normal probability distributions," IEEE Trans. Inf. Theory **63**, 1829–1857 (2017).

6.21 Marco Tomamichel and Masahito Hayashi, "A hierarchy of information quantities for finite block length analysis of quantum tasks," IEEE Trans. Inf. Theory **59**, 7693–7710 (2013).

6.22 Ke Li, "Second-order asymptotics for quantum hypothesis testing," Ann. Stat. **42**, 171–189 (2014).

6.23 Marco Tomamichel and Vincent Y. F. Tan, "Second-order asymptotics for the classical capacity of image-additive quantum channels," Commun. Math. Phys. **338**, 103–137 (2015).

6.24 Nilanjana Datta, Marco Tomamichel, and Mark M. Wilde, "On the second-order asymptotics for entanglement-assisted communication," Quantum Inf. Process. **15**, 2569–2591 (2016).

6.25 Nilanjana Datta and Felix Leditzky, "Second-order asymptotics for source coding, dense coding, and pure-state entanglement conversions," IEEE Trans. Inf. Theory **61**, 582–608 (2015).

6.26 Mark M. Wilde, Joseph M. Renes, and Saikat Guha, "Second-order coding rates for pure-loss bosonic channels," Quantum Inf. Process. **15**, 1289–1308 (2016).

6.27 Marco Tomamichel, Mario Berta, and Joseph M. Renes, "Quantum coding with finite resources," Nat. Commun. **7**, 11419 (2016).

6.28 Mark M. Wilde, Marco Tomamichel, Seth Lloyd, and Mario Berta, "Gaussian hypothesis testing and quantum illumination," Phys. Rev. Lett. **119**, 120501 (2017).

6.29 John Goold, Marcus Huber, Arnau Riera, Lídia del Rio, and Paul Skrzypczyk, "The role of quantum information in thermodynamics – a topical review," J. Phys. A **49**, 143001 (2016).

6.30 Johan Åberg, "Catalytic coherence," Phys. Rev. Lett. **113**, 150402 (2014).

6.31 Matteo Lostaglio, David Jennings, and Terry Rudolph, "Description of quantum coherence in thermodynamic processes requires constraints beyond free energy," Nat. Commun. **6**, 6383 (2015b).

6.32 Raam Uzdin, Amikam Levy, and Ronnie Kosloff, "Equivalence of quantum heat machines, and quantum-thermodynamic signatures," Phys. Rev. X **5**, 031044 (2015).

6.33 Kamil Korzekwa, Matteo Lostaglio, Jonathan Oppenheim, and David Jennings, "The extraction of work from quantum coherence," New J. Phys. **18**, 023045 (2016).

6.34 Robert Alicki and Mark Fannes, "Entanglement boost for extractable work from ensembles of quantum batteries," Phys. Rev. E **87**, 042123 (2013).

6.35 Martí Perarnau-Llobet, Karen V Hovhannisyan, Marcus Huber, Paul Skrzypczyk, Nicolas Brunner, and Antonio Acín, "Extractable work from correlations," Phys. Rev. X **5**, 041011 (2015).

6.36 Jonatan Bohr Brask, Géraldine Haack, Nicolas Brunner, and Marcus Huber, "Autonomous quantum thermal machine for generating steady-state entanglement," New J. Phys. **17**, 113029 (2015).

6.37 H. Umegaki, "Conditional expectation in an operator algebra," Kodai Math. Sem. Rep. **14**, 59–85 (1962).

6.38 A. Uhlmann, "The transition probability for states of star-algebras," Ann. Phys. (N. Y). **497**, 524–532 (1985).

6.39 Kosuke Ito, Wataru Kumagai, and Masahito Hayashi, "Asymptotic compatibility between local-operations-and-classical-communication conversion and recovery," Phys. Rev. A **92**, 052308 (2015).

6.40 Nelly Huei Ying Ng, Mischa Prebin Woods, and Stephanie Wehner, "Surpassing the Carnot efficiency by extracting imperfect work," New J. Phys. **19**, 113005 (2017).

6.41 F. G. S. L. Brandão, M. Horodecki, N. H. Y. Ng, J. Oppenheim, and S. Wehner, "The second laws of quantum thermodynamics," Proc. Natl. Acad. Sci. U.S.A. **112**, 3275 (2015).

6.42 Kamil Korzekwa, *Coherence, thermodynamics and uncertainty relations*, Ph.D. thesis, Imperial College London (2016).

6.43 Rajendra Bhatia, *Matrix analysis* (Springer Science & Business Media, 2013).

6.44 Ernst Ruch, Rudolf Schranner, and Thomas H. Seligman, "The mixing distance," J. Chem. Phys. **69** (1978).

6.45 Joseph M Renes, "Relative submajorization and its use in quantum resource theories," J. Math. Phys. **57**, 122202 (2016).

6.46 Remco Van Der Meer, Nelly Huei Ying Ng, and Stephanie Wehner, "Smoothed generalized free energies for thermodynamics," Phys. Rev. A **96**, 062135 (2017).

6.47 Michał Horodecki, Karol Horodecki, Paweł Horodecki, Ryszard Horodecki, Jonathan Oppenheim, Aditi Sen(De), and Ujjwal Sen, "Local information as a resource in distributed quantum systems," Phys. Rev. Lett. **90**, 100402 (2003).

6.48 Guifré Vidal, Daniel Jonathan, and MA Nielsen, "Approximate transformations and robust manipulation of bipartite pure-state entanglement," Phys. Rev. A **62**, 012304 (2000).

6.49 Marco Tomamichel, *Quantum information processing with finite resources: mathematical foundations*, Vol. 5 (Springer, 2015).

6.50 Johan Åberg, "Truly work-like work extraction via a single-shot analysis," Nat. Commun. **4**, 1925 (2013).

6.51 Dario Egloff, Oscar CO Dahlsten, Renato Renner, and Vlatko Vedral, "A measure of majorization emerging from single-shot statistical mechanics," New J. Phys. **17**, 073001 (2015).

6.52 Chris Perry, Piotr Ćwikliński, Janet Anders, Michał Horodecki, and Jonathan Oppenheim, "A sufficient set of experimentally implementable thermal operations," arXiv:1511.06553 (2015).

6.53 Michele Campisi, Peter Talkner, and Peter Hänggi, "Finite bath fluctuation theorem," Phys. Rev. E **80**, 031145 (2009).

6.54 David Reeb and Michael M Wolf, "An improved Landauer principle with finite-size corrections," New J. Phys. **16**, 103011 (2014).

6.55 Jonathan G Richens, Álvaro M Alhambra, and Lluis Masanes, "Finite-bath corrections to the second law of thermodynamics," Phys. Rev. E **97**, 062132 (2018).

6.56 Jakob Scharlau and Markus P Mueller, "Quantum Horn's lemma, finite heat baths, and the third law of thermodynamics," Quantum **2**, 54 (2018).

6.57 Hiroyasu Tajima and Masahito Hayashi, "Finite-size effect on optimal efficiency of heat engines," Phys. Rev. E **96**, 012128 (2017).

6.58 Kosuke Ito and Masahito Hayashi, "Optimal performance of generalized heat engines with finite-size baths of arbitrary multiple conserved quantities based on non-iid scaling," arXiv:1612.04047 (2016).

6.59 Mischa P Woods, Nelly Ng, and Stephanie Wehner, "The maximum efficiency of nano heat engines depends on more than temperature," arXiv:1506.02322 (2015).

6.60 P. Faist, J. Oppenheim, and R. Renner, "Gibbs-preserving maps outperform thermal operations in the quantum regime," New J. of Phys. **17**, 043003 (2015).

6.61 Kamil Korzekwa, "Structure of the thermodynamic arrow of time in classical and quantum theories," Phys. Rev. A **95**, 052318 (2017).

6.62 Masahito Hayashi, *Quantum Information — An Introduction* (Springer, 2006).

6.63 C.T. Chubb, V.Y.F. Tan, and M. Tomamichel, "Moderate Deviation Analysis for Classical Communication over Quantum Channels," Commun. Math. Phys. **355** (2017).

6.64 Hao-Chung Cheng and Min-Hsiu Hsieh, "Moderate Deviation Analysis for Classical-Quantum Channels and Quantum Hypothesis Testing," (2017), arXiv:1701.03195 .

Chapter 7

Moderate deviation analysis of majorisation-based resource interconversion

Christopher T. Chubb[1], Marco Tomamichel[2], and Kamil Korzekwa[1]

[1] *Centre for Engineered Quantum Systems, School of Physics, University of Sydney, Sydney NSW 2006, Australia*

[2] *Centre for Quantum Software and Information, School of Software, University of Technology Sydney, Sydney NSW 2007, Australia*

ArXiv: 1809.07778

Abstract

We consider the problem of interconverting a finite amount of resources within all theories whose single-shot transformation rules are based on a majorisation relation, e.g. the resource theories of entanglement and coherence (for pure state transformations), as well as thermodynamics (for energy-incoherent transformations). When only finite resources are available we expect to see a non-trivial trade-off between the rate r_n at which n copies of a resource state ρ can be transformed into nr_n copies of another resource state σ, and the error level ϵ_n of the interconversion process, as a function of n. In this work we derive the optimal trade-off in the so-called moderate deviation regime, where the rate of interconversion r_n approaches its optimum in the asymptotic limit of unbounded resources ($n \to \infty$), while the error ϵ_n vanishes in the same limit. We find that the moderate deviation analysis exhibits a resonance behaviour which implies that certain pairs of resource states can be interconverted at the asymptotically optimal rate with negligible error, even in the finite n regime.

7.1 Introduction

In principle, while processing quantum information, any initial state can be transformed into any final state. One could thus conclude that all quantum states are equally valuable or resourceful. In reality, however, some transformations are harder to implement than others, which results in a partial ordering

of the set of quantum states, with the hardest to prepare at the top, and easiest at the bottom. Such a resource hierarchy arises naturally when we face any kind of restrictions: from the locality constraint, through experimental difficulties in preparing particular superpositions, to fundamental constraints induced by physical laws like energy conservation. The mathematical framework developed to study possible state transformations under such restrictions is known under the collective name of resource theories [7.1, 7.2].

Inspired by classical information theory, the early resource-theoretic works considered optimal conversion rates between different resource states in the asymptotic regime, i.e., the limit of processing infinitely many copies of a given state. This led to the discovery of asymptotic resource measures, which provided operational meaning to quantities such as entropy of entanglement [7.3] and non-equilibrium free energy [7.4]. Namely, a given transformation becomes asymptotically possible if and only if the corresponding asymptotic resource measure is non-increasing, which allows one to reversibly interconvert between all resource states.

On the other hand, almost simultaneously to the asymptotic studies, the single-shot regime was investigated, where one aims at deciding whether it is possible to convert a single copy of an initial state into the final state. Here, probably the most famous contributions are the Nielsen's theorem [7.5] within the resource theory of entanglement, and more recently the family of second laws for the resource theory of thermodynamics [7.6]. In general, in the single-shot regime simple asymptotic transformation rules get replaced by more complex sets of conditions, which also give rise to irreversible transformation.

In this paper we focus on the interconversion process in the intermediate regime, when the number of processed resource states is large, but finite. This way we aim at keeping the simplicity of the asymptotic analysis, but also at preserving the irreversible nature of single-shot regime. The first steps in this direction were recently made in Refs. [7.7] and [7.8] for the resource theories of entanglement and thermodynamics, where the corrections to asymptotic conversion rates were found in the scenario with a constant transformation error (i.e., in the small deviation regime [7.9]). Here, we present a moderate deviation analysis [7.10] (see also [7.11, 7.12] for applications in the quantum domain) of the interconversion problem within a unified framework that includes all resource theories for which the single-shot transformation rules can be expressed via majorisation or thermo-majorisation. This way we find finite-size corrections to conversion rates in resource theories of entanglement [7.13], coherence [7.14] and thermodynamics [7.15], in the regime where the transformation error, measured by either infidelity or total variation distance, asymptotically vanishes.

Our results can be directly applied to the study of important problems such as entanglement distillation [7.16] or coherence dilution [7.17], but also allow one for a rigorous analysis of the irreversibility arising when finite-size resources are interconverted. Most intriguingly, we find that if a pair of states satisfies a particular resonance condition, one can achieve lossless interconversion, i.e., transformation that is arbitrarily close to reversible even for finite n. In the accompanying paper [7.18] we discuss how this effect can be employed to avoid irreversibility, which directly affects, e.g., the performance of heat engines working with finite-size working bodies [7.19].

This paper is structured in the following way. First, in Sec. 7.2, we set the scene by introducing necessary tools and concepts. Next, in Sec. 7.3 we state our main result concerning moderate deviation corrections to the asymptotic interconversion rates for majorisation-based resource theories. We then proceed to Sec. 7.4 that contains auxiliary technical results concerning tail bounds, which are used in the formal proof that can be found in Sec. 7.5. Finally, we provide conclusions and outlook in Sec. 7.6.

7.2 Setting the scene

7.2.1 Resource theories in different regimes

Every quantum resource theory [7.1, 7.2] is defined by a set of quantum operations that are considered free, and a set of restrictions that make other operations impossible without an additional cost. Such restrictions may arise from practical difficulties, e.g., when preparing a system in a superposition of particular states is experimentally challenging, but may also be of fundamental nature, as with the laws of thermodynamics constraining possible transformations to preserve energy and increase entropy. A resource is then defined as a quantum system that allows one to lift a given restriction. Typical examples of resources include an excited pure state that acts as a work storage, and thus can be used to reduce the entropy of another system (overcoming thermodynamic constraints [7.15]); an entangled Bell pair, which can be used to teleport a quantum state (overcoming locality constraints [7.13]); or a system in the superposition of energy eigenstates, which can be used as a reference frame for time (overcoming symmetry constraints [7.20]).

Once the restrictions and the corresponding resources are defined, the central question concerns resource interconversion, i.e., what final states can be obtained from the initial one. This problem so far was mainly approached in either the *single-shot regime* [7.21], or in an idealised *asymptotic limit* [7.22]. The first approach, due to its generality and the corresponding complexity of the answer, provides only a limited insight into the nature of different resource states. The second one provides an elegant and simple answer in the form of optimal *conversion rate*, which tells us how many copies of the final state can be obtained per one copy of the initial state, if one assumes access to an infinite source of initial states. From a practical point of view, however, such an assumption is unjustified, as most quantum resources will be available only in small amounts in the foreseeable future. More fundamentally, finite-size effects may be of interest themselves, as it is the case within quantum thermodynamics [7.23], where one aims at accurate description of heat and work processes involving small number of particles.

Very recently the first steps have been made to study the intermediate regime, where one focuses on the interconversion of large but finite number n of resource states. First, in Ref. [7.7] the authors focused on transformations within the resource theory of entanglement. Their results were then generalised and adapted to the studies of the interconversion process in the resource theory of thermodynamics by the present authors [7.8]. In both these works the second-order correction to the asymptotic rate was found in the so-called *small deviation* regime [7.9], where the conversion rate approaches the asymptotic one for $n \to \infty$, but the transformation is realised with a constant error. In the current work we solve the issue of constant error by deriving corrections to the asymptotic rate in the *moderate deviation* regime [7.10], where the correction term still vanishes as $n \to \infty$, but also the transformation is asymptotically error-free. For the completeness of discussion, we also note that the interconversion problem may be studied in the *large deviation* regime [7.24], where the error is exponentially vanishing for the price of the constant gap between the realised conversion rate and the asymptotic one. In Table 7.1 we collect references to central results concerning state interconversion within the investigated resource theories in various regimes.

7.2.2 Exact single-shot interconversion

Irrespective of the investigated regime, the first step is to find single-shot interconversion rules, which form the basis of further analysis. In this work we study the interconversion problem within all

	Ent.	Coh.	Thermo.
$n = 1,\ \epsilon = 0$	[7.5]	[7.25]	[7.15]
$n \to \infty,\ \epsilon \to 0$	[7.3]		[7.4]
$n < \infty,\ \epsilon > 0$	[7.7]		[7.8]
$n < \infty,\ \epsilon \to 0$	This work		

Table 7.1: *Interconversion in various regimes.* Exposition of works on state interconversion within resource theories of ent(anglement), coh(erence) and thermo(dynamics) in single-shot regime ($n = 1, \epsilon = 0$), asymptotic limit ($n \to \infty, \epsilon \to 0$), small deviation regime ($n < \infty, \epsilon > 0$) and moderate deviation regime ($n < \infty, \epsilon \to 0$).

majorisation-based resource theories, i.e., when conditions for single-shot transformations can be expressed as majorisation partial order [7.26], or a variant known as thermo-majorisation [7.15, 7.27]. Within such theories, each resource state can be represented by a probability distribution, and the conversion process is possible when the distribution representing the initial state majorises (or is majorised) by the distribution representing the final state, with majorisation $\succ$ defined by

$$\boldsymbol{a} \succ \boldsymbol{b} \quad \Longleftrightarrow \quad \forall j: \ \sum_{i=1}^{j} a_i^{\downarrow} \geq \sum_{i=1}^{j} b_i^{\downarrow}, \tag{7.1}$$

where $\boldsymbol{a}^{\downarrow}$ denotes the vector $\boldsymbol{a}$ in a decreasing order.

Three prominent examples of majorisation-based resource theories include the resource theories of entanglement, coherence and thermodynamics. These are defined via the relevant sets of free operations and free states: Local Operations and Classical Communication (LOCC) and separable states in entanglement theory [7.13]; Incoherent Operations and incoherent states in coherence theory [7.14]; Thermal Operations and the thermal equilibrium state γ in the resource theory of thermodynamics (with respect to a fixed background temperature $T = 1/\beta$) [7.28]. As mentioned above, within each of these theories there exists a representation of initial and target quantum states, ρ and σ, as probability distributions $\boldsymbol{p}$ and $\boldsymbol{q}$. For entanglement theory, given initial and target pure bipartite states, $\rho = |\Psi\rangle\langle\Psi|$ and $\sigma = |\Phi\rangle\langle\Phi|$, with the Schmidt decomposition given by

$$|\Psi\rangle = \sum_i a_i |\psi_i \psi_i\rangle, \quad |\Phi\rangle = \sum_i b_i |\phi_i \phi_i\rangle, \tag{7.2}$$

we can represent them via probability distributions

$$p_i = |a_i|^2, \quad q_i = |b_i|^2. \tag{7.3}$$

For coherence theory, with respect to a fixed basis $\{|i\rangle\}$, one can represent pure initial and target states, $\rho = |\psi\rangle\langle\psi|$ and $\sigma = |\phi\rangle\langle\phi|$, using

$$p_i = |\langle i|\psi\rangle|^2, \quad q_i = |\langle i|\phi\rangle|^2. \tag{7.4}$$

Finally, in the resource theory of thermodynamics, the initial and target energy-incoherent mixed states ρ and σ can be represented by

$$p_i = \langle E_i| \rho |E_i\rangle, \quad q_i = \langle E_i| \sigma |E_i\rangle, \tag{7.5}$$

where $\{|E_i\rangle\}$ denotes the energy eigenbasis of the system. We will denote distributions representing free states by $\boldsymbol{f}$. In entanglement and coherence theories these are represented by sharp probability distributions $\boldsymbol{s}$ with a single non-zero entry; whereas in the thermodynamic case $\boldsymbol{f}$ is given by a thermal Gibbs distribution $\boldsymbol{\gamma}$ with $\gamma_i \propto \exp(-\beta E_i)$.

