

Initialization and Characterization of Open Quantum Systems

by

Christopher James Wood

A thesis
presented to the University of Waterloo
in fulfillment of the
thesis requirement for the degree of
Doctor of Philosophy
in
Physics

Waterloo, Ontario, Canada, 2015

© Christopher James Wood 2015

Author's Declaration

I hereby declare that I am the sole author of this thesis. This is a true copy of the thesis, including any required final revisions, as accepted by my examiners.

I understand that my thesis may be made electronically available to the public.

Abstract

In this thesis we address three topics of open quantum systems theory. The first concerns the operational characterization of an open quantum system. We develop a graphical calculus for open quantum systems and use it to review the operational representation and characterization of noisy evolution using completely-positive maps, or quantum channels. These graphical techniques facilitate an intuitive unification of the various representations of CP-maps, and the transformations between these representations. We generalize the channel formalism by introducing quantum superchannels — effective quantum channels which take another channel as input. Superchannels may be constructed using graphical techniques to rearrange the tensor network for a sequence of channels, and these constructions allow for the description of a strictly greater set of dynamics than is possible with the standard formalism. We demonstrate this by providing a method for the complete characterization of an open system initially correlated with its environment. In doing so we introduce and develop several novel quantitative measures for characterizing the strength and presence of initial correlations in a quantum system. These techniques may be implemented using measurement of the system alone, and do not require any access or prior knowledge about the environment.

The second topic we discuss is the parallel initialization of an ensemble quantum system into a high purity state. We describe a dissipative state engineering technique where an ensemble of identical spins may be prepared in its ground state by coupling the spins to the resolved sidebands of a high-Q resonator. This allows for parallel removal of entropy from a large number of quantum systems, enabling an ensemble to achieve a polarization greater than thermal equilibrium, and in principle on a time scale much shorter than thermal relaxation processes. This is achieved by the collective enhancement in coupling strength of the coupled angular momentum subspaces which behave as larger effective spins. Cavity cooling is shown to cool each of these subspaces to their respective ground state, however by including a local T_2 dephasing dissipator on each spin we show how one may use this to break the identical spin degeneracy and couple the subspaces to enable cooling to the full ground state of the joint system.

The third topic concerns quantum correlations between the momentum and spin subsystems of a neutron in a neutron interferometer. We explore the robustness of these correlations in the presence of dephasing noise due to randomized phase differences between interferometer paths. We demonstrate that even in the presence of strong noise quantum correlations persist and can be detected by introducing post-selected spin measurements at the output of the interferometer. We relate these post-selected experiments to which-way measurements and a quantum eraser.

Acknowledgements

I am extremely grateful to my supervisor David G. Cory for his seemingly endless supply of interesting research problems, and for the guidance, support, and insightful suggestions he has provided throughout my PhD. I'm also very glad to have known the many members who have passed through Corylab over my time here, with a special mention to Christopher Granade, Holger Haas, Daniel Puzzuoli, Ian Hincks, Troy Borneman, Pat Gumann and Dmitry Pushin. We've become a large family, and I thank you for the friendship, the beers, and for all the interesting and insightful discussions about physics and life. I also thank all my other colleagues and collaborators, especially Rob Spekkens, Daniel Terno and Alexei Gilchrist for everything they taught me at the start of my research career. And to Mark Butler, whose contagious enthusiasm for physics went above and beyond any requirements for a high school teacher, thank you for convincing me to study physics in the first place. To my family, though I may not always be in contact, I am always thinking of you and I'm extremely grateful for your love and support. A huge thank you to all the friends I've made during my time in Waterloo, there are too many of you to name, but you've made the last several years a blast. Jane Bond and staff, cheers for being Waterloo's (reluctant) physicist bar, and the best local one could hope for. Finally, thank you Explosions In The Sky for providing me with a soundtrack for my studies, travels, life, the universe, and everything.

Dedication

To Gail, Graeme and Michelle.