The celebrated Nielsen's theorem [7.5] (for entanglement) and the recent result of Ref. [7.25] (for coherence) state that the initial state represented by $\boldsymbol{p}$ can be transformed into the target state represented by $\boldsymbol{q}$ if and only if $\boldsymbol{p} \prec \boldsymbol{q}$. Similarly, in Ref. [7.15], it was found that a thermodynamic transformations between states represented by $\boldsymbol{p}$ and $\boldsymbol{q}$ is possible if and only if $\hat{\boldsymbol{p}} \succ \hat{\boldsymbol{q}}$, where $\hat{\boldsymbol{a}}$ can be obtained from $\boldsymbol{a}$ via a straightforward application of an *embedding map* Γ^β [7.6, 7.29]. For the sake of our analysis, it is only crucial to note that Γ^β maps d-dimensional distributions to $\hat{d}$-dimensional ones with $\hat{d} \geq d$; and that an embedded version of the free thermal distribution is given by a maximally mixed distribution on a larger subspace, i.e., $\hat{\boldsymbol{\gamma}} = \boldsymbol{\eta}$ with $\boldsymbol{\eta} = [1/\hat{d}, \ldots, 1/\hat{d}]$.

7.2.3 Approximate multi-copy interconversion

When considering transformations between many copies of initial and target states, represented by $\boldsymbol{p}^{\otimes n}$ and $\boldsymbol{q}^{\otimes m}$, we need to make sure that the dimensionality of the input and output spaces match. Since one can always append any number of free states $\boldsymbol{f}$ to both the initial and target states, we introduce total initial and target distributions,

$$\boldsymbol{P}^{n,m} := \boldsymbol{p}^{\otimes n} \otimes \boldsymbol{f}^{\otimes m}, \quad \boldsymbol{Q}^{n,m} := \boldsymbol{q}^{\otimes m} \otimes \boldsymbol{f}^{\otimes n}. \tag{7.6}$$

Our main object of interest will be the *conversion rate* $r_n := m/n$, i.e., the number of target states one can obtain per one copy of the initial state. For notational clarity we will denote total initial and target distributions by $\boldsymbol{P}^n$ and $\boldsymbol{Q}^n$, with the dependence on m (so, in fact, on r_n) kept implicit. The single-shot interconversion conditions can now be expressed as $\boldsymbol{P}^n \prec \boldsymbol{Q}^n$ for the entanglement and coherence transformations, and $\hat{\boldsymbol{P}}^n \succ \hat{\boldsymbol{Q}}^n$ for the thermodynamic transformations.

We also need to introduce the concept of approximate interconversion. Assume that for given $\boldsymbol{P}^n$ and $\boldsymbol{Q}^n$ the relevant majorisation relation does not hold, so that the interconversion is impossible. However, there may exist $\tilde{\boldsymbol{Q}}^n$ that is ϵ-close to $\boldsymbol{Q}^n$ and such that the interconversion is possible. We then say that an approximate transformation is possible with the error level ϵ quantified by either the infidelity, $1 - F$, or total variation distance (TVD), δ, between target and final states, with

$$F(\boldsymbol{Q}^n, \tilde{\boldsymbol{Q}}^n) := \left(\sum_i \sqrt{Q_i^n \tilde{Q}_i^n} \right)^2, \tag{7.7a}$$

$$\delta(\boldsymbol{Q}^n, \tilde{\boldsymbol{Q}}^n) := \frac{1}{2} \sum_i \left| Q_i^n - \tilde{Q}_i^n \right|. \tag{7.7b}$$

The concept of approximate interconversion gives rise to two notions of approximate majorisation introduced in Ref. [7.8], ϵ-post-majorisation $\succ_\epsilon$ and ϵ-pre-majorisation ${}_\epsilon\succ$, defined by

$$\boldsymbol{a} \succ_\epsilon \boldsymbol{b} \quad \Longleftrightarrow \quad \exists\, \tilde{\boldsymbol{b}} : \ \boldsymbol{a} \succ \tilde{\boldsymbol{b}} \text{ and } \delta(\boldsymbol{b}, \tilde{\boldsymbol{b}}) \leq \epsilon, \tag{7.8a}$$

$$\boldsymbol{a} \ {}_\epsilon\succ \boldsymbol{b} \quad \Longleftrightarrow \quad \exists\, \tilde{\boldsymbol{a}} : \ \tilde{\boldsymbol{a}} \succ \boldsymbol{b} \text{ and } \delta(\boldsymbol{a}, \tilde{\boldsymbol{a}}) \leq \epsilon, \tag{7.8b}$$

where, depending on the context, δ can be replaced by $1 - F$. Crucially, in Ref. [7.8] the present authors showed that these two notions are equivalent and, moreover, that ϵ-post-majorisation between embedded vectors, $\hat{\boldsymbol{a}} \succ_\epsilon \hat{\boldsymbol{b}}$, is a necessary and sufficient condition for the existence of an approximate thermodynamic transformation between $\boldsymbol{a}$ and $\boldsymbol{b}$ with error level ϵ.

We conclude that an approximate transformation between initial and target states, represented by $\boldsymbol{p}^{\otimes n}$ and $\boldsymbol{q}^{\otimes n r_n}$, is possible within resource theories of entanglement and coherence if and only if

$$\boldsymbol{P}^n \prec_\epsilon \boldsymbol{Q}^n, \tag{7.9}$$

with the free state $\boldsymbol{f} = \boldsymbol{s}$. We will refer to the above relation as the approximate majorisation relation for the *entanglement direction*. Similarly, such a transformation is possible within resource theory of thermodynamics if and only if

$$\hat{\boldsymbol{P}}^n \succ_\epsilon \hat{\boldsymbol{Q}}^n, \tag{7.10}$$

with the free state $\boldsymbol{f} = \boldsymbol{\gamma}$. We will refer to this relation as the approximate majorisation relation for the *thermodynamic direction*.

7.2.4 Information-theoretic notions

The main role in the quantitative analysis of the interconversion process for the entanglement direction will be played by the Shannon entropy H and entropy variance V. For a given probability distribution $\boldsymbol{a}$ these are defined by

$$H(\boldsymbol{a}) = -\sum_i a_i \ln a_i, \tag{7.11a}$$

$$V(\boldsymbol{a}) = \sum_i a_i \left[\ln a_i + H(\boldsymbol{a})\right]^2. \tag{7.11b}$$

The analogous role for the thermodynamic direction will be played by the relative entropy D and relative entropy variance V. Given two probability distributions, $\boldsymbol{a}$ and $\boldsymbol{b}$, these are defined by

$$D(\boldsymbol{a}||\boldsymbol{b}) = \sum_i a_i \ln \frac{a_i}{b_i}, \tag{7.12a}$$

$$V(\boldsymbol{a}||\boldsymbol{b}) = \sum_i a_i \left[\ln \frac{a_i}{b_i} - D(\boldsymbol{a}||\boldsymbol{b})\right]^2. \tag{7.12b}$$

An important fact, that can be verified by direct calculation, is that the relative quantities are invariant under embedding, i.e., $D(\boldsymbol{a}||\boldsymbol{b}) = D(\hat{\boldsymbol{a}}||\hat{\boldsymbol{b}})$ and $V(\boldsymbol{a}||\boldsymbol{b}) = V(\hat{\boldsymbol{a}}||\hat{\boldsymbol{b}})$ [7.8].

In order to formally state our main result we also need to introduce the notion of a *moderate sequence*:

Definition 7.1 (Moderate sequence)**.** *A sequence of real numbers $\{t_n\}_n$ is a* moderate sequence *if its scaling is strictly between $1/\sqrt{n}$ and 1, meaning that $t_n \to 0$ and $\sqrt{n}t_n \to +\infty$ as $n \to \infty$.*

Note that an important family of moderate sequences is given by $t_n \sim n^{-\alpha}$ for $\alpha \in (0, 1/2)$, which can be used to obtain a particularly simple version of our main results.

Finally, as we will be interested in asymptotic expansions in n, we will employ the standard asymptotic notation: $o(f(n))$, $O(f(n))$ and $\Theta(f(n))$. We will also use $\overset{\text{ev.}}{>}$ and $\overset{\text{ev.}}{<}$ to denote eventual inequalities, specifically we write $a_n \overset{\text{ev.}}{>} b_n$ if and only if there exists N such that $a_n > b_n$ for all $n \geq N$. Moreover, we will denote equalities and inequalities up to terms of order $o(t_n)$ by $\simeq$, $\lesssim$ and $\gtrsim$.

7.3 Interconversion rates beyond the asymptotic regime

We are now ready to state our central technical result, which may be of interest outside the resource-theoretic studies due to ubiquity of majorisation partial order in the broad field of applied mathematics [7.26]. We split it into three theorems. The first two concern state interconversion below the asymptotic rate and with asymptotically vanishing error (one for each majorisation direction). The third one concerns practically less relevant scenario of state interconversion above the asymptotic rate and with error asymptotically approaching 1.

For the entanglement direction we introduce the *optimal conversion rate* $R_n^{\text{ent}}(\epsilon)$ as the largest conversion rate r_n for which the approximate majorisation relation for the entanglement direction, $\boldsymbol{P}^n \prec_\epsilon \boldsymbol{Q}^n$, holds. Due to the discussion presented in Sec. 7.2, $R_n^{\text{ent}}(\epsilon)$ is the maximal rate for which the approximate interconversion, with error ϵ, is possible between states represented by $\boldsymbol{p}$ and $\boldsymbol{q}$ within resource theories of entanglement and coherence. We also define the *asymptotic rate*,

$$R_\infty^{\text{ent}} = \frac{H(\boldsymbol{p})}{H(\boldsymbol{q})}, \tag{7.13}$$

and the *irreversibility parameter*,

$$\nu^{\text{ent}} = \frac{V(\boldsymbol{p})/H(\boldsymbol{p})}{V(\boldsymbol{q})/H(\boldsymbol{q})}. \tag{7.14}$$

We then have:

Theorem 7.1 (Entanglement direction)**.** *For any moderate sequence t_n and the accepted error level of*

$$\epsilon_n = e^{-nt_n^2}, \tag{7.15}$$

the asymptotic expansion of the optimal conversion rate $R_n^{\text{ent}}(\epsilon_n)$ is

$$R_n^{\text{ent}}(\epsilon_n) \simeq R_\infty^{\text{ent}} - \sqrt{\frac{2V(\boldsymbol{p})}{H(\boldsymbol{q})^2}}\left|1 - 1/\sqrt{\nu^{\text{ent}}}\right| t_n. \tag{7.16}$$

Analogously, for the thermodynamic direction we introduce the *optimal conversion rate* $R_n^{\text{th}}(\epsilon)$ as the largest conversion rate r_n for which the approximate majorisation relation for the thermodynamic direction, $\hat{\boldsymbol{P}}^n \succ_\epsilon \hat{\boldsymbol{Q}}^n$, holds. As before, $R_n^{\text{th}}(\epsilon)$ is the maximal rate for which the approximate interconversion, with error ϵ, is possible between states represented by $\boldsymbol{p}$ and $\boldsymbol{q}$ within the resource theory of thermodynamics. We also define the *asymptotic rate*,

$$R_\infty^{\text{th}} = \frac{D(\boldsymbol{p}||\boldsymbol{\gamma})}{D(\boldsymbol{q}||\boldsymbol{\gamma})}, \tag{7.17}$$

and the *irreversibility parameter*,

$$\nu^{\text{th}} = \frac{V(\boldsymbol{p}||\boldsymbol{\gamma})/D(\boldsymbol{p}||\boldsymbol{\gamma})}{V(\boldsymbol{q}||\boldsymbol{\gamma})/D(\boldsymbol{q}||\boldsymbol{\gamma})}. \tag{7.18}$$

We then have:

Theorem 7.2 (Thermodynamic direction)**.** *For any moderate sequence t_n and the accepted error level of*

$$\epsilon_n = e^{-nt_n^2}, \tag{7.19}$$

the asymptotic expansion of the optimal conversion rate $R_n^{\text{th}}(\epsilon_n)$ is

$$R_n^{\text{th}}(\epsilon_n) \simeq R_\infty^{\text{th}} - \sqrt{\frac{2V(\boldsymbol{p}||\boldsymbol{\gamma})}{D(\boldsymbol{q}||\boldsymbol{\gamma})^2}}\left|1 - 1/\sqrt{\nu^{\text{th}}}\right| t_n. \tag{7.20}$$

Finally, one expects that conversion above the asymptotic rate leads to transformation error approaching 1. This is formalised in the following theorem which, unlike the previous two theorems (that hold for the error level measured by both infidelity and total variation distance), applies only to TVD. In Appendix 7.A, where we relate our current results to the small deviation analysis of Refs. [7.7, 7.8], we also conjecture the analogue of Theorem 7.3 with the error measured by infidelity.

Theorem 7.3 (Converse regime)**.** *For any moderate sequence t_n and the accepted TVD error of*

$$\epsilon_n = 1 - e^{-nt_n^2}, \tag{7.21}$$

the asymptotic expansion of the optimal conversion rate $R_n^{\text{ent}}(\epsilon_n)$ is

$$R_n^{\text{ent}}(\epsilon_n) \simeq R_\infty^{\text{ent}} + \sqrt{\frac{2V(\boldsymbol{p})}{H(\boldsymbol{q})^2}}\left(1 + 1/\sqrt{\nu^{\text{ent}}}\right) t_n, \tag{7.22a}$$

and similarly for $R_n^{\text{th}}(\epsilon_n)$ we have

$$R_n^{\text{th}}(\epsilon_n) \simeq R_\infty^{\text{th}} + \sqrt{\frac{2V(\boldsymbol{p}||\boldsymbol{\gamma})}{D(\boldsymbol{q}||\boldsymbol{\gamma})^2}}\left(1 + 1/\sqrt{\nu^{\text{th}}}\right) t_n. \tag{7.22b}$$

We present the proofs in Sec. 7.5, after we introduce the necessary tools in Sec. 7.4. Before that let us make two important remarks.

Remark 1. *For initial and target states satisfying $\nu^{\text{ent}} = 1$, the optimal conversion rate R_n^{ent} in the regime of vanishing error is given by the asymptotic rate R_∞^{ent}. This means that, up to terms of order $o(t_n)$, such a transformation is reversible even for finite n. Analogous observation holds for the thermodynamic direction. We discuss the implications of this particularly interesting scenario in an accompanying paper [7.18].*

Remark 2. *When $V(\boldsymbol{p}) = 0$, resulting in $1/\sqrt{\nu^{\text{ent}}}$ diverging to infinity and the apparent multiplication of zero times infinity, one can simply use the definition of ν^{ent} to replace Eq. (7.16) with*

$$R_n^{\text{ent}}(\epsilon_n) \simeq R_\infty^{\text{ent}} \pm \sqrt{\frac{2V(\boldsymbol{q})H(\boldsymbol{p})}{H(\boldsymbol{q})^3}} t_n. \tag{7.23}$$

Analogous observation holds for the thermodynamic direction.

7.4 Moderate deviation toolkit

7.4.1 Preliminaries

The central result of the moderate deviation analysis can be stated as follows.

Lemma 7.1 (Moderate deviation bound)**.** *Let $\{X_i\}_{1\leq i\leq n}$ be independent and identically distributed (i.i.d.) random variables with zero-mean and variance v. For any moderate sequence $\{t_n\}_n$ the following hold:*

$$\lim_{n\to\infty}\frac{1}{nt_n^2}\ln\left[\Pr\left(\frac{1}{n}\sum_{i=1}^{n}X_i\geq t_n\right)\right]=-\frac{1}{2v}, \tag{7.24a}$$

$$\lim_{n\to\infty}\frac{1}{nt_n^2}\ln\left[\Pr\left(\frac{1}{n}\sum_{i=1}^{n}X_i\leq -t_n\right)\right]=-\frac{1}{2v}. \tag{7.24b}$$

The proof of the above lemma can be found, e.g., in Appendix A of Ref. [7.11]. For the remainder of the paper, consider $\{t_n\}_n$ to be a fixed moderate sequence. For clarity we will henceforth omit the dependence of all implicit constants on this sequence. It should be noted that the above lemma also holds when $v=0$, where we henceforth adopt the convention that $1/v=+\infty$ in this case.

7.4.2 Two variations on tail bounds

We now want to adapt Lemma 7.1 to our purposes of majorisation-based analysis. For a probability vector $\boldsymbol{a}$ we thus introduce the following quantity

$$k_n(\boldsymbol{a},x):=\exp\left(H(\boldsymbol{a}^{\otimes n})+xnt_n\right), \tag{7.25}$$

which allows us to formulate the magnitude-based version of the moderate deviation bound for products of distributions.

Lemma 7.2 (Magnitude-based tail bound)**.** *Consider an arbitrary probability distribution $\boldsymbol{a}$. For $x\leq 0$ we have*

$$\lim_{n\to\infty}\frac{1}{nt_n^2}\ln\sum_i\left\{\left(\boldsymbol{a}^{\otimes n}\right)_i\middle|\left(\boldsymbol{a}^{\otimes n}\right)_i\geq\frac{1}{k_n(\boldsymbol{a},x)}\right\}=\frac{-x^2}{2V(\boldsymbol{a})}, \tag{7.26a}$$

and similarly for $x\geq 0$ we have

$$\lim_{n\to\infty}\frac{1}{nt_n^2}\ln\sum_i\left\{\left(\boldsymbol{a}^{\otimes n}\right)_i\middle|\left(\boldsymbol{a}^{\otimes n}\right)_i\leq\frac{1}{k_n(\boldsymbol{a},x)}\right\}=\frac{-x^2}{2V(\boldsymbol{a})}. \tag{7.26b}$$

Proof. Consider the random variable $L:=-\log a$, distributed according to $\boldsymbol{a}$, such that the expectation value $\langle L\rangle$ and the variance $\mathrm{Var}(L)$ are equal to $H(\boldsymbol{a})$ and $V(\boldsymbol{a})$ respectively. We can express k_n in terms of L as

$$\log k_n(\boldsymbol{a},x)=n\langle L\rangle_{\boldsymbol{a}}+xnt_n. \tag{7.27}$$

If we let $\{L_j\}_{1\leq j\leq n}$ be i.i.d. copies of L, then we can write the tail bound of $\boldsymbol{a}^{\otimes n}$ in terms of tail bounds on the average of these variables,

$$\begin{aligned}\sum_i \left\{ \left(\boldsymbol{a}^{\otimes n}\right)_i \middle| \left(\boldsymbol{a}^{\otimes n}\right)_i \geq \frac{1}{k_n(\boldsymbol{a},x)} \right\} &= \sum_{i_1,\ldots,i_n} \left\{ \prod_{j=1}^n a_{ij} \middle| \prod_{j=1}^n a_{i_j} \geq \frac{1}{k_n(\boldsymbol{a},x)} \right\} \\ &= \sum_{i_1,\ldots,i_n} \left\{ \prod_{j=1}^n a_{ij} \middle| \sum_{j=1}^n \log a_{i_j} \geq \log \frac{1}{k_n(\boldsymbol{a},x)} \right\} \\ &= \Pr\left[\sum_{j=1}^n L_j \leq n\langle L\rangle + xnt_n \right]. \end{aligned} \tag{7.28}$$

For $x<0$, we can now apply Lemma 7.2 to the variables $X_j := (L_j - \langle L\rangle)/x$ to obtain Eq. (7.26a). An analogous argument can be employed for $x>0$, with all of the above inequalities reversed, yielding Eq. (7.26b). Finally, for $x=0$ case, we can appeal to the Central Limit Theorem, which gives

$$\sum_i \left\{ \left(\boldsymbol{a}^{\otimes n}\right)_i \,\middle|\, \left(\boldsymbol{a}^{\otimes n}\right)_i \geq \frac{1}{k_n(\boldsymbol{a},0)} \right\} = \Pr\left[\frac{1}{n}\sum_{j=1}^n L_j \leq \langle L\rangle \right] \xrightarrow{n\to\infty} \frac{1}{2}, \tag{7.29}$$

implying Eqs. (7.26a)-(7.26b). □

Using the above result we can now prove the majorisation-based version of the moderate deviation bound.

Lemma 7.3 (Majorisation-based tail bound)**.** *Consider an arbitrary probability distribution $\boldsymbol{a}$ satisfying $V(\boldsymbol{a})>0$. For $x\leq 0$ we have*

$$\lim_{n\to\infty} \frac{1}{nt_n^2} \ln\left[\sum_{i\leq k_n(\boldsymbol{a},x)} \left(\boldsymbol{a}^{\otimes n}\right)_i^{\downarrow} \right] = -\frac{x^2}{2V(\boldsymbol{a})}, \tag{7.30a}$$

and similarly for $x\geq 0$ we have

$$\lim_{n\to\infty} \frac{1}{nt_n^2} \ln\left[\sum_{i\geq k_n(\boldsymbol{a},x)} \left(\boldsymbol{a}^{\otimes n}\right)_i^{\downarrow} \right] = -\frac{x^2}{2V(\boldsymbol{a})}. \tag{7.30b}$$

Proof. Here we follow the proof of the small-deviation analogue of this result, Lemmas 15 and 16 of Ref. [7.7]. Consider first the $x\leq 0$ case, and define two sets of indices

$$S_n(x) := \{1,\ldots,\lfloor k_n(\boldsymbol{a},x)\rfloor\}, \tag{7.31a}$$

$$\tilde{S}_n(x) := \left\{ i \,\middle|\, \left(\boldsymbol{a}^{\otimes n}\right)_i^{\downarrow} \geq 1/k_n(\boldsymbol{a},x) \right\}. \tag{7.31b}$$

We note that Lemma 7.2 gives that

$$\lim_{n\to\infty} \frac{1}{nt_n^2} \ln \left[\sum_{i\in\tilde{S}(x)} \left(\boldsymbol{a}^{\otimes n}\right)_i^{\downarrow} \right] = -\frac{x^2}{2V(\boldsymbol{a})} \tag{7.32}$$

for any $x \leq 0$, and we wish to show an analogous result for $S_n(x)$. We will achieve this by showing, for any $\delta > 0$, that $\tilde{S}_n(x) \subseteq S_n(x) \subseteq \tilde{S}_n(x+\delta)$ holds eventually, i.e., for large enough n. The first inclusion follows trivially from the normalisation of our distribution, and so it is left only to show that $\tilde{S}_n(x) \subseteq S_n(x+\delta)$.

Noting that $(\boldsymbol{a}^{\otimes n})_i^{\downarrow} - 1/k_n(\boldsymbol{a}, x+\delta/2) \geq 0$ if and only if $i \in \tilde{S}_n(x+\delta/2)$, we see that

$$\begin{aligned} &\sum_{i\in\tilde{S}_n(x+\delta/2)} \left[(\boldsymbol{a}^{\otimes n})_i^{\downarrow} - \frac{1}{k_n(\boldsymbol{a}, x+\delta/2)} \right] \\ &\qquad \geq \sum_{i\in T} \left[(\boldsymbol{a}^{\otimes n})_i^{\downarrow} - \frac{1}{k_n(\boldsymbol{a}, x+\delta/2)} \right], \end{aligned} \tag{7.33}$$

for any set of indices T. Taking $T = \tilde{S}_n(x+\delta)$, this gives

$$\frac{\left|\tilde{S}_n(x+\delta) \setminus \tilde{S}_n(x+\delta/2)\right|}{k_n(\boldsymbol{a}, x+\delta/2)} \geq \sum_{i\in\tilde{S}_n(x+\delta)\setminus\tilde{S}_n(x+\delta/2)} \left(\boldsymbol{a}^{\otimes n}\right)_i^{\downarrow}. \tag{7.34}$$

Lemma 7.2 tells us that the summation on the RHS scales as $e^{-\Theta(nt_n^2)}$, specifically that there is a lower bound of the form $e^{-Cnt_n^2}$ for some constant C. Since $t_n \to 0$, we eventually have that $Ct_n < \delta/2$, and so this sum can be lower bounded as follows

$$\sum_{i\in\tilde{S}_n(x+\delta)\setminus\tilde{S}_n(x+\delta/2)} \left(\boldsymbol{a}^{\otimes n}\right)_i^{\downarrow} \overset{\text{ev.}}{>} e^{-\delta n t_n/2}. \tag{7.35}$$

Applying this bound to Eq. (7.34) allows us to conclude

$$\begin{aligned} \left|\tilde{S}_n(x+\delta)\right| &\geq \left|\tilde{S}_n(x+\delta) \setminus \tilde{S}_n(x+\delta/2)\right| \\ &\overset{\text{ev.}}{>} e^{-\delta n t_n} k_n(\boldsymbol{a}, x+\delta/2), \\ &= k_n(\boldsymbol{a}, x), \end{aligned} \tag{7.36}$$

and therefore that $S_n(x) \subseteq \tilde{S}_n(x+\delta)$ as required.

The inclusions $\tilde{S}_n(x) \subseteq S_n(x) \subseteq \tilde{S}_n(x+\delta)$, together with Lemma 7.2, give us the following inequalities

$$\liminf_{n\to\infty} \frac{1}{nt_n^2} \ln \left[\sum_{i\in S_n(x)} \left(\boldsymbol{a}^{\otimes n}\right)_i^{\downarrow} \right] \geq -\frac{(x+\delta)^2}{2V(\boldsymbol{a})}, \tag{7.37a}$$

$$\limsup_{n\to\infty} \frac{1}{nt_n^2} \ln \left[\sum_{i\in S_n(x)} \left(\boldsymbol{a}^{\otimes n}\right)_i^{\downarrow} \right] \leq -\frac{x^2}{2V(\boldsymbol{a})}. \tag{7.37b}$$

As this holds for any $\delta > 0$, we conclude that

$$\lim_{n\to\infty} \frac{1}{nt_n^2} \ln \left[\sum_{i \in S_n(x)} \left(\boldsymbol{a}^{\otimes n}\right)^{\downarrow}_i \right] = -\frac{x^2}{2V(\boldsymbol{a})}, \tag{7.38}$$

which is equivalent to Equation (7.30a). An analogous proof can be performed for $x \geq 0$, resulting in Eq. (7.30b). □

Remark 3. *One can extend Lemma 7.3 to probability distributions $\boldsymbol{a}$ with $V(\boldsymbol{a}) = 0$ by a direct calculation, since $V(\boldsymbol{a}) = 0$ means all non-zero entries of $\boldsymbol{a}$ are equal. One then obtains that Eq. (7.30a) holds for $x < 0$ and Eq. (7.30b) for $x > 0$, i.e., both expressions diverge to $-\infty$.*

7.4.3 Tail bounds for total distributions

Recall that in Sec. 7.2.3 we defined total initial and target states for a given rate r_n as

$$\boldsymbol{P}^n := \boldsymbol{p}^{\otimes n} \otimes \boldsymbol{f}^{\otimes nr_n}, \quad \boldsymbol{Q}^n := \boldsymbol{q}^{\otimes nr_n} \otimes \boldsymbol{f}^{\otimes n}, \tag{7.39}$$

where $\boldsymbol{f}$ stands for the free state of a given resource theory, i.e., $\boldsymbol{f}$ is a sharp state $\boldsymbol{s}$ for entanglement and coherence transformations, and $\boldsymbol{f}$ is the maximally mixed state $\boldsymbol{\eta}$ in the case of thermodynamic transformations (corresponding to the embedded thermal state $\boldsymbol{\gamma}$). For notational clarity we will henceforth omit the $\downarrow$ superscripts on these total states, assuming them to be ordered (i.e. we denote $\boldsymbol{P}^{n\downarrow}$ and $\boldsymbol{Q}^{n\downarrow}$ simply by $\boldsymbol{P}^n$ and $\boldsymbol{Q}^n$).

Analogous to the quantity which appears in our moderate deviation bounds, consider the quantity

$$K_n(x) := \exp\left(H(\boldsymbol{Q}^n) + xnt_n\right). \tag{7.40}$$

Using Lemma 7.3 we can prove the following tail bounds for the total distributions.

Lemma 7.4 (Tail bound for $\boldsymbol{P}^n$ and $\boldsymbol{Q}^n$)**.** *For any $\mu \in \mathbb{R}$, consider the conversion rate*

$$r_n(\mu) = \frac{H(\boldsymbol{f}) - H(\boldsymbol{p}) + \mu t_n}{H(\boldsymbol{f}) - H(\boldsymbol{q})}, \tag{7.41}$$

and the irreversibility parameter

$$\nu := \frac{V(\boldsymbol{p})}{V(\boldsymbol{q})} \cdot \frac{H(\boldsymbol{f}) - H(\boldsymbol{q})}{H(\boldsymbol{f}) - H(\boldsymbol{p})}. \tag{7.42}$$

The total output state $\boldsymbol{Q}^n$ has the tail bounds

$$x \leq 0: \quad \lim_{n\to\infty} \frac{1}{nt_n^2} \ln \left[\sum_{i \leq K_n(x)} Q_i^n \right] = -\frac{\nu x^2}{2V(\boldsymbol{p})}, \tag{7.43a}$$

$$x \geq 0: \quad \lim_{n\to\infty} \frac{1}{nt_n^2} \ln \left[\sum_{i \geq K_n(x)} Q_i^n \right] = -\frac{\nu x^2}{2V(\boldsymbol{p})}. \tag{7.43b}$$

Similarly, the total input state $\boldsymbol{P}^n$ has the tail bounds

$$x \leq \mu : \quad \lim_{n\to\infty} \frac{1}{nt_n^2} \ln \left[\sum_{i \leq K_n(x)} P_i^n \right] = -\frac{(x-\mu)^2}{2V(\boldsymbol{p})}, \tag{7.44a}$$

$$x \geq \mu : \quad \lim_{n\to\infty} \frac{1}{nt_n^2} \ln \left[\sum_{i \geq K_n(x)} P_i^n \right] = -\frac{(x-\mu)^2}{2V(\boldsymbol{p})}. \tag{7.44b}$$

Proof. For $\boldsymbol{f} = \boldsymbol{s}$ we have

$$\begin{aligned} \sum_{i \leq K_n(x)} Q_i^n &= \sum_{i \leq k_{nr_n}(\boldsymbol{q}, x/r_n)} \left(\boldsymbol{q}^{\otimes nr_n} \otimes \boldsymbol{s}^{\otimes n}\right)_i^{\downarrow} \\ &= \sum_{i \leq k_{nr_n}(\boldsymbol{q}, x/r_n)} \left(\boldsymbol{q}^{\otimes nr_n}\right)_i^{\downarrow}. \end{aligned} \tag{7.45}$$

Similarly for $\boldsymbol{f} = \boldsymbol{\eta}$ we have

$$\begin{aligned} \sum_{i \leq K_n(x)} Q_i^n &= \sum_{i \leq d^n k_{nr_n}(\boldsymbol{q}, x/r_n)} \left(\boldsymbol{q}^{\otimes nr_n} \otimes \boldsymbol{\eta}^{\otimes n}\right)_i^{\downarrow} \\ &= \sum_{i \leq k_{nr_n}(\boldsymbol{q}, x/r_n)} \left(\boldsymbol{q}^{\otimes nr_n}\right)_i^{\downarrow}. \end{aligned} \tag{7.46}$$

Applying Lemma 7.3 to both of the above equations yields the desired bounds.

Next, define $K_n^P(y) := \exp\left(H(\boldsymbol{P}^n) + ynt_n\right)$. By analogy to $\boldsymbol{Q}^n$, we have the following tail bounds on $\boldsymbol{P}^n$

$$y \leq 0 : \quad \lim_{n\to\infty} \frac{1}{nt_n^2} \ln \left[\sum_{i \leq K_n^P(y)} P_i^n \right] = -\frac{y^2}{2V(\boldsymbol{p})}, \tag{7.47a}$$

$$y \geq 0 : \quad \lim_{n\to\infty} \frac{1}{nt_n^2} \ln \left[\sum_{i \geq K_n^P(y)} P_i^n \right] = -\frac{y^2}{2V(\boldsymbol{p})}. \tag{7.47b}$$

Using the rate $r_n(\mu)$, and expanding out both K_n and K_n^P, we find that $K_n^P(x-\mu) = K_n(x)$. Substituting this into the above expressions, we get the desired tail bounds purely in terms of $K_n(x)$. □

We now want to consider the regions in which our two total distributions majorise each other. To do this, we first define the values of x for which the tail bounds for $\boldsymbol{P}^n$ and $\boldsymbol{Q}^n$ coincide. Let us introduce

$$z_{\mathrm{C}} := \frac{\mu}{1-\sqrt{\nu}}, \qquad \text{and} \qquad z_{\mathrm{T}} := \frac{\mu}{1+\sqrt{\nu}}. \tag{7.48}$$

These correspond to the values of x for which the moderate deviation tail bounds of the total distributions meet on the same side (cis) or on opposite sides (trans), respectively. More precisely, as a

consequence of Lemma 7.4, z_{C} and z_{T} are the solutions to the following equations

$$\lim_{n\to\infty}\frac{1}{nt_n^2}\ln\left[\sum_{i\leq K_n(z_{\mathrm{C}})}P_i^n\right]=\lim_{n\to\infty}\frac{1}{nt_n^2}\ln\left[\sum_{i\leq K_n(z_{\mathrm{C}})}Q_i^n\right], \tag{7.49a}$$

$$\lim_{n\to\infty}\frac{1}{nt_n^2}\ln\left[\sum_{i\geq K_n(z_{\mathrm{C}})}P_i^n\right]=\lim_{n\to\infty}\frac{1}{nt_n^2}\ln\left[\sum_{i\geq K_n(z_{\mathrm{C}})}Q_i^n\right], \tag{7.49b}$$

$$\lim_{n\to\infty}\frac{1}{nt_n^2}\ln\left[\sum_{i\leq K_n(z_{\mathrm{T}})}P_i^n\right]=\lim_{n\to\infty}\frac{1}{nt_n^2}\ln\left[\sum_{i\geq K_n(z_{\mathrm{T}})}Q_i^n\right], \tag{7.49c}$$

$$\lim_{n\to\infty}\frac{1}{nt_n^2}\ln\left[\sum_{i\geq K_n(z_{\mathrm{T}})}P_i^n\right]=\lim_{n\to\infty}\frac{1}{nt_n^2}\ln\left[\sum_{i\leq K_n(z_{\mathrm{T}})}Q_i^n\right]. \tag{7.49d}$$

We schematically present the positions of z_{C} and z_{T} in Fig. 7.1, which also serves to illustrate the proof of the following lemma.

Lemma 7.5 (Dominance of total states)**.** *For a bounded interval $[a,b]$, such that $a>z_{\mathrm{C}}$ (for $\nu<1$) or $b<z_{\mathrm{C}}$ (for $\nu>1$), the inequalities*

$$\sum_{i\leq K_n(x)}P_i^n>\sum_{i\leq K_n(x)}Q_i^n \tag{7.50}$$

hold for all $x\in[a,b]$, for sufficiently large n. Similarly, for any bounded interval $[a,b]$ with $b<z_{\mathrm{T}}$, the inequalities

$$\sum_{i\geq K_n(x)}P_i^n>\sum_{i\leq K_n(x)}Q_i^n \tag{7.51}$$

hold for all $x\in[a,b]$, for sufficiently large n.

Proof. We will prove Eq. (7.50) and explain how Eq. (7.51) can be proven in an analogous way. Consider the function $L(y)=\log\frac{y}{1-y}$, which is strictly increasing for $y\in(0,1)$. Next, define two sequences of functions

$$f_n(x):=\frac{2V(\boldsymbol{p})}{nt_n^2}L\left(\sum_{i\leq K_n(x)}P_i^n\right), \tag{7.52a}$$

$$g_n(x):=\frac{2V(\boldsymbol{p})}{nt_n^2}L\left(\sum_{i\leq K_n(x)}Q_i^n\right). \tag{7.52b}$$

We can combine the direct and converse parts of Lemma 7.3, i.e., Eqs. (7.30a) and (7.30b), to obtain the limits $f_n\to f$ and $g_n\to g$ as $n\to\infty$, where

$$f(x)=\frac{(x-\mu)^3}{|x-\mu|}\quad\text{and}\quad g(x)=\frac{\nu x^3}{|x|}. \tag{7.53a}$$

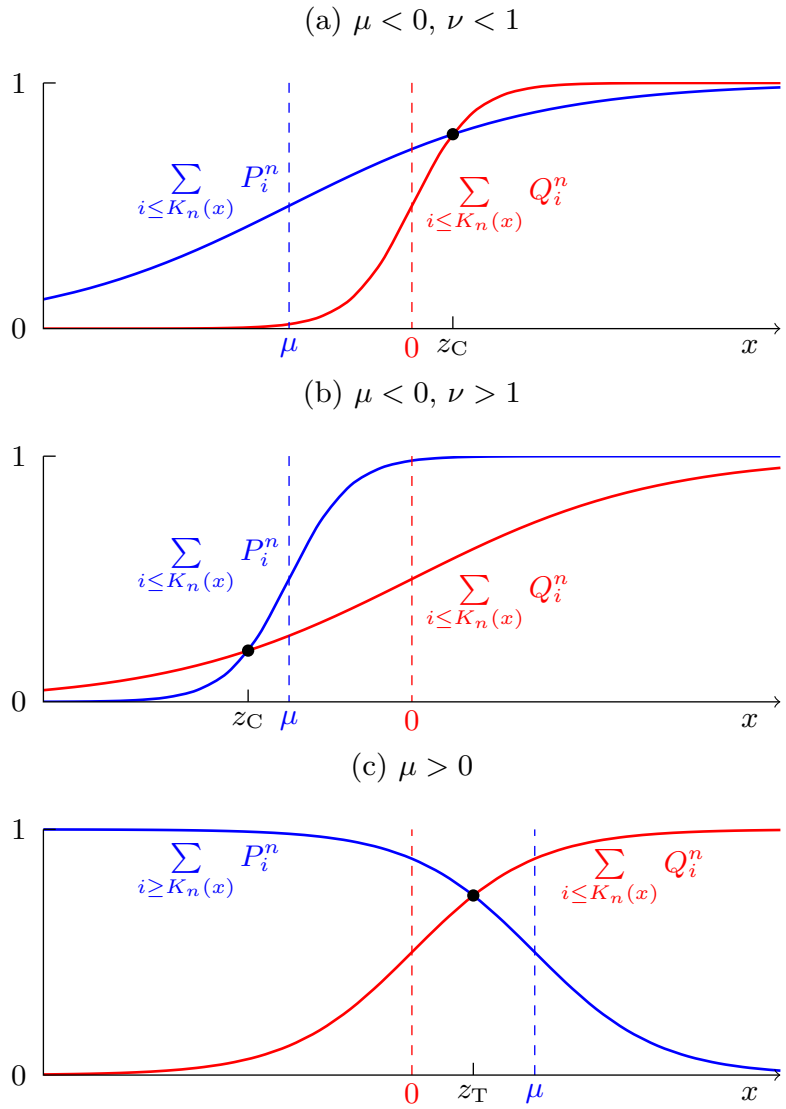


Figure 7.1: Schematic representation of cumulative distribution functions for $\boldsymbol{P}^n$ and $\boldsymbol{Q}^n$, and the positions of z_{C} and z_{T} in different regimes.

We now find that

$$f(x) > g(x) \quad \Longleftrightarrow \quad \begin{cases} x > z_{\mathrm{C}} & \text{for } \nu < 1, \\ x < z_{\mathrm{C}} & \text{for } \nu > 1. \end{cases} \tag{7.54}$$

Therefore, for any x in the above regions we have that $f_n(x) \overset{\text{ev.}}{>} g_n(x)$. As L is strictly monotone, this in turn implies that Eq. (7.50) holds eventually in the same region.

The above argument only ensures that $f_n \to f$ and $g_n \to g$ *point-wise*. This does not yet allow us to conclude that there exists an N such that Eq. (7.50) will hold *for all* $x \in [a, b]$ and all $n \geq N$. In

Section 7.B we close this gap by proving that this convergence is compact.

Finally, Eq. (7.51) can be proven in an analogous way by substituting: $g_n \to -g_n$ and $g \to -g$. □

7.5 Proof of the main theorem

We are now ready to prove Theorems 7.1, 7.2 and 7.3. We will achieve this in a series of steps. First, we will prove the following result.

Proposition 7.6 (Thermodynamic direction, $\boldsymbol{\gamma} = \boldsymbol{\eta}$, TVD)**.** *For a TVD error level of*

$$\epsilon_n^- = e^{-nt_n^2} \quad \text{or} \quad \epsilon_n^+ = 1 - e^{-nt_n^2}, \tag{7.55}$$

the approximation majorisation condition

$$\boldsymbol{P}^n \succ_{\epsilon_n^\pm} \boldsymbol{Q}^n \tag{7.56}$$

with $\boldsymbol{f} = \boldsymbol{\eta}$ holds with an optimal interconversion rate of

$$R_n^{\text{th}}(\epsilon_n^\pm) \simeq \frac{D(\boldsymbol{p}\|\boldsymbol{\eta}) \pm \sqrt{2V(\boldsymbol{p}\|\boldsymbol{\eta})}\left|1 \pm 1/\sqrt{\nu^{\text{th}}}\right| t_n}{D(\boldsymbol{q}\|\boldsymbol{\eta})}. \tag{7.57}$$

The proof of Proposition 7.6 will consist of two parts: first, in Sec. 7.5.1, we will show that the claimed rate is achievable for the given error; and then, in Sec. 7.5.2, that it is also optimal. This way we will prove a special case of Theorem 7.2 for the case of infinite temperature (when $\boldsymbol{\gamma} = \boldsymbol{\eta}$) and error level measured only by TVD; and Theorem 7.3 for the thermodynamic direction and infinite temperature.

The next step is to generalise Proposition 7.6 to arbitrary finite temperatures (arbitrary thermal state $\boldsymbol{\gamma}$). It is enough to note that the approximate interconversion condition for the thermodynamic direction, $\hat{\boldsymbol{P}}^n \succ_\epsilon \hat{\boldsymbol{Q}}^n$, is exactly captured by Eq. (7.56) if one only replaces $\boldsymbol{p}$ and $\boldsymbol{q}$ with $\hat{\boldsymbol{p}}$ and $\hat{\boldsymbol{q}}$, respectively. Moreover, since the relative entropy and relative entropy variance are invariant under the embedding map [7.8], one can obtain the optimal rate by "unembeddinig" Eq. (7.57), i.e., replacing $\hat{\boldsymbol{p}}$, $\hat{\boldsymbol{q}}$ and $\boldsymbol{\eta} = \hat{\boldsymbol{\gamma}}$ with $\boldsymbol{p}$, $\boldsymbol{q}$ and $\boldsymbol{\gamma}$ respectively. Thus, by proving Proposition 7.6, we in fact prove Theorem 7.2 for any temperature and the error level measured by TVD; and Theorem 7.3 for the thermodynamic direction with arbitrary temperature.

Then, in Sec. 7.5.3, we will prove the following result

Proposition 7.7 (Entanglement direction, TVD)**.** *For a TVD error level of*

$$\epsilon_n^- = e^{-nt_n^2} \quad \text{or} \quad \epsilon_n^+ = 1 - e^{-nt_n^2}, \tag{7.58}$$

the approximation majorisation condition

$$\boldsymbol{P}^n \prec_{\epsilon_n^\pm} \boldsymbol{Q}^n \tag{7.59}$$

with $\boldsymbol{f} = \boldsymbol{s}$ holds with an optimal interconversion rate of

$$R_n^{\text{ent}}(\epsilon_n^\pm) \simeq \frac{H(\boldsymbol{p}) \pm \sqrt{2V(\boldsymbol{p})}\left|1 \pm 1/\sqrt{\nu^{\text{ent}}}\right| t_n}{H(\boldsymbol{q})}. \tag{7.60}$$

To prove Proposition 7.7 we will leverage the proof of Proposition 7.6. More precisely, we will explain how to adapt that proof, so that the thermodynamic direction gets replaced by the entanglement direction. This way we will prove Theorem 7.1 for the error level measured by TVD; and Theorem 7.3 for the entanglement direction.

The final missing piece is to show that Theorems 7.1 and 7.2 also hold for the error level measured by infidelity. We will prove this in Sec. 7.5.4, again by explaining the necessary modifications of the reasoning that will result in replacing TVD with infidelity distance.

7.5.1 Proof of Proposition 7.6 (Achieveability)

We will start by considering the *achieveability* of Proposition 7.6, i.e., a lower bound on the optimal conversion rate for the thermodynamic direction. For notational convenience, we will drop the superscripts on both R_n^{th} and ν^{th}, adopt the convention $D(\cdot) := D(\cdot\|\boldsymbol{\eta})$, and note that $V(\cdot\|\boldsymbol{\eta}) = V(\cdot)$. Specifically, we will prove the following:

Lemma 7.8 (Proposition 7.6: Achieveability)**.** *For a TVD error level of*

$$\epsilon_n^- = e^{-nt_n^2} \quad \text{or} \quad \epsilon_n^+ = 1 - e^{-nt_n^2}, \tag{7.61}$$

the optimal rate is lower bounded,

$$R_n(\epsilon_n^\pm) \gtrsim \frac{D(\boldsymbol{p}) \pm \sqrt{2V(\boldsymbol{p})}\,|1 \pm 1/\sqrt{\nu}|\,t_n}{D(\boldsymbol{q})}. \tag{7.62}$$

We will prove this lemma by constructing a family of distributions $\tilde{\boldsymbol{P}}^{(\mu)}$, which eventually obey the required majorisation condition $\tilde{\boldsymbol{P}}^{(\mu)} \succ \boldsymbol{Q}^n$. We will then show that by picking appropriate values of $\mu = \mu_+$ and $\mu = \mu_-$, one can obtain distributions $\tilde{\boldsymbol{P}}^{(\mu_\pm)}$ such that $\delta(\boldsymbol{P}^n, \tilde{\boldsymbol{P}}^{(\mu_\pm)}) \overset{\text{ev.}}{\leq} \epsilon_n^\pm$.

Constructing the approximate distribution $\tilde{\mathbf{P}}^{(\mu)}$

To prove achieveability, we will construct a family of distributions $\tilde{\boldsymbol{P}}^{(\mu)}$ which, for any fixed μ, eventually majorise $\boldsymbol{Q}^n$. As in Lemma 7.4, consider the rate

$$r_n(\mu) = \frac{D(\boldsymbol{p}) + \mu t_n}{D(\boldsymbol{q})}, \tag{7.63}$$

where $\mu \in \mathbb{R}$ is a parameter of our construction. We will construct $\tilde{\boldsymbol{P}}^{(\mu)}$ by the *cut-and-pile* method. Specifically we will consider starting with $\boldsymbol{P}^n$, removing mass from its tail, and adding it to the largest element. This construction allows us to construct a nearby state which is higher in the majorisation order. We start by defining the cutting point,

$$z_{\mu,\nu} := \begin{cases} 2\mu - z_{\text{C}} & : \mu < 0, \quad \nu < 1, \\ z_{\text{C}} & : \mu < 0, \quad \nu > 1, \\ z_{\text{T}} & : \mu > 0. \end{cases} \tag{7.64}$$

If we let $\zeta > 0$ be a small slack parameter, then $\tilde{\boldsymbol{P}}^{(\mu)}$ is defined as

$$\tilde{P}_i^{(\mu)} := \begin{cases} P_1^n + \sum_{i \geq K_n(y)} P_i^n & : i = 1, \\ P_i^n & : 1 < i < K_n(y), \\ 0 & : i \geq K_n(y), \end{cases} \tag{7.65}$$

with $y = z_{\mu,\nu} - \zeta$.

Showing majorisation $\tilde{\mathbf{P}}^{(\mu)} \succ \tilde{\mathbf{Q}}^n$

Given the above construction, we now want to prove the majorisation condition $\tilde{\boldsymbol{P}}^{(\mu)} \succ \boldsymbol{Q}^n$ eventually holds. The idea here is to leverage Lemma 7.5, and show that a cut-and-pile construction with a cut at $K_n(x)$ for any $x < z_{\mu,\nu}$ will always eventually majorise $\boldsymbol{Q}^n$.

Lemma 7.9. *For any fixed μ, $\tilde{\boldsymbol{P}}^{(\mu)} \overset{ev.}{\succ} \boldsymbol{Q}^n$.*

Proof. To prove majorisation we need to show that, eventually, the inequalities

$$\sum_{i=1}^{k} \tilde{P}_i^{(\mu)} \geq \sum_{i=1}^{k} Q_i^n \tag{7.66}$$

hold for all k. The 'cut' of the cut-and-pile construction implies Eq. (7.66) for large k. Specifically, the restricted support of $\tilde{\boldsymbol{P}}^{(\mu)}$,

$$\sum_{i < K_n(z_{\mu,\nu} - \zeta)} \tilde{P}_i^{(\mu)} = 1, \tag{7.67}$$

implies that Eq. (7.66) holds trivially for any $k \geq K_n(z_{\mu,\nu} - \zeta)$.

The idea now is to show that the 'pile' similarly gives us majorisation for small k, and then to leverage Lemma 7.5 to argue that $\boldsymbol{P}^n$ already majorises $\boldsymbol{Q}^n$ for intermediate k. We will split this argument into the three cases given in the definition of $z_{\mu,\nu}$, and illustrated by panels (a)-(c) of Fig. 7.1.

Case 1: $\mu < 0$, $\nu < 1$. Noticing that $\boldsymbol{P}^n$ tail bounds in Lemma 7.4 are symmetric under $x \to 2\mu - x$, we have that

$$\lim_{n\to\infty} \frac{1}{nt_n^2} \ln \left[\sum_{i \geq K_n(z_{\mu,\nu} - \zeta)} P_i^n \right] = \lim_{n\to\infty} \frac{1}{nt_n^2} \ln \left[\sum_{i \leq K_n(z_{\mathrm{C}} + \zeta)} P_i^n \right]. \tag{7.68}$$

Applying Lemma 7.5, we therefore have that

$$\sum_{i \geq K_n(z_{\mu,\nu} - \zeta)} P_i^n \overset{\text{ev.}}{>} \sum_{i \leq K_n(z_{\mathrm{C}} + \zeta)} Q_i^n. \tag{7.69}$$

Using this, for any $k \leq K_n(z_{\mathrm{C}} + \zeta)$ we can leverage the 'pile' in the $\tilde{\boldsymbol{P}}^{(\mu)}$ construction to yield

$$\sum_{i \leq k} \tilde{P}_i^{(\mu)} \geq \tilde{P}_1^{(\mu)} > \sum_{i \geq K_n(z_{\mu,\nu} - \zeta)} P_i^n \overset{\text{ev.}}{>} \sum_{i \leq K_n(z_{\mathrm{C}} + \zeta)} Q_i^n \geq \sum_{i \leq k} Q_i^n, \tag{7.70}$$

which implies Eq. (7.66). Applying Lemma 7.5, we have that Eq. (7.66) must also eventually hold on the remaining intermediate indices $k \in [K_n(z_{\mathrm{C}} + \zeta), K_n(z_{\mu,\nu} - \zeta)]$.

Case 2: $\mu < 0$, $\nu > 1$. Noticing that $\boldsymbol{Q}^n$ tail bounds in Lemma 7.4 are symmetric under $x \to -x$, we have that

$$\lim_{n\to\infty} \frac{1}{nt_n^2} \ln \left[\sum_{i \geq K_n(z_{\mu,\nu})} Q_i^n \right] = \lim_{n\to\infty} \frac{1}{nt_n^2} \ln \left[\sum_{i \leq K_n(-z_{\mathrm{C}})} Q_i^n \right]. \tag{7.71}$$

We now use Lemma 7.5 again, giving for any $k \leq K_n(-z_{\mathrm{C}})$ that the 'pile' of $\tilde{\boldsymbol{P}}^{(\mu)}$ implies Eq. (7.66),

$$\sum_{i \leq k} \tilde{P}_i^{(\mu)} > \sum_{i \geq K_n(z_{\mu,\nu} - \zeta)} P_i^n \overset{\text{ev.}}{>} \sum_{i \leq K_n(-z_{\mathrm{C}})} Q_i^n \geq \sum_{i \leq k} Q_i^n. \tag{7.72}$$

A direct application of Eq. (7.5) implies Equation (7.66) also holds on all of the remaining intermediate indices $k \in [K_n(-z_{\mathrm{C}}), K_n(z_{\mu,\nu} - \zeta)]$.

Case 3: $\mu > 0$. For $k \leq K_n(z_{\mathrm{T}} - \zeta)$, we use Lemma 7.5 to eventually give

$$\sum_{i \leq k} \tilde{P}_i^{(\mu)} > \sum_{i \geq K_n(z_{\mathrm{T}} - \zeta)} P_i^n \overset{\text{ev.}}{>} \sum_{i \leq K_n(z_{\mathrm{T}} - \zeta)} Q_i^n \geq \sum_{i \leq k} Q_i^n. \tag{7.73}$$

In this case there is no intermediate region, which implies Eq. (7.66) holds for all k. □

Showing that $\tilde{\mathbf{P}}^{(\mu)}$ is close to $\mathbf{P}^n$

Now that we have shown that $\tilde{\boldsymbol{P}}^{(\mu)}$ eventually majorises $\boldsymbol{Q}^n$, we want to ask how close $\tilde{\boldsymbol{P}}^{(\mu)}$ is to $\boldsymbol{P}^n$ in total variation distance. The answer is provided by the following lemma.

Lemma 7.10. *For a fixed $\mu < 0$,*

$$\delta(\tilde{\boldsymbol{P}}^{(\mu)}, \boldsymbol{P}^n) \overset{ev.}{<} \exp\left(-\frac{(z_{\mu,\nu} - \mu - 2\zeta)^2}{2V(\boldsymbol{p})} nt_n^2 \right), \tag{7.74}$$

and similarly for $\mu > 0$,

$$\delta(\tilde{\boldsymbol{P}}^{(\mu)}, \boldsymbol{P}^n) \overset{ev.}{<} 1 - \exp\left(-\frac{(z_{\mu,\nu} - \mu - 2\zeta)^2}{2V(\boldsymbol{p})} nt_n^2 \right). \tag{7.75}$$

Proof. A convenient feature of the cut-and-pile construction is that the total variation distance to the original distribution takes a particularly simple form. Specifically,

$$\delta(\tilde{\boldsymbol{P}}^{(\mu)}, \boldsymbol{P}^n) = \sum_{i \geq K_n(z_{\mu,\nu} - \zeta)} P_i^n. \tag{7.76}$$

We can now apply Lemma 7.4 to give

$$\lim_{n\to\infty} \frac{1}{nt_n^2} \ln \left[\delta(\tilde{\boldsymbol{P}}^{(\mu<0)}, \boldsymbol{P}^n) \right] = -\frac{(z_{\mu,\nu} - \mu - \zeta)^2}{2V(\boldsymbol{p})}, \tag{7.77a}$$

$$\lim_{n\to\infty} \frac{1}{nt_n^2} \ln \left[1 - \delta(\tilde{\boldsymbol{P}}^{(\mu>0)}, \boldsymbol{P}^n) \right] = -\frac{(z_{\mu,\nu} - \mu - \zeta)^2}{2V(\boldsymbol{p})}. \tag{7.77b}$$

This, in turn, implies the eventual inequalities

$$\ln\left[\delta(\tilde{\boldsymbol{P}}^{(\mu<0)}, \boldsymbol{P}^n)\right] \overset{\text{ev.}}{<} -\frac{(z_{\mu,\nu} - \mu - 2\zeta)^2}{2V(\boldsymbol{p})} n t_n^2, \tag{7.78a}$$

$$\ln\left[1 - \delta(\tilde{\boldsymbol{P}}^{(\mu>0)}, \boldsymbol{P}^n)\right] \overset{\text{ev.}}{>} -\frac{(z_{\mu,\nu} - \mu - 2\zeta)^2}{2V(\boldsymbol{p})} n t_n^2, \tag{7.78b}$$

which are equivalent to our desired bounds. □

Proof of Achieveability

We now put together the above to prove the achieveability of Proposition 7.6.

Proof of Lemma 7.8. Consider our construction of $\tilde{\boldsymbol{P}}^{(\mu)}$ for a specific choice of μ. Specifically, let

$$\mu_\pm = \left|1 \pm 1/\sqrt{\nu}\right| \left(\pm\sqrt{2V(\boldsymbol{p})} - 2\zeta\right), \tag{7.79}$$

where μ_- will give the direct result (with vanishing error ϵ_n^-), and μ_+ will give the converse result (with error ϵ_n^+ approaching 1).

From Lemma 7.9 we have that $\tilde{\boldsymbol{P}}^{(\mu_\pm)} \overset{\text{ev.}}{\succ} \boldsymbol{Q}^n$ as required. Substituting $\mu_\pm$ in Lemma 7.10, we have that the TVD error is bounded,

$$\delta\left(\tilde{\boldsymbol{P}}^{(\mu_\pm)}, \boldsymbol{P}^n\right) \overset{\text{ev.}}{<} \epsilon_n^\pm, \tag{7.80}$$

as required. The rate in these cases takes the form

$$\begin{aligned} r_n(\mu_\pm) &= \frac{D(\boldsymbol{p}) + \mu_\pm t_n}{D(\boldsymbol{q})} \\ &= \frac{D(\boldsymbol{p}) + \left|1 \pm 1/\sqrt{\nu}\right| \left(\pm\sqrt{2V(\boldsymbol{p})} - 2\zeta\right) t_n}{D(\boldsymbol{q})}. \end{aligned} \tag{7.81}$$

As the above analysis holds for any $\zeta > 0$, we can therefore conclude that

$$R_n(\epsilon_n^\pm) \gtrsim \frac{D(\boldsymbol{p}) \pm \sqrt{2V(\boldsymbol{p})}\left|1 \pm 1/\sqrt{\nu}\right| t_n}{D(\boldsymbol{q})}, \tag{7.82}$$

as desired. □

7.5.2 Proof of Proposition 7.6 (Optimality)

We now move on to showing the *optimality* of Proposition 7.6, i.e. an upper bound on the optimal conversion rate for the thermodynamic direction. For convenience we will reuse the notation used in the proof of achieveability.

Lemma 7.11 (Proposition 7.6: Optimality)**.** *For a TVD error level of*

$$\epsilon_n^- = e^{-nt_n^2} \quad \text{or} \quad \epsilon_n^+ = 1 - e^{-nt_n^2}, \tag{7.83}$$

the optimal rate is upper bounded,

$$R_n(\epsilon_n^\pm) \lesssim \frac{D(\boldsymbol{p}) \pm \sqrt{2V(\boldsymbol{p})}\,|1 \pm 1/\sqrt{\nu}|\,t_n}{D(\boldsymbol{q})}. \tag{7.84}$$

We will prove the above by showing that, for any distribution $\tilde{\boldsymbol{P}}^n$ obeying the majorisation condition $\tilde{\boldsymbol{P}}^n \succ \boldsymbol{Q}^n$ with a rate $r_n(\mu)$, the TVD distance between $\tilde{\boldsymbol{P}}^n$ and $\boldsymbol{P}^n$ is eventually lower bounded by $\epsilon_n^\pm$. To achieve this, we will use the fact that the total variation distance is monotonically decreasing under coarse-graining. Specifically, for any distributions $\boldsymbol{a}$ and $\boldsymbol{b}$, and index k, the triangle inequality gives

$$\delta(\boldsymbol{a}, \boldsymbol{b}) \geq \left| \sum_{i \leq k} a_i - \sum_{j \leq k} b_j \right|, \tag{7.85}$$

Applying this to distributions $\tilde{\boldsymbol{P}}^n$ and $\boldsymbol{P}^n$, and index $k = K_n(x)$, gives

$$\delta(\tilde{\boldsymbol{P}}^n, \boldsymbol{P}^n) \geq \left| \sum_{i \leq K_n(x)} \tilde{P}_i^n - \sum_{j \leq K_n(x)} P_j^n \right|. \tag{7.86}$$

The idea is that, with a careful choice of x, we will be able to use the majorisation $\tilde{\boldsymbol{P}}^n \succ \boldsymbol{Q}^n$ to replace the summations over $\tilde{\boldsymbol{P}}^n$ with those over $\boldsymbol{Q}^n$. This will then allow us to apply Lemma 7.4 to arrive at our final bound on the error.

We will first present this argument in detail for the case where $\mu < 0$ and $\nu < 1$, and then present the modifications required for the remaining cases of $\mu < 0$ and $\nu > 1$, and $\mu > 0$.

Case 1: $\mu < 0$**,** $\nu < 1$

Here we will perform our coarse-grained binning at $x = z_{\mathrm{C}} - \zeta$. Recalling that $\tilde{\boldsymbol{P}}^n \succ \boldsymbol{Q}^n$, we have that

$$\sum_{i \leq K_n(z_{\mathrm{C}} - \zeta)} \tilde{P}_i^n \geq \sum_{i \leq K_n(z_{\mathrm{C}} - \zeta)} Q_i^n. \tag{7.87}$$

Using the positivity of ζ, Lemma 7.5 allows us to also conclude that

$$\sum_{i \leq K_n(z_{\mathrm{C}} - \zeta)} \tilde{P}_i^n \overset{\text{ev.}}{>} \sum_{i \leq K_n(z_{\mathrm{C}} - \zeta)} P_i^n. \tag{7.88}$$

We now use the fact that $|\alpha - \beta|$ is monotonically increasing in α for $\alpha \geq \beta$. Using this, we have an eventual lower bound on TVD purely in terms of the total states $\boldsymbol{P}^n$ and $\boldsymbol{Q}^n$,

$$\delta(\tilde{\boldsymbol{P}}^n, \boldsymbol{P}^n) \geq \left| \sum_{i \leq K_n(z_{\mathrm{C}} - \zeta)} Q_i^n - \sum_{j \leq K_n(z_{\mathrm{C}} - \zeta)} P_j^n \right|. \tag{7.89}$$

Applying Lemma 7.4, we see that the tail sum of $\boldsymbol{Q}^n$ asymptotically dominates. Specifically, we have

$$\liminf_{n\to\infty} \frac{1}{nt_n^2} \ln\left[\delta(\tilde{\boldsymbol{P}}^n, \boldsymbol{P}^n)\right] \geq -\frac{\nu(z_{\mathrm{C}}-\zeta)^2}{2V(\boldsymbol{p})}, \tag{7.90}$$

and therefore

$$\delta(\tilde{\boldsymbol{P}}^n, \boldsymbol{P}^n) \overset{\text{ev.}}{>} \exp\left(-\frac{\nu(z_{\mathrm{C}}-2\zeta)^2}{2V(\boldsymbol{p})} nt_n^2\right). \tag{7.91}$$

We now choose

$$\mu = \left|1-1/\sqrt{\nu}\right| \left(-\sqrt{2V(\boldsymbol{p})} + 2\zeta\sqrt{\nu}\right). \tag{7.92}$$

This gives us that $\delta(\tilde{\boldsymbol{P}}^n, \boldsymbol{P}^n) \overset{\text{ev.}}{>} \epsilon_n^-$, with a rate of

$$r_n(\mu) = \frac{D(\boldsymbol{p}) + \left|1-1/\sqrt{\nu}\right| \left(-\sqrt{2V(\boldsymbol{p})} + 2\zeta\sqrt{\nu}\right) t_n}{D(\boldsymbol{q})}. \tag{7.93}$$

As this is true for all $\zeta > 0$, we can conclude that

$$R_n(\epsilon_n^-) \lesssim \frac{D(\boldsymbol{p}) - \sqrt{2V(\boldsymbol{p})}\left|1-1/\sqrt{\nu}\right| t_n}{D(\boldsymbol{q})}. \tag{7.94}$$

Case 2: $\mu < 0$, $\nu > 1$

Now we consider the case of $\mu < 0$ and $\nu > 1$. The proof here is similar, starting with a cut at $x = z_{\mathrm{C}} + \zeta$. Here the tail sum of $\boldsymbol{P}^n$ asymptomatically dominates, with Lemma 7.4 giving

$$\delta(\tilde{\boldsymbol{P}}^n, \boldsymbol{P}^n) \overset{\text{ev.}}{>} \exp\left(-\frac{(z_{\mathrm{C}}-\mu+2\zeta)^2 nt_n^2}{2V(\boldsymbol{p})}\right). \tag{7.95}$$

We now make the choice

$$\mu = \left|1-1/\sqrt{\nu}\right| \left(-\sqrt{2V(\boldsymbol{p})} + 2\zeta\right), \tag{7.96}$$

which gives that $\delta(\tilde{\boldsymbol{P}}^n, \boldsymbol{P}^n) \overset{\text{ev.}}{>} \epsilon_n^-$, and thus

$$R_n(\epsilon_n^-) \lesssim \frac{D(\boldsymbol{p}) - \sqrt{2V(\boldsymbol{p})}\left|1-1/\sqrt{\nu}\right| t_n}{D(\boldsymbol{q})}. \tag{7.97}$$

Case 3: $\mu > 0$

Lastly, we consider the case of $\mu > 0$. Here, we perform our cut at $x = z_{\mathrm{T}}$ and, instead of using an argument based on majorisation *from above*, we use majorisation *from below*. Specifically, we note that $\tilde{\boldsymbol{P}}^n \succ \boldsymbol{Q}^n$ implies

$$\sum_{i>K_n(z_{\mathrm{T}})} \tilde{P}_i^n \leq \sum_{i>K_n(z_{\mathrm{T}})} Q_i^n. \tag{7.98}$$

We now use that $|\alpha - \beta|$ is monotonically *decreasing* in α for $\alpha \leq \beta$. This gives us the analogous upper bound

$$\delta(\tilde{\boldsymbol{P}}^n, \boldsymbol{P}^n) \geq \left| \sum_{i \leq K_n(z_{\mathrm{T}})} Q_i^n - \sum_{j \leq K_n(z_{\mathrm{T}})} P_j^n \right|. \tag{7.99}$$

Applying Lemma 7.4, we find that the tail sums of $\boldsymbol{P}^n$ and $\boldsymbol{Q}^n$ both dominate[1], yielding

$$\delta(\tilde{\boldsymbol{P}}^n, \boldsymbol{P}^n) \overset{\text{ev.}}{>} 1 - \exp\left(-\frac{(z_{\mathrm{T}} - \mu + \zeta)^2 n t_n^2}{2V(\boldsymbol{p})}\right). \tag{7.100}$$

for sufficiently large n. Substituting

$$\mu = \left|1 + 1/\sqrt{\nu}\right| \left(\sqrt{2V(\boldsymbol{p})} + \zeta\right), \tag{7.101}$$

we eventually have that $\delta(\tilde{\boldsymbol{P}}^n, \boldsymbol{P}^n) > \epsilon_n^+$, and thus

$$R_n(\epsilon_n^+) \lesssim \frac{D(\boldsymbol{p}) + \sqrt{2V(\boldsymbol{p})}\,|1 + 1/\sqrt{\nu}|\,t_n}{D(\boldsymbol{q})}. \tag{7.102}$$

7.5.3 Proof of Proposition 7.7

Proposition 7.6 states that the largest rate R_n such that the majorisation condition

$$\boldsymbol{p}^{\otimes n} \otimes \boldsymbol{\eta}^{\otimes nR_n} \succ_{\epsilon_n^\pm} \boldsymbol{q}^{\otimes nR_n} \otimes \boldsymbol{\eta}^{\otimes n} \tag{7.103}$$

holds is of the form

$$R_n(\epsilon_n^\pm) \simeq \frac{H(\boldsymbol{\eta}) - H(\boldsymbol{p}) \pm \sqrt{2V(\boldsymbol{p})}\,|1 \pm 1/\sqrt{\nu}|\,t_n}{H(\boldsymbol{\eta}) - H(\boldsymbol{q})}, \tag{7.104}$$

where

$$\nu = \frac{V(\boldsymbol{p})}{V(\boldsymbol{q})} \cdot \frac{H(\boldsymbol{\eta}) - H(\boldsymbol{q})}{H(\boldsymbol{\eta}) - H(\boldsymbol{p})}. \tag{7.105}$$

The proof of the above relied entirely on Lemma 7.4, which holds for $\boldsymbol{f} = \boldsymbol{\eta}$ as well as $\boldsymbol{f} = \boldsymbol{s}$. Thus, in order to rigorously prove Proposition 7.7, one could perform steps analogous to the ones presented in Sec. 7.5.1-7.5.2, with $\boldsymbol{\eta}$ replaced by $\boldsymbol{s}$. Instead, below we present a shorter proof that directly employs the result stated by Proposition 7.6.

Proof of Proposition 7.7. Since Lemma 7.4 holds both for $\boldsymbol{f} = \boldsymbol{\eta}$ and for $\boldsymbol{f} = \boldsymbol{s}$, the statement of Proposition 7.6 can be extended to show that the smallest R_n such that the majorisation condition

$$\boldsymbol{p}^{\otimes n} \otimes \boldsymbol{s}^{\otimes nR_n} \succ_{\epsilon_n^\pm} \boldsymbol{q}^{\otimes nR_n} \otimes \boldsymbol{s}^{\otimes n} \tag{7.106}$$

[1] The lack of a slack parameter in the cut means that these two sums compete. In the direct case (vanishing error ϵ_n^-) this leads to a catastrophic cancellation, but in the converse case, this causes no problem.

holds is of the form

$$R_n(\epsilon_n^{\pm}) \simeq \frac{H(\boldsymbol{s}) - H(\boldsymbol{p}) \pm \sqrt{2V(\boldsymbol{p})}\,|1 \pm 1/\sqrt{\nu}|\,t_n}{H(\boldsymbol{s}) - H(\boldsymbol{q})}, \tag{7.107}$$

where

$$\nu = \frac{V(\boldsymbol{p})}{V(\boldsymbol{q})} \cdot \frac{H(\boldsymbol{s}) - H(\boldsymbol{q})}{H(\boldsymbol{s}) - H(\boldsymbol{p})}. \tag{7.108}$$

We now want to reverse the direction of majorisation. The first step is simply to swap $\boldsymbol{p} \leftrightarrow \boldsymbol{q}$, and use the fact ϵ-post-majorisation is equivalent to ϵ-pre-majorisation [7.8]. This transforms the considered majorisation relation into the following form

$$\boldsymbol{p}^{\otimes nR_n} \otimes \boldsymbol{s}^{\otimes n} \prec_{\epsilon_n^{\pm}} \boldsymbol{q}^{\otimes n} \otimes \boldsymbol{s}^{\otimes nR_n}. \tag{7.109}$$

We see that R_n now forms the 'inverse rate' of the transformation between states represented by $\boldsymbol{p}$ and the ones represented by $\boldsymbol{q}$. To find the true rate of this transformation, we make the following substitutions

$$n \leftarrow nR_n, \quad t_n \leftarrow t_n/\sqrt{R_n}, \quad R_n \leftarrow 1/R_n. \tag{7.110}$$

As a result, the desired majorisation condition,

$$\boldsymbol{p}^{\otimes n} \otimes \boldsymbol{s}^{\otimes nR_n} \prec_{\epsilon_n^{\pm}} \boldsymbol{q}^{\otimes nR_n} \otimes \boldsymbol{s}^{\otimes n} \tag{7.111}$$

holds with the optimal rate of the form

$$\begin{aligned} R_n(\epsilon_n^{\pm}) &\simeq \frac{H(\boldsymbol{p})}{H(\boldsymbol{q}) \mp \sqrt{2V(\boldsymbol{q})\frac{H(\boldsymbol{q})}{H(\boldsymbol{p})}}\,|1 \pm \sqrt{\nu}|\,t_n} \\ &\simeq \frac{H(\boldsymbol{p}) \pm \sqrt{2V(\boldsymbol{p})}\,|1 \pm 1/\sqrt{\nu}|}{H(\boldsymbol{q})}, \end{aligned} \tag{7.112}$$

and

$$\nu = \frac{V(\boldsymbol{p})}{V(\boldsymbol{q})} \cdot \frac{H(\boldsymbol{q})}{H(\boldsymbol{p})}. \tag{7.113}$$

□

7.5.4 Extension to infidelity

We now want to argue that, in the direct regime (for vanishing error ϵ_n^-), our results extend to the case where we consider error in terms of infidelity instead of TVD. To show that the achieveability argument, presented in Sec. 7.5.1, extends to infidelity, we leverage the Fuchs-van de Graaf inequality [7.30], specifically

$$1 - \sqrt{F(\boldsymbol{P}^n, \tilde{\boldsymbol{P}}^n)} \leq \delta(\boldsymbol{P}^n, \tilde{\boldsymbol{P}}^n). \tag{7.114}$$

Using this, in the direct regime where $\delta(\boldsymbol{P}^n, \tilde{\boldsymbol{P}}^n) \to 0$, we have that the corresponding moderate exponential of infidelity must be bounded by that of the TVD,

$$\limsup_{n\to\infty} \frac{1}{nt_n^2} \ln\left[1 - F(\boldsymbol{P}^n, \tilde{\boldsymbol{P}}^n)\right] \leq \lim_{n\to\infty} \frac{1}{nt_n^2} \ln\left[\delta(\boldsymbol{P}^n, \tilde{\boldsymbol{P}}^n)\right]. \tag{7.115}$$

Therefore, the TVD upper bounds employed in the proof of achieveability naturally extend to the infidelity.

In the proof of optimality, presented in Sec. 7.5.2, we used coarse-graining and a monotonicity argument to bound the TVD as

$$\delta(\boldsymbol{P}^n, \tilde{\boldsymbol{P}}^n) \geq \left|\sum_{i\leq k} P_i^n - \sum_{j\leq k} Q_j^n\right|. \tag{7.116}$$

By using the data-processing inequality and an analogous monotonicity argument, we can similarly bound the fidelity as

$$\sqrt{F(\boldsymbol{P}^n, \tilde{\boldsymbol{P}}^n)} \leq \sqrt{\sum_{i\leq k} P_i^n \cdot \sum_{j\leq k} Q_j^n} + \sqrt{\sum_{i>k} P_i^n \cdot \sum_{j>k} Q_j^n}. \tag{7.117}$$

In the case where $\nu < 1$, we chose k such that

$$\sum_{i\leq k} P_i^n \ll \sum_{i\leq k} Q_i^n \ll 1, \tag{7.118}$$

and similarly for $\nu > 1$

$$\sum_{i>k} Q_i^n \ll \sum_{i>k} P_i^n \ll 1. \tag{7.119}$$

In either case, a single tail sum dominates in the bound upon both fidelity and TVD, similarly allowing us to lower bound the moderate exponent of the infidelity,

$$\liminf_{n\to\infty} \frac{1}{nt_n^2} \ln\left[1 - F(\boldsymbol{P}^n, \tilde{\boldsymbol{P}}^n)\right] \geq \lim_{n\to\infty} \frac{1}{nt_n^2} \ln\left[\delta(\boldsymbol{P}^n, \tilde{\boldsymbol{P}}^n)\right]. \tag{7.120}$$

7.6 Outlook

We have performed the moderate deviation analysis of resource interconversion problems for which single-shot transformation rules are based on majorisation and thermo-majorisation. As a result, in the regime of asymptotically vanishing error, we have found unified expressions for second-order corrections to asymptotic conversion rates within resource theories of entanglement, coherence and thermodynamics. More precisely, we obtained a family of results that specifies the optimal trade-off between the speed at which the conversion rate approaches the asymptotic rate, and the speed at which the error vanishes, when the number of transformed states n grows. Crucially, we have found that the correction term can vanish independently of n when a certain resonance condition between

the initial and final states is satisfied. This opens the path to transformation reversibility beyond the asymptotic limit, the phenomenon that we discuss in detail in the accompanying paper [7.18].

There are quite a few research directions that one may want to take in order to generalise and extend the results presented in this paper. Since the small deviation analysis of the majorisation-based resource interconversion has been performed in Refs. [7.7,7.8], and the moderate deviation analysis was the focus of the current work, the straightforward extension would be to investigate the interconversion problem in the large deviation regime. This may be of particular interest in the context of fluctuation-free work extraction, where one may want to sacrifice a constant fraction of possible work output for its quality [7.31]. On the other hand, one could also look for the exact expression for the third-order term of the asymptotic expansion of R_n, which one can conjecture to scale as $O(t_n^2 + \log n)$ [7.11]. Another obvious generalisation is to go beyond the restrictions of pure states (for entanglement direction) and energy-incoherent states (for thermodynamic direction). This, however, is a much harder task, as the single-shot transformation rules are still not known for these unrestricted cases. Finally, since the second-order analysis of resource interconversion led us not only to quantitative results, but also to qualitatively new predictions concerning finite-size reversibility, a similar analysis for other resource theories is now very well justified.

Acknowledgements

CC and KK acknowledge support from the ARC via the Centre of Excellence in Engineered Quantum Systems (EQuS), project number CE110001013. CC also acknowledges support from the Australian Institute for Nanoscale Science and Technology Postgraduate Scholarship (John Makepeace Bennett Gift). MT is funded by an ARC Discovery Early Career Researcher Award, project number DE160100821.

7.A Relation to small deviation bound

In this section we explore the relationship between our moderate deviation results of Theorems 7.1–7.3, and the small deviation results of Refs. [7.7,7.8]. In these papers, a second-order expansion of the rate is given for a *constant* infidelity error, in terms of the Rayleigh-normal distribution. Below we will consider the expansions of the Rayleigh-normal cumulative distribution function around $-\infty$, and show consistency with Theorems 7.1 and 7.2. Moreover, we will explain how the expansion around $+\infty$ leads to a conjecture analogous to Theorem 7.3 with error level measured by infidelity.

The Rayleigh-normal distributions are a parameterised family of distributions, depending on a parameter $\nu \geq 0$. For a formal definition and properties, see Ref. [7.7]. We will denote the associated cumulative distribution functions by $Z_\nu(\mu)$, for $\nu \geq 0$ and $\mu \in \mathbb{R}$. We will also adopt the notation for

Gaussian probability and cumulative distribution functions of

$$\phi(x) = \frac{1}{\sqrt{2\pi}} e^{-x^2/2}, \tag{7.121}$$

$$\phi_{\mu,\nu}(x) = \frac{1}{\sqrt{2\pi\nu}} e^{-(x-\mu)^2/2\nu}, \tag{7.122}$$

$$\Phi(x) = \int_{-\infty}^{x} \phi(t)\,\mathrm{d}t, \tag{7.123}$$

$$\Phi_{\mu,\nu}(x) = \int_{-\infty}^{x} \phi_{\mu,\nu}(t)\,\mathrm{d}t. \tag{7.124}$$

7.A.1 Expansion around $\mu = -\infty$

The crossing point $\alpha_{\mu,\nu}$ is defined by the equation [7.7]

$$\frac{\phi(\alpha_{\mu,\nu})}{\phi_{\mu,\nu}(\alpha_{\mu,\nu})} = \frac{\Phi(\alpha_{\mu,\nu})}{\Phi_{\mu,\nu}(\alpha_{\mu,\nu})}. \tag{7.125}$$

As $\mu \to -\infty$, $\alpha_{\mu,\nu} \to +\infty$. We will use the $x \to \infty$ approximation $\Phi(x) \approx 1$, which leads to $\phi(\alpha_{\mu,\nu})/\phi_{\mu,\nu}(\alpha_{\mu,\nu}) \approx 1$, resulting in

$$\alpha_{\mu,\nu} \approx \frac{\mu}{1-\sqrt{\nu}}. \tag{7.126}$$

We now look at the Rayleigh-normal distribution, which takes the form

$$\sqrt{1 - Z_\nu(\mu)} = \sqrt{\frac{2\sqrt{\nu}}{1+\nu}} e^{-\mu^2/4(1+\nu)} \Phi_{\frac{-\mu}{1+\nu}, \frac{2\nu}{1+\nu}}(-\alpha_{\mu,\nu}) + \sqrt{\Phi(\alpha_{\mu,\nu})\Phi_{\mu,\nu}(\alpha_{\mu,\nu})}. \tag{7.127}$$

We now note that, as $\alpha_{\mu,\nu} \to +\infty$, the first term on the RHS is exponentially vanishing, and second exponentially approaching 1. Specifically we have

$$\ln\left[1 - \sqrt{\Phi(\alpha_{\mu,\nu})\Phi_{\mu,\nu}(\alpha_{\mu,\nu})}\right] \approx -\frac{1}{2}\frac{\mu^2}{(1-\sqrt{\nu})^2}, \tag{7.128a}$$

$$\ln\left[e^{-\mu^2/4(1+\nu)} \Phi_{\frac{-\mu}{1+\nu}, \frac{2\nu}{1+\nu}}(\alpha_{\mu,\nu})\right] \approx -\frac{1}{2}\frac{\mu^2}{(1-\sqrt{\nu})^2}, \tag{7.128b}$$

and thus for $\mu \to -\infty$ we have

$$\ln\left[Z_\nu(\mu)\right] \approx -\frac{1}{2}\frac{\mu^2}{(1-\sqrt{\nu})^2}. \tag{7.129}$$

7.A.2 Consistency with Theorems 7.1 and 7.2

The small deviation analyses of Refs. [7.7,7.8] consider transformations with a fixed infidelity, bounded away from zero. Despite this, using the above expansion one can naïvely substitute $1 - F = \epsilon_n^-$ into the expressions for the optimal rates obtained within the small deviation regime. Whilst this analysis

is no longer rigorous, we will see that it gives a rate which is consistent with our rigorous results, Theorems 7.1 and 7.2.

Inverting the expansion around $\mu = -\infty$, we have

$$Z_\nu^{-1}(\epsilon) \approx \left|1 - \sqrt{\nu}\right| \sqrt{-2\ln\epsilon}, \tag{7.130}$$

for small positive ϵ. In particular, for $\epsilon_n^- := e^{-nt_n^2}$ we have

$$Z_\nu^{-1}(\epsilon_n^-) \simeq \left|1 - \sqrt{\nu}\right| \sqrt{2n}t_n. \tag{7.131}$$

Substituting the above into the results of Refs. [7.7, 7.8] yields the expressions for optimal rates from Theorems 7.1 and 7.2.

7.A.3 Expansion around $\mu = +\infty$.

The crossing point $\alpha_{\mu,\nu}$ is defined by the equation [7.7]

$$\frac{\phi(\alpha_{\mu,\nu})}{\phi_{\mu,\nu}(\alpha_{\mu,\nu})} = \frac{\Phi(\alpha_{\mu,\nu})}{\Phi_{\mu,\nu}(\alpha_{\mu,\nu})}. \tag{7.132}$$

As $\mu \to +\infty$, $\alpha_{\mu,\nu} \to -\infty$. We can now use the approximation $\Phi(x) \approx \phi(x)/x$ for $x \to -\infty$. Applying this to the above, we have

$$\frac{\phi(\alpha_{\mu,\nu})}{\phi_{\mu,\nu}(\alpha_{\mu,\nu})} \approx \frac{\phi(\alpha_{\mu,\nu})}{\phi_{\mu,\nu}(\alpha_{\mu,\nu})} \cdot \frac{\alpha_{\mu,\nu} - \mu}{\nu\alpha_{\mu,\nu}}, \tag{7.133}$$

which in turn implies that

$$\alpha_{\mu,\nu} \approx \frac{\mu}{1-\nu}. \tag{7.134}$$

Returning to the Rayleigh-normal distribution, we note that $\sqrt{1 - Z_\nu(\mu)}$ exponentially vanishes as $\alpha_{\mu,\nu} \to -\infty$, specifically

$$\ln\left[\sqrt{\Phi(\alpha_{\mu,\nu})\Phi_{\mu,\nu}(\alpha_{\mu,\nu})}\right] \approx -\frac{1+\nu}{(1-\nu)^2}\frac{\mu^2}{4}, \tag{7.135a}$$

$$\ln\left[e^{-\mu^2/4(1+\nu)}\Phi_{\frac{-\mu}{1+\nu},\frac{2\nu}{1+\nu}}(\alpha_{\mu,\nu})\right] \approx -\frac{1}{1+\nu}\frac{\mu^2}{4}. \tag{7.135b}$$

As such, we can see that the second expression dominates, leading us to conclude that, for $\mu \to +\infty$, the Rayleigh-normal distribution may be approximated as

$$\ln\left[1 - Z_\nu(\mu)\right] \approx -\frac{\mu^2}{4(1+\nu)}. \tag{7.136}$$

7.A.4 Conjectured infidelity analogue of Theorem 7.3

Similar to the case of the expansion around $\mu = -\infty$, we can use the expansion around $\mu = +\infty$ to (non-rigorously) obtain a conjecture for the analogue of Theorem 7.3 for error measured in infidelity. Inverting the expansion around $\mu = +\infty$, we find

$$Z_\nu^{-1}(\epsilon) \approx \sqrt{-4(1+\nu)\ln(1-\epsilon)}, \tag{7.137}$$

for ϵ close to 1. Specifically, for $\epsilon_n^+ = 1 - e^{-nt_n^2}$ we have

$$Z_\nu^{-1}(\epsilon_n^+) \approx \sqrt{4(1+\nu)n}t_n. \tag{7.138}$$

Substituting this approximation into the results of Refs. [7.7, 7.8], we have that the optimal rates for an infidelity error level of ϵ_n^+ are of the following form

$$R_n^{\text{ent}}(\epsilon_n) \simeq R_\infty^{\text{ent}} + \sqrt{\frac{2V(\boldsymbol{p})}{H(\boldsymbol{q})^2}}\sqrt{1+1/\nu^{\text{th}}}\, t_n, \tag{7.139a}$$

$$R_n^{\text{th}}(\epsilon_n) \simeq R_\infty^{\text{th}} + \sqrt{\frac{2V(\boldsymbol{p}||\boldsymbol{\gamma})}{D(\boldsymbol{q}||\boldsymbol{\gamma})^2}}\sqrt{1+1/\nu^{\text{th}}}\, t_n. \tag{7.139b}$$

7.B Compact convergence lemma

Lemma 7.12. *Let $\{f_n\}_n$ and $\{g_n\}_n$ be sequences of non-decreasing real functions, both of which point-wise converge to continuous functions f and g, respectively. If $f > g$, then $f_n|_X > g_n|_X$ eventually on all compact X.*

Proof. First we note (see, e.g., page 1 of Ref. [7.32]) that sequences of non-decreasing real functions which point-wise converge to continuous functions do so compactly. Let $\Delta_n : X \to \mathbb{R}$ be defined by $\Delta_n(x) = f_n(x) - g_n(x)$ for all $x \in X$. As $\Delta = f - g$, we know that $\Delta > 0$. Indeed, because Δ is a continuous function on a compact domain, the Extreme Value Theorem tells us that Δ is bounded away from zero, i.e., there exists an $\epsilon > 0$ such that $\Delta \geq \epsilon$. As $\Delta_n \to \Delta$ uniformly, we must eventually have that $\Delta_n \geq \Delta - \epsilon/2 \geq \epsilon/2$, and so $f_n|_X > g_n|_X$. □

Bibliography

7.1 Bob Coecke, Tobias Fritz, and Robert W Spekkens, "A mathematical theory of resources," Inform. Comput. **250**, 59–86 (2016).

7.2 E. Chitambar and G. Gour, "Quantum Resource Theories," arXiv:1806.06107 (2018).

7.3 Charles H Bennett, Herbert J Bernstein, Sandu Popescu, and Benjamin Schumacher, "Concentrating partial entanglement by local operations," Phys. Rev. A **53**, 2046 (1996).

7.4 Fernando G. S. L. Brandão, Michał Horodecki, Jonathan Oppenheim, Joseph M. Renes, and Robert W. Spekkens, "Resource theory of quantum states out of thermal equilibrium," Phys. Rev. Lett. **111**, 250404 (2013).

7.5 Michael A Nielsen, "Conditions for a class of entanglement transformations," Phys. Rev. Lett. **83**, 436 (1999).

7.6 F. G. S. L. Brandão, M. Horodecki, N. H. Y. Ng, J. Oppenheim, and S. Wehner, "The second laws of quantum thermodynamics," Proc. Natl. Acad. Sci. U.S.A. **112**, 3275 (2015).

7.7 Wataru Kumagai and Masahito Hayashi, "Second-order asymptotics of conversions of distributions and entangled states based on Rayleigh-normal probability distributions," IEEE Trans. Inf. Theory **63**, 1829–1857 (2017).

7.8 Christopher T Chubb, Marco Tomamichel, and Kamil Korzekwa, "Beyond the thermodynamic limit: finite-size corrections to state interconversion rates," arXiv:1711.01193 (2017a).

7.9 V Strassen, "Asymptotische Abschätzungen in Shannons Informationstheorie," in *Trans. Third Prague Conf. Inf. Theory* (Prague, 1962) pp. 689–723.

7.10 Yucel Altug and Aaron B. Wagner, "Moderate Deviations in Channel Coding," IEEE Trans. Inf. Theory **60**, 4417–4426 (2014).

7.11 C.T. Chubb, V.Y.F. Tan, and M. Tomamichel, "Moderate Deviation Analysis for Classical Communication over Quantum Channels," Commun. Math. Phys. **355** (2017b).

7.12 Hao-Chung Cheng and Min-Hsiu Hsieh, "Moderate Deviation Analysis for Classical-Quantum Channels and Quantum Hypothesis Testing," (2017), arXiv:1701.03195 .

7.13 Ryszard Horodecki, Paweł Horodecki, Michał Horodecki, and Karol Horodecki, "Quantum entanglement," Rev. Mod. Phys. **81**, 865–942 (2009).

7.14 T Baumgratz, M Cramer, and MB Plenio, "Quantifying coherence," Phys. Rev. Lett. **113**, 140401 (2014).

7.15 M. Horodecki and J. Oppenheim, "Fundamental limitations for quantum and nanoscale thermodynamics," Nat. Commun. **4**, 2059 (2013).

7.16 Paul G Kwiat, Salvador Barraza-Lopez, Andre Stefanov, and Nicolas Gisin, "Experimental entanglement distillation and hidden non-locality," Nature **409**, 1014 (2001).

7.17 Andreas Winter and Dong Yang, "Operational resource theory of coherence," Phys. Rev. Lett. **116**, 120404 (2016).

7.18 Kamil Korzekwa, Christopher T Chubb, and Marco Tomamichel, "Avoiding irreversibility: engineering resonant conversions of quantum resources," (In preparation).

7.19 Hiroyasu Tajima and Masahito Hayashi, "Finite-size effect on optimal efficiency of heat engines," Phys. Rev. E **96**, 012128 (2017).

7.20 Stephen D Bartlett, Terry Rudolph, and Robert W Spekkens, "Reference frames, superselection rules, and quantum information," Rev. Mod. Phys. **79**, 555 (2007).

7.21 Gilad Gour, "Quantum resource theories in the single-shot regime," Phys. Rev. A **95**, 062314 (2017).

7.22 Fernando GSL Brandão and Gilad Gour, "Reversible framework for quantum resource theories," Phys. Rev. Lett. **115**, 070503 (2015).

7.23 John Goold, Marcus Huber, Arnau Riera, Lídia del Rio, and Paul Skrzypczyk, "The role of quantum information in thermodynamics – a topical review," J. Phys. A **49**, 143001 (2016).

7.24 Amir Dembo and Ofer Zeitouni, *Large Deviations Techniques and Applications*, 2nd ed., Stochastic Modelling and Applied Probability (Springer, 1998).

7.25 Shuanping Du, Zhaofang Bai, and Yu Guo, "Conditions for coherence transformations under incoherent operations," Phys. Rev. A **91**, 052120 (2015).

7.26 Albert W Marshall, Ingram Olkin, and Barry C Arnold, *Inequalities: theory of majorization and its applications*, Vol. 143 (Springer, 1979).

7.27 Ernst Ruch, Rudolf Schranner, and Thomas H. Seligman, "The mixing distance," J. Chem. Phys. **69** (1978).

7.28 D. Janzing, P. Wocjan, R. Zeier, R. Geiss, and Th. Beth, "Thermodynamic cost of reliability and low temperatures: tightening Landauer's principle and the second law," Int. J. Theor. Phys. **39**, 2717–2753 (2000).

7.29 Matteo Lostaglio, "Thermodynamic laws for populations and quantum coherence: A self-contained introduction to the resource theory approach to thermodynamics," arXiv:1807.11549 (2018).

7.30 Christopher A Fuchs and Jeroen Van De Graaf, "Cryptographic distinguishability measures for quantum-mechanical states," IEEE Trans. Inf. Theory **45**, 1216–1227 (1999).

7.31 Johan Åberg, "Truly work-like work extraction via a single-shot analysis," Nat. Commun. **4**, 1925 (2013).

7.32 Sidney I Resnick, *Extreme values, regular variation and point processes* (Springer, 2013).

Chapter 8

Avoiding irreversibility: engineering resonant conversions of quantum resources

Kamil Korzekwa[1], Christopher T. Chubb[1], and Marco Tomamichel[2]
[1] *Centre for Engineered Quantum Systems, School of Physics, University of Sydney, Sydney NSW 2006, Australia*
[2] *Centre for Quantum Software and Information, School of Software, University of Technology Sydney, Sydney NSW 2007, Australia*

To appear in Physical Review Letters
ArXiv: 1810.02366

Abstract

We identify and explore the intriguing property of resource resonance arising within resource theories of entanglement, coherence and thermodynamics. While the theories considered are reversible asymptotically, the same is generally not true in realistic scenarios where the available resources are bounded. The finite-size effects responsible for this irreversibility could potentially prohibit small quantum information processors or thermal machines from achieving their full potential. Nevertheless, we show here that by carefully engineering the resource interconversion process any such losses can be greatly suppressed. Our results are predicted by higher order expansions of the trade-off between the rate of resource interconversion and the achieved fidelity, and are verified by exact numerical optimisations of appropriate approximate majorisation conditions.

8.1 Introduction

Rapid progress in experimental techniques to control intermediate-scale quantum systems [8.1] may soon bring the advent of new technologies that will utilise quantum effects to overcome current limitations of electromechanical systems and information processors [8.2]. From a theoretical perspective, the first step to achieve this is to identify which components of quantum theory can provide such an

advantage, i.e., to recognise what actually constitutes quantum resources. Once identified, it is then crucial to understand when different resources can be interconverted, and how efficiently this can be done. The ultimate goal is to find *reversible* resource transformations, i.e., transformations that can be perfectly reversed and that do not dissipate any resources. Such questions can be treated in great generality under the umbrella of quantum resource theories [8.3, 8.4].

With the above applications in mind, we focus on the question of reversibility of resource interconversion for intermediate-scale quantum systems, specifically in the resource theories of entanglement [8.5] and coherence [8.6] (for pure state transformations), and thermodynamics [8.7] (for energy-incoherent transformations). In particular, we ask under what precise circumstances we can engineer reversible, and thus dissipationless, transformations between a relatively small number of resource states. In the above resource theories reversibility always holds in the asymptotic regime, where one assumes access to an infinite source of resource states [8.8]. This means that, in the appropriate asymptotic limit, we can reverse the conversion process perfectly, recovering exactly the resource states we started with. Such an assumption, however, is unjustified in many practically relevant and fundamentally interesting scenarios. On the one hand, quantum resources like entanglement and coherence will only be available in small amounts in the near future; on the other hand, in quantum thermodynamics we want to explore how the known macroscopic laws change when we go beyond the thermodynamic limit and consider thermal processes involving a finite number of particles [8.9].

A first attempt to address irreversibility beyond the asymptotic limit was made in [8.10] in the context of entanglement concentration. Then, in [8.12] the authors developed a mathematical framework that allows us to study corrections to reversible entanglement transformations arising due to finite-size effects, while the conditions for these corrections to vanish were discussed in [8.11]. The framework of [8.12] has recently been extended to the resource theory of thermodynamics by the present authors [8.13]. On a technical level all these prior works deal with conversion processes that asymptotically recover the resource states only up to a constant error. The accompanying work [8.14] takes a complementary approach that allows us to consider more relevant processes with asymptotically vanishing error (that are still asymptotically reversible).

These analytical results together with the numerical results presented here reveal the phenomenon of *resource resonance*, a surprising feature of resource interconversion that ensures that resource dissipation can be avoided by carefully engineering small quantum systems. Exploring this phenomenon and discussing its potential impact on quantum information processing and thermodynamic protocols will be the main focus of this paper.

8.2 Motivating example

Before we delve into the details and present our results in full generality, let us first illustrate their spirit using the following simple example. Consider a heat engine, with a finite-size working body consisting of $n = 200$ non-interacting two-level systems and operating between two thermal baths at temperatures $T_h > T_c$, performing work on a battery system initially in the ground state. As we explain later in the text, such a process can be elegantly phrased as an interconversion problem within the resource theory of thermodynamics, with the working body at cold temperature T_c acting as a non-equilibrium resource in the presence of the hot bath at temperature T_h. While the working body heats up from T_c to $T_{c'}$, part of its resource content can be converted into work by exciting the battery

system. Now, a perfect engine with an infinite working body, $n \to \infty$, and constantly operating at the Carnot efficiency would extract the amount of work nW_C equal to the free energy change of the working body. However, since n is small, the energy fluctuations of the working body are not negligible compared to the average energy, and thus the engine operates beyond the thermodynamic limit. We can then expect two kinds of effects. On the one hand, the quality of work [8.15, 8.16] will not be perfect, as some of the energy fluctuations will be transferred to the battery. In Fig. 1a we present the optimal work quality as a function of T_c and $T_{c'}$ for an engine extracting $W = 0.95W_C$ per qubit of the working body. On the other hand, if we demand the work quality to be above some threshold level, the optimal efficiency of the engine may be affected so that it cannot achieve the Carnot limit. In Fig. 1b we plot the optimal fraction of W_C that can be extracted when its quality is bounded by a constant, again as a function of T_c and $T_{c'}$.

In both plots we see clear resonant lines, indicating near-perfect quality and near-Carnot efficiency corresponding to reversible (and thus dissipationless) processes, obtained while operating well outside the asymptotic regime with a finite-size working body. As we argue in this paper, it is in fact a genuine feature of all resource theories whose single-shot transformation laws are based on majorisation relation. These include the considered resource theories of entanglement, coherence and thermodynamics, but also applies to, e.g., the studies of randomness conversions [8.12]. In all these cases, we find that if a pair of states satisfies a particular *resonance condition*, one can achieve interconversion with dramatically reduced losses, i.e., transformation that is arbitrarily close to reversible even for finite number n of resource states being processed. In what follows we first introduce necessary concepts, then state the resonance condition, and finally discuss how it can be employed to avoid irreversibility.

8.3 Setting the scene

The interconversion and dissipation of general quantum resources can be suitably analysed within a resource-theoretic framework [8.3, 8.4]. In a resource theory we allow only a subset of all possible quantum operations, and this in turn restricts the quantum states that can be prepared — the so-called free states of the theory. Natural restrictions may arise from practical difficulties, e.g., when preparing a system in a superposition of particular states is experimentally challenging, but may also be of fundamental nature, as with the laws of thermodynamics constraining possible transformations to preserve energy and increase entropy. A quantum resource is any quantum system that, in conjunction with the allowed operations, allows one to overcome these restrictions. A paradigmatic example is provided by the resource theory of entanglement [8.5]: Alice, facing the restriction of not having access to a quantum channel to Bob, is unable to share an entangled quantum state $|\psi\rangle$ with him. However, if they are in possession of a maximally entangled state acting as a resource, they can use it to teleport Bob's share of $|\psi\rangle$ using only local operations and classical communication.

In this work we focus on the interconversion problem within the resource theories of entanglement, coherence and thermodynamics. These are defined via the relevant sets of free operations and free states, the latter of which we will collectively denote by ω. More precisely, in entanglement theory one is restricted to local operations and classical communication (LOCC) and the allowed free states are given by separable states [8.5]; in coherence theory one is restricted to incoherent operations and the allowed free states are incoherent in the preferred basis [8.6]; finally, in the resource theory of thermodynamics (with respect to a fixed background temperature $T = 1/\beta$) one is restricted to thermal

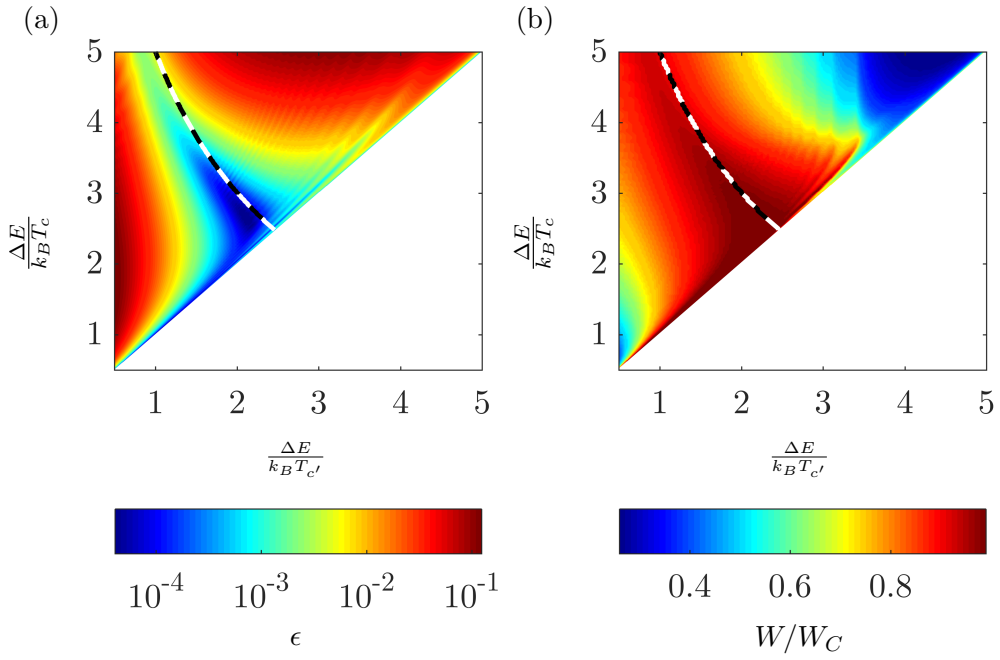


Figure 8.1: *Resonance in work extraction.* Performance of the heat engine with a working body consisting of $n = 200$ non-interacting qubits, each with energy gap ΔE. Hot bath temperature it set by $k_B T_h = 10\Delta E$ (where k_B is the Boltzmann constant), the working body is initially at cold bath temperature T_c and heats up to $T_{c'}$ in the process. (a) The optimal quality of work, measured by the infidelity ϵ between the final and excited battery state, while extracting $W = 0.95 W_C$ per qubit. (b) The optimal fraction W_C that can be extracted per qubit when the quality of work is bounded by $\epsilon < 10^{-3}$. The dashed line in both plots indicates the position of the resonance predicted by Eq. (8.9).

operations and the only free state is a thermal equilibrium state γ at temperature T [8.17]. Rather surprisingly, all these three prominent examples of resource theories are formally very strongly related, as single-shot interconversion in each of them is ruled by a majorisation partial order [8.18,8.19] (or its variant known as thermo-majorisation [8.7]). This not only allows for a unified treatment, but also for a simplified representation of initial and target quantum states, ρ and σ, as probability distributions, $\boldsymbol{p}$ and $\boldsymbol{q}$.

Before we describe this mapping, let us first explain the notions of multi-copy and approximate transformations. When considering transformations between multiple copies of initial and target states, $\rho^{\otimes n}$ and $\sigma^{\otimes Rn}$, the number of initial and target states will generally not be the same, i.e., conversion rate R may differ from 1. However, since one can always append any number of free states ω to both the initial and target states, without loss of generality one can focus on transformations between total initial and target states,

$$\tau_i = \rho^{\otimes n} \otimes \omega^{\otimes Rn}, \quad \tau_f = \sigma^{\otimes Rn} \otimes \omega^{\otimes n}. \tag{8.1}$$

Now, it may happen that for given τ_i and τ_f the interconversion is impossible, but there exists a state $\tilde{\tau}_f$ that is ϵ-close to τ_f and such that the interconversion between τ_i and $\tilde{\tau}_f$ is possible. We then say that an approximate transformation is possible with the error level ϵ, as quantified by the infidelity

$1-F$ with $F=[\mathrm{Tr}(\tau_f^{\frac{1}{2}}\tilde{\tau}_f\tau_f^{\frac{1}{2}})^{\frac{1}{2}}]^2$.

We then say that an approximate transformation is possible with the error level ϵ quantified by the infidelity between target state τ_f and final state $\tilde{\tau}_f$.

As mentioned above, the interconversion conditions for all considered resource theories can be expressed using theory-dependent representations of initial and target quantum states. For entanglement theory, given initial and target pure bipartite states, $\rho=|\Psi\rangle\langle\Psi|$ and $\sigma=|\Phi\rangle\langle\Phi|$, with the Schmidt decomposition given by $|\Psi\rangle=\sum_i a_i|\psi_i\rangle\otimes|\psi_i\rangle$ and $|\Phi\rangle=\sum_i b_i|\phi_i\rangle\otimes|\phi_i\rangle$, we can represent them via probability distributions

$$p_i=|a_i|^2,\quad q_i=|b_i|^2. \tag{8.2}$$

For coherence theory with respect to a fixed basis $\{|i\rangle\}$ one can represent pure initial and target states, $\rho=|\psi\rangle\langle\psi|$ and $\sigma=|\phi\rangle\langle\phi|$, using

$$p_i=|\langle i|\psi\rangle|^2,\quad q_i=|\langle i|\phi\rangle|^2. \tag{8.3}$$

Finally, in the resource theory of thermodynamics, the initial and target energy-incoherent mixed states ρ and σ can be represented by

$$p_i=\langle E_i|\rho|E_i\rangle,\quad q_i=\langle E_i|\sigma|E_i\rangle, \tag{8.4}$$

where $\{|E_i\rangle\}$ denotes the energy eigenbasis of the system. We will denote distributions representing free states ω by $\boldsymbol{f}$. In entanglement and coherence theories these are represented by sharp probability distributions $\boldsymbol{s}$ with a single non-zero entry; whereas in the thermodynamic case $\boldsymbol{f}$ is given by a thermal Gibbs distribution $\boldsymbol{\gamma}$ with $\gamma_i\propto\exp(-\beta E_i)$.

8.4 Reversibility condition

The optimal conversion rate R between n copies of an initial state ρ, represented by $\boldsymbol{p}$, and Rn copies of the target state σ, represented by $\boldsymbol{q}$, can now be expressed in a unified way for all considered resource theories using a general asymptotic expansion,

$$R\simeq R_\infty(\boldsymbol{p},\boldsymbol{q})-r(\boldsymbol{p},\boldsymbol{q},n,\epsilon). \tag{8.5}$$

In the above, R_∞ denotes the rate achieved in the asymptotic limit, whereas r is the correction term, which depends on the number of initial states n and the accepted error level ϵ, and vanishes for $n\to\infty$. Expressions for asymptotic rates R_∞ are well known for the studied resource theories, with $R_\infty=H(\boldsymbol{p})/H(\boldsymbol{q})$ for coherence and entanglement [8.20], and $R_\infty=D(\boldsymbol{p}||\boldsymbol{\gamma})/D(\boldsymbol{q}||\boldsymbol{\gamma})$ for thermodynamics [8.21], where H and D denote the Shannon entropy and the relative entropy, respectively. What is important for the current analysis, is the fact that $R_\infty(\boldsymbol{p},\boldsymbol{q})=1/R_\infty(\boldsymbol{q},\boldsymbol{p})$, i.e., the asymptotic rate of a given process is equal to the inverse of the asymptotic rate for the reverse process. As a result, every resource interconversion is reversible in the asymptotic limit, so that no dissipation of resources takes place.

In the accompanying paper [8.14] we find the next order in the asymptotic expansion of the correction term r for all considered resource theories in the *moderate deviation* regime, i.e., when one

demands both error and correction term to vanish with growing n [1]. The crucial observation is that for all α the correction term r depends on a parameter $\nu = \nu(\boldsymbol{p}, \boldsymbol{q})$, which we will refer to as the *irreversibility parameter*. Its importance lies in the fact that as $\nu \geq 0$ gets closer to 1, the magnitude of r diminishes, and in the limit $\nu = 1$ it vanishes. Therefore, in the moderate deviation approximation, states satisfying $\nu = 1$ can be interconverted with vanishing error at the asymptotic rate, $R = R_\infty$, even for finite n, and thus are reversibly interconvertible.

For entanglement and coherence transformations ν is given by [8.12]

$$\nu = \frac{V(\boldsymbol{p})/H(\boldsymbol{p})}{V(\boldsymbol{q})/H(\boldsymbol{q})}, \tag{8.6}$$

with V denoting entropy variance [8.23, 8.24],

$$V(\boldsymbol{p}) = \sum_i p_i \left(\log p_i + H(\boldsymbol{p})\right)^2, \tag{8.7}$$

and the correction term reads

$$r(\boldsymbol{p}, \boldsymbol{q}, n, e^{-nt_n^2}) = \sqrt{\frac{2V(\boldsymbol{p})}{H(\boldsymbol{q})^2}}\left|1 - 1/\sqrt{\nu}\right| t_n. \tag{8.8}$$

with t_n denoting arbitrary moderate sequence, i.e., $t_n \to 0$ and $\sqrt{n}t_n \to +\infty$ as $n \to \infty$. Note that since the Shannon entropy H quantifies the average (asymptotic) resource content of a state, entropy variance V can be understood as quantifying its fluctuations. Indeed, V vanishes only for distributions that are uniform on their support, which correspond to maximally coherent and maximally entangled states. Hence, the ratio V/H tells us about relative strength of resource fluctuations in a single copy of a given state, and so the resonance condition is satisfied for states with equal relative resource fluctuations.

Similarly, in the thermodynamic case we have [8.13]

$$\nu = \frac{V(\boldsymbol{p}||\boldsymbol{\gamma})/D(\boldsymbol{p}||\boldsymbol{\gamma})}{V(\boldsymbol{q}||\boldsymbol{\gamma})/D(\boldsymbol{q}||\boldsymbol{\gamma})}, \tag{8.9}$$

with V denoting relative entropy variance [8.23],

$$V(\boldsymbol{p}||\boldsymbol{\gamma}) = \sum_i p_i \left(\log \frac{p_i}{\gamma_i} - D(\boldsymbol{p}||\boldsymbol{\gamma})\right)^2, \tag{8.10}$$

and the correction term reads

$$r(\boldsymbol{p}, \boldsymbol{q}, n, e^{-nt_n^2}) = \sqrt{\frac{2V(\boldsymbol{p}||\boldsymbol{\gamma})}{D(\boldsymbol{q}||\boldsymbol{\gamma})^2}}\left|1 - 1/\sqrt{\nu}\right| t_n. \tag{8.11}$$

Note that D, quantifying resource content in the asymptotic (thermodynamic) limit, can be interpreted as non-equilibrium free energy of a state [8.21]. Thus, the relative entropy variance describes fluctuations of free energy and, again, states with equal relative fluctuations are in resonance. Moreover, as noted in [8.13], $V(\boldsymbol{\gamma}'||\boldsymbol{\gamma})$ (for $\boldsymbol{\gamma}'$ being a thermal state at temperature $T' \neq T$) is proportional to the heat capacity at T' and to the squared Carnot factor for T and T'.

[1] Complementary to these results, are expressions for r in the *small deviation* regime (when one allows for constant error ϵ and vanishing r), obtained in [8.12] and our previous work [8.13], for resource theories of entanglement and thermodynamics, respectively.

8.5 Resource resonance

The moderate deviation analysis of the asymptotic expansion for the resource conversion rate predicts a resonance phenomenon. However, the question that remains open is: how well does this analytical result capture the actual errors and rates arising in resource interconversion of a small number n of resource states involved? To address this problem, we numerically investigate the interconversion problem for small-dimensional systems, and compare the results with our analytical predictions. More precisely, building on an algorithm developed by the present authors in [8.13], we find the actual optimal conversion rates for a fixed error level, and also the actual error level for a fixed conversion rate. The details can be found in the Supplementary Material, and here we focus on presenting a few illustrative examples.

We start by considering the following scenario: imagine one has access to n systems, each initially in a state ρ_1 or ρ_2, and wants to transform them to a target state σ. Moreover, assume that the asymptotic resource values of ρ_1 and ρ_2 are equal, i.e., asymptotically one can obtain the same number of copies $R_\infty n$ of σ from either n copies of ρ_1 or n copies of ρ_2. One can achieve this asymptotic conversion rate also for finite n, but for the price of error ϵ. This error depends on the irreversibility parameter ν, and thus for finite n the states ρ_1 and ρ_2 no longer have the same value. Crucially, however, it may happen that they are incompatible with σ in opposite ways, such that the irreversibility parameter for ρ_1 and σ is smaller than 1, while for ρ_2 and σ it is larger than 1. Thus, by taking λn copies of ρ_1 and $(1-\lambda)n$ copies of ρ_2 one can tune the initial state to be in resonance with the target state, i.e., to have irreversibility parameter close to 1. In Fig. 2a we show how such tuning can reduce the infidelity of entanglement transformation by several orders of magnitude. Note that this effect is much stronger than the expected increase in fidelity due to the increased number n of processed states that can be observed for $\lambda \in \{0, 1\}$.

We also consider an alternative situation where, instead of demanding conversion at the asymptotic rate, we enforce the error to be below some fixed threshold value ϵ. For the sake of simplicity, we can again focus on a set of initial states $\{\rho_i\}$ that are asymptotically equivalent, and ask how many copies of the target state σ can we optimally get for $\rho_i^{\otimes n}$, with error not exceeding ϵ. Now, based on our analytical results, we expect that as n grows the conversion rate R will approach the asymptotic value R_∞ quicker for states ρ_i that are closer to resonance with σ. Indeed, this is the case, as the numerical results presented in Fig. 2b show.

Finally, let us come back to our initial motivating example (see Fig. 8.1) to explain it in more detail. First of all, we note that it is a particular instance of the paradigmatic thermodynamic protocol of work extraction [8.7, 8.15, 8.25], where the state we extract work from is given by a thermal state at a colder temperature than the background temperature T_h. More precisely, our aim is to transform n copies of $\gamma_c \otimes |0\rangle\langle 0|$ (where the first system refers to the working body and second to the battery consisting of n qubits) into n copies of $\gamma_{c'} \otimes |1\rangle\langle 1|$, with subscripts c and c' denoting initial and final temperatures of the working body, T_c and $T_{c'}$, and $|1\rangle\langle 1|$ being the excited state of the battery with energy W. The amount of extracted work is then given by nW and its quality is measured by the infidelity between the battery's final state and its target state $|1\rangle\langle 1|^{\otimes n}$. It is known that in the asymptotic limit the optimal thermodynamic process extracts nW_C with $W_C = k_B T_h [D(\boldsymbol{\gamma}_c||\boldsymbol{\gamma}_h) - D(\boldsymbol{\gamma}_{c'}||\boldsymbol{\gamma}_h)]$ [8.21], which coincides with the amount of work that would be extracted by an engine operating at Carnot efficiency between an infinite bath at fixed temperature T_h and a finite working body that heats up during the process from T_c to $T_{c'}$ [8.13]. For small n, however, fluctuations can decrease this optimal amount

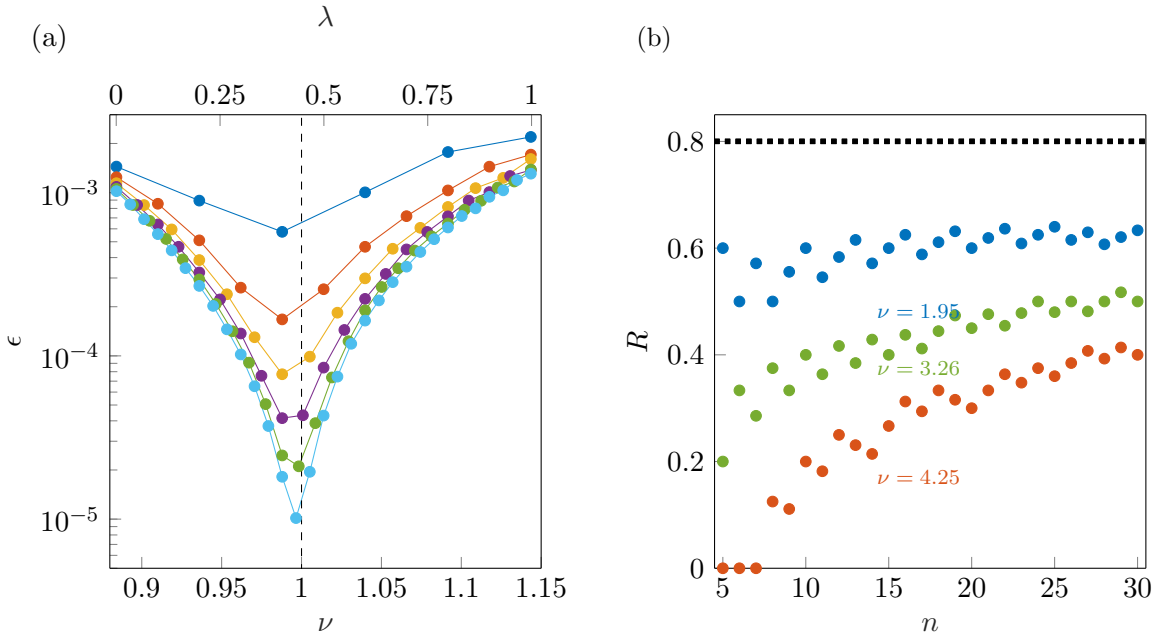


Figure 8.2: *Tuning resources to resonance.* State interconversion under LOCC for a bipartite system consisting of n pairs of qutrits. (a) The infidelity ϵ between the target state $|\Phi\rangle^{\otimes n}$ and the optimal final state obtained from $|\Psi_1\rangle^{\otimes \lambda n} \otimes |\Psi_2\rangle^{\otimes (1-\lambda)n}$. Different plots correspond to varying numbers of interconverted states $n \in \{5, 10, \ldots, 30\}$, from top to bottom. The location of the resonant mixing factor λ_* (dashed line) can be found using Eq. (8.6). (b) The optimal conversion rate R between the initial state $|\Psi_i\rangle$ and the target state $|\Phi\rangle^{\otimes Rn}$ with transformation infidelity bounded by $\epsilon < 0.01$. Different plots correspond to initial states with equal asymptotic conversion rate R_∞ (dashed line), but varying irreversibility parameters ν. The Schmidt coefficients of all states can be found in the Supplementary Material.

(alternatively, decrease the quality of extracted work), as already discussed before and presented in Fig. 8.1. It is also important to note that, since $V(\boldsymbol{\gamma}_c||\boldsymbol{\gamma}_h)$ is positive and vanishes for $T_c = T_h$ and $T_c = 0$, for all T_h there exist resonant pairs of temperatures T_c and $T_{c'}$, for which fluctuations vanish.

8.6 Conclusions

We have shown that finite-size analysis can bring qualitatively new insights into the nature of interconversion processes within resource theories of entanglement, coherence and thermodynamics. This is in rather stark contrast to the small or moderate deviation analysis for information theoretic tasks based on quantum hypothesis testing, including many channel coding problems (see, e.g., [8.26–8.30]), where we instead see a behaviour that can be understood as a rather immediate manifestation of the central limit theorem and its moderate deviation analogue. Here, we have demonstrated that the predicted *finite resource resonance* is clearly visible numerically and is not dominated by higher order asymptotics neglected in the analytical approximations. This makes us believe that the resonance effect could also be observed in noisy intermediate-scale quantum devices in the near future. Moreover, we explained how the predicted phenomenon can be employed to significantly improve the quality of the conversion

process.

On the one hand, our result may serve as a guiding principle for devising optimal protocols for resource conversion that would minimise losses by operating near resonance. On the other hand, it strongly motivates the investigation of the second order asymptotic corrections to interconversion rates in other resource theories. In particular, one could look for such corrections within the resource theory of $U(1)$ asymmetry [8.31, 8.32], which could contribute to our understanding of quantum thermodynamics beyond energy-incoherent states [8.33, 8.34]. Finally, while it is true that the resonance phenomenon is strongly related to the majorisation condition, it is possible that similar behaviour can be observed in other quantum information processing tasks [8.35].

Acknowledgements

We thank Masahito Hayashi for helpful comments. We acknowledge support from the ARC via the Centre of Excellence in Engineered Quantum Systems (EQUS), project number CE110001013. CC also acknowledges support from the Australian Institute for Nanoscale Science and Technology Postgraduate Scholarship (John Makepeace Bennett Gift). MT acknowledges an Australian Research Council Discovery Early Career Researcher Award, project no. DE160100821.

8.A Supplementary Material

Here we explain how the numerical data presented in the plots of the main text was obtained. First, note that the approximate interconversion conditions are ultimately specified by approximate majorisation and thermo-majorisation relations. In order to find the final distribution that is closest in infidelity distance to the target state, and at the same time satisfies the required majorisation relation, we employ the algorithm developed in our previous work [8.13] (and conceptually based on [8.36]). The precise details can be found in Appendix C of [8.13] (the `python` code is also attached to the arXiv submission of [8.13]).

The data presented in Figs. 1a and 1b was obtained in the following way. First, the temperature and energy units were chosen such that the Boltzmann constant $k_B = 1$, and the energy difference between the ground and excited state of the qubit $\Delta E = 1$. Thus, thermal distributions at temperature T_x of qubits comprising the working body are given by

$$\boldsymbol{\gamma}_x = \frac{1}{1 + e^{-1/T_x}} \left[1, e^{-1/T_x}\right]. \tag{8.12}$$

The hot temperature was set to $T_h = 10$, while the initial and final cold temperatures, T_c and $T_{c'}$, varied between 0.5 and 5 (with $T_{c'} > T_c$). Now, for each pair of points $(T_c, T_{c'})$ the following free energy difference was calculated,

$$W_C(T_c, T_{c'}) := T_h[D(\boldsymbol{\gamma}_c||\boldsymbol{\gamma}_h) - D(\boldsymbol{\gamma}_{c'}||\boldsymbol{\gamma}_h)]. \tag{8.13}$$

Note that this is the optimal amount of work that can be extracted in the asymptotic limit per one qubit of the working body. The process of extracting work W per copy is modelled by the following interconversion process

$$(\boldsymbol{\gamma}_c \otimes [1,0])^{\otimes n} \longrightarrow (\boldsymbol{\gamma}_{c'} \otimes [0,1])^{\otimes n}, \tag{8.14}$$

where the second system is a qubit battery with the energy difference between the ground and excited state set to W. For Fig. 1a we set $W = 0.95W_C$ and run the algorithm yielding the optimal infidelity ϵ between the right hand side of Eq. (8.14) (the target state) and a distribution thermo-majorised by the left hand side of Eq. (8.14) (the final state). For Fig. 1b we set $W = xW_C$ with x varying from 1 to 0, until the same algorithm does not yield infidelity ϵ below the threshold value set to 10^{-3}. For both plots we set $n = 200$.

We now proceed to Figs. 2a and 2b. Pure state transformations under LOCC are fully determined by majorisation relations between Schmidt coefficients of the transformed states. Let us then denote the Schmidt coefficients of a pure bipartite state $|\Psi\rangle$ by $\boldsymbol{\zeta}_\Psi$. The data presented in Fig. 2a was obtained for initial and target states characterised by the following Schmidt coefficients (up to 4 significant digits):

$$\boldsymbol{\zeta}_{\Psi_1} = [0.4309, 0.4300, 0.1391], \quad H(\boldsymbol{\zeta}_{\Psi_1}) = 1, \quad V(\boldsymbol{\zeta}_{\Psi_1}) = 0.1529, \tag{8.15a}$$
$$\boldsymbol{\zeta}_{\Psi_2} = [0.5499, 0.2300, 0.2201], \quad H(\boldsymbol{\zeta}_{\Psi_2}) = 1, \quad V(\boldsymbol{\zeta}_{\Psi_2}) = 0.1977, \tag{8.15b}$$
$$\boldsymbol{\zeta}_{\Phi_1} = [0.5121, 0.3300, 0.1579], \quad H(\boldsymbol{\zeta}_{\Phi}) = 1, \quad V(\boldsymbol{\zeta}_{\Phi}) = 0.1729. \tag{8.15c}$$

Similarly, Fig. 2b was obtained using:

$$\boldsymbol{\zeta}_{\Psi_1} = [0.5436, 0.4264, 0.0300], \quad H(\boldsymbol{\zeta}_{\Psi_1}) = 0.8, \quad V(\boldsymbol{\zeta}_{\Psi_1}) = 0.2406, \tag{8.16a}$$
$$\boldsymbol{\zeta}_{\Psi_2} = [0.6594, 0.2806, 0.0600], \quad H(\boldsymbol{\zeta}_{\Psi_2}) = 0.8, \quad V(\boldsymbol{\zeta}_{\Psi_2}) = 0.4024, \tag{8.16b}$$
$$\boldsymbol{\zeta}_{\Psi_3} = [0.7119, 0.1481, 0.1400], \quad H(\boldsymbol{\zeta}_{\Psi_3}) = 0.8, \quad V(\boldsymbol{\zeta}_{\Psi_3}) = 0.5236, \tag{8.16c}$$
$$\boldsymbol{\zeta}_{\Phi} = [0.4514, 0.4086, 0.1400], \quad H(\boldsymbol{\zeta}_{\Phi}) = 1, \quad V(\boldsymbol{\zeta}_{\Phi}) = 0.1541. \tag{8.16d}$$

As in the thermodynamic case, here also we employed the algorithm yielding the optimal infidelity between the final and target states.

Bibliography

8.1 J. Preskill, Quantum **2**, 79 (2018).

8.2 J. P. Dowling and G. J. Milburn, Philos. Trans. Royal Soc. A **361**, 1655 (2003).

8.3 B. Coecke, T. Fritz, and R. W. Spekkens, Inform. Comput. **250**, 59 (2016).

8.4 E. Chitambar and G. Gour, arXiv:1806.06107 (2018).

8.5 R. Horodecki, P. Horodecki, M. Horodecki, and K. Horodecki, Rev. Mod. Phys. **81**, 865 (2009).

8.6 T. Baumgratz, M. Cramer, and M. Plenio, Phys. Rev. Lett. **113**, 140401 (2014).

8.7 M. Horodecki and J. Oppenheim, Nat. Commun. **4**, 2059 (2013).

8.8 F. G. Brandão and G. Gour, Phys. Rev. Lett. **115**, 070503 (2015).

8.9 J. Goold, M. Huber, A. Riera, L. del Rio, and P. Skrzypczyk, J. Phys. A **49**, 143001 (2016).

8.10 W. Kumagai and M. Hayashi, Phys. Rev. Lett. **111**, 130407 (2013).

8.11 K. Ito, W. Kumagai, and M. Hayashi, Phys. Rev. A **92**, 052308 (2015).

8.12 W. Kumagai and M. Hayashi, IEEE Trans. Inf. Theory **63**, 1829 (2017).

8.13 C. T. Chubb, M. Tomamichel, and K. Korzekwa, arXiv:1711.01193 (2017).

8.14 C. T. Chubb, M. Tomamichel, and K. Korzekwa, arXiv:1809.07778 (2018).

8.15 J. Åberg, Nat. Commun. **4**, 1925 (2013).

8.16 N. H. Y. Ng, M. P. Woods, and S. Wehner, New J. Phys. **19**, 113005 (2017).

8.17 D. Janzing, P. Wocjan, R. Zeier, R. Geiss, and T. Beth, Int. J. Theor. Phys. **39**, 2717 (2000).

8.18 M. A. Nielsen, Phys. Rev. Lett. **83**, 436 (1999).

8.19 S. Du, Z. Bai, and Y. Guo, Phys. Rev. A **91**, 052120 (2015).

8.20 C. H. Bennett, H. J. Bernstein, S. Popescu, and B. Schumacher, Phys. Rev. A **53**, 2046 (1996).

8.21 F. G. S. L. Brandão, M. Horodecki, J. Oppenheim, J. M. Renes, and R. W. Spekkens, Phys. Rev. Lett. **111**, 250404 (2013).

8.22 Complementary to these results, are expressions for r in the *small deviation* regime (when one allows for constant error ϵ and vanishing r), obtained in [8.12] and our previous work [8.13], for resource theories of entanglement and thermodynamics, respectively.

8.23 V. Strassen, in *Trans. Third Prague Conf. Inf. Theory* (Prague, 1962) pp. 689–723.

8.24 M. Hayashi, IEEE Trans. Inf. Theory **54**, 4619 (2008).

8.25 P. Skrzypczyk, A. J. Short, and S. Popescu, Nat. Commun. **5**, 4185 (2014).

8.26 M. Tomamichel and M. Hayashi, IEEE Trans. Inf. Theory **59**, 7693 (2013).

8.27 K. Li, Ann. Stat. **42**, 171 (2014).

8.28 M. Tomamichel and V. Y. F. Tan, Commun. Math. Phys. **338**, 103 (2015).

8.29 C. T. Chubb, V. Y. F. Tan, and M. Tomamichel, Commun. Math. Phys. **355**, 1283 (2017).

8.30 H.-C. Cheng and M.-H. Hsieh, IEEE Trans. Inf. Theory **64**, 1385 (2018).

8.31 I. Marvian, *Symmetry, Asymmetry and Quantum Information*, Ph.D. thesis, University of Waterloo (2012).

8.32 I. Marvian and R. W. Spekkens, New J. Phys. **15**, 033001 (2013).

8.33 M. Lostaglio, K. Korzekwa, D. Jennings, and T. Rudolph, Phys. Rev. X **5**, 021001 (2015).

8.34 P. Ćwikliński, M. Studziński, M. Horodecki, and J. Oppenheim, Phys. Rev. Lett. **115**, 210403 (2015).

8.35 M. Tomamichel, *Quantum information processing with finite resources: Mathematical foundations*, Springer Briefs in Mathematical Physics, Vol. 5 (Springer, 2015).

8.36 G. Vidal, D. Jonathan, and M. Nielsen, Phys. Rev. A **62**, 012304 (2000).

Chapter 9

Conclusion

In this thesis we looked at the role of noise in quantum information processing. In Part I this was considered from the perspective of quantum error correction, while in Part II we considered fundamental resource-error trade-offs of problem in quantum communication and quantum resource theories.

Quantum error correcting codes are typically analysed with respect to simple toy models of noise. A common theme in both Chapters 2 and 3 was extending these analyses to more realistic noise models. Specifically in Chapter 2 we provide a multitude of techniques for studying noise which is not spatially iid, but has some local correlations, and in Chapter 3 we analyse the surface code under the influence of dephasing-biased noise.

A natural application of these techniques would be in the design of codes which not only take into account—but exploit—the features of realistic noise processes, such as biases and correlations. Follow the ideas of Chapter 3, we could start by considering modifications of existing codes which help tailor them to certain noise models. The correlated decoder discussed in Chapter 2 would be a useful tool for this, as it provides a method of approximating the optimal decoder even when the performance of a code with respect to a given error model is not well understood.

One significant limitation of the methods considered in Chapters 2 and 3 is that they only apply to stochastic Pauli noise. Within the framework of stabiliser/subsystem codes, this is a particularly convenient noise model both for numerical and theoretical analysis, but there exist important noise models which lie outside this framework, such as amplitude damping or coherent errors. If it can be extended, a statistical mechanical mapping outside of Pauli noise would be a very significant tool for studying the performance of codes under such noise models.

In Chapter 4 we consider approximate quantum error correction, in codes possessing logical operators which only satisfy approximate commutation relations. Whilst we have shown that such operators imply the existence of an underlying code, how one would utilise such operators to actually perform error correction and decoding on such a code is not well understood. To what degree error correction procedures can be generalised to approximate error correcting codes is therefore an interesting direction for future research.

In Chapter 5 we provided an analysis of the finite-size effect of quantum communication in the so-called moderate deviation regime, in which the rate approaches the capacity *and* the error rate vanishes. At the time, this was the first work looking at applying moderate deviation techniques to the quantum regime. This analysis proceeded by first developing a moderate deviation analysis of

hypothesis testing, and then utilising a connection between this and c-q channel coding. There are however a multitude of other coding tasks which are known to have connections to hypothesis testing, such as quantum and entanglement-assisted classical coding, and this likely provides an avenue via which a moderate deviation analysis of these may also proceed.

In Chapters 6 and 7 we applied a finite-resource analysis to the problem of state interconversion in resource theories such as pure-state entanglement theory, and energy-incoherent quantum thermodynamics. In Chapter 6 this was considered in the regime of constant error, and refined in Chapter 7 to account for non-exponentially decreasing error.

The analysis of these resource theories relied on the fact that, mathematically, these resource theories are all based on majorisation conditions in the single-shot regime. A natural future direction is to perform similar finite-resource analyses to resource theories which are not known to be based on majorisation conditions. One of the most intriguing of these is the resource theory of magic, which is of critical importance in the area of quantum computing.

Finally in Chapter 8 we explored the phenomenon of *resource resonance*, which followed as a physical consequence of the analyses in Chapters 6 and 7. Here we found that, by careful engineering the states being used, detrimental finite-size effects in entanglement and thermodynamic transformation may be dramatically reduced. Whilst our theory is asymptotic in nature, we also showed numerically that this effect persists for small system sizes.

Resource resonance arose from a refinement of the asymptotic analysis. Even though this phenomenon is still relevant at small system sizes, an analysis how this phenomenon arises from the single-shot theory is still poorly understood. As above, whether this resonance is seen in other resource theories, most notably that of magic, would also be of significant importance. Such a resonance behaviour in the theory of magic would allow for significantly more efficient use of magic states in computational protocols.