Table of Contents

List of Tables	xii
List of Figures	xiii
Abbreviations	xviii
1 Introduction	1
1.1 Noise in Quantum Systems	1
1.2 Thesis Outline and Main Results	2
1.3 Introduction to Open Quantum Systems	7
1.3.1 Mathematical Preliminaries	7
1.3.2 Closed Quantum Systems	10
1.3.3 Open Quantum Systems	11
1.3.4 Measures to Compare Quantum States	13
1.3.5 Measures of Correlations	14
2 Quantum Channels	16
2.1 Introduction	16
2.2 Tensor Networks and Graphical Calculus	19
2.2.1 Bipartite Matrix Operations	28
2.2.2 Vectorization of Matrices	29

2.3	Representations of Completely Positive Maps	33
2.3.1	Kraus Representation	33
2.3.2	System-Environment Representation	34
2.3.3	Superoperator Representation	37
2.3.4	Choi-Matrix Representation	38
2.3.5	Process Matrix Representation	41
2.4	Transforming Between Representations	43
2.4.1	Transformations Between the Choi-Matrix and Superoperator	44
2.4.2	Transformations to the Superoperator	45
2.4.3	Transformations to the Choi-Matrix Representation	47
2.4.4	Transformations to the Kraus Representation	49
2.4.5	Transformations to the System-Environment representation	54
2.5	Composite Channels	56
2.5.1	Vectorization of Composite Systems	57
2.5.2	Composing Superoperators and Choi-Matrices	61
2.5.3	Reduced Channels	63
2.5.4	Partial Trace of a Channel	67
2.6	Matrix Operations as Channels	68
2.7	Summary	71
3	Characterization of Open Quantum Systems	74
3.1	Introduction	74
3.2	Measures for Comparing Quantum Channels	77
3.2.1	Completely Bounded Trace Norm	77
3.2.2	Entanglement Fidelity	84
3.2.3	Average Gate Fidelity	87
3.3	Partial Characterization of Quantum Channels	91
3.3.1	Unitary Designs for Average Gate Fidelity Estimation	91

3.3.2	Randomized Benchmarking	93
3.4	Quantum State Tomography	95
3.4.1	Maximum Likelihood Estimation	98
3.4.2	Tomography with Two Outcome Measurements	102
3.5	Quantum Process Tomography	104
3.5.1	Ancilla Assisted Process Tomography	105
3.5.2	Robust Quantum Process Tomography	108
3.6	Summary	109
4	Quantum Superchannels	112
4.1	Introduction	112
4.2	Definition of Quantum Superchannels	114
4.3	Constructing Quantum Superchannels	116
4.3.1	Post-Selected Superchannels	121
4.4	Characterizing Superchannels	124
4.4.1	Quantum Superchannel Tomography	125
4.5	Twirling Superchannel	127
4.5.1	Subsystem Twirling Superchannel	131
4.6	Summary	134
5	Characterization of Initially Correlated Open Quantum Systems	136
5.1	Introduction	136
5.2	Initial Correlation Quantum Superchannel	139
5.2.1	Properties of the IC Superchannel	141
5.2.2	IC Quantum Superchannel Tomography	143
5.3	Initial Correlation Measures	143
5.3.1	Initial Correlation Norm	144
5.3.2	Preparation Fidelity	147

5.4	Example IC Superchannels	149
5.4.1	Classically Correlated Environment	150
5.4.2	Maximally Entangled Environment	154
5.4.3	Qubit System	157
5.5	Experimental Demonstration	160
5.5.1	Reconstructed IC Superchannels	162
5.5.2	Initial-Correlation Norm and Preparation Fidelity	164
5.6	Summary	165
6	Parallel State Initialization in an Ensemble Quantum System	168
6.1	Introduction	168
6.2	Collective Representations of Ensemble Quantum Systems	172
6.2.1	Dicke Model	175
6.2.2	Generalized Dicke Model	178
6.2.3	Symmetric Subspace	182
6.3	Spin-Cavity Master Equation	187
6.3.1	Adiabatic Elimination of a Dissipative Cavity	192
6.4	Cavity Cooling Master Equation	197
6.4.1	Subspace Equilibrium State	199
6.4.2	Simulation of Subspace Cooling Rate	201
6.4.3	Assumptions and Validity of the Markovian Approximation	203
6.4.4	Cavity Cooling in the Full State Space	207
6.5	Cavity Cooling with Local Dephasing	213
6.5.1	Magnus Expansion for Dissipative Interaction Frames	217
6.5.2	Average Cavity Cooling Dissipator in the T_2 Interaction Frame	219
6.5.3	Equilibrium Value and Cooling Rate for $\langle J_z(t) \rangle$	222
6.5.4	Simulation Results	225
6.6	Summary	227

7	Quantum Correlations in Neutron Interferometry	230
7.1	Introduction	230
7.2	Theoretical Model of a Neutron Interferometer	231
7.2.1	Output Intensities	234
7.2.2	Contrast	235
7.3	Quantum Correlations in a Neutron Interferometer	235
7.3.1	Output Intensity with Spin-filtering	237
7.3.2	Contrast with spin-filtering	238
7.3.3	Interpretation of Spin-Post Selection Experiments	240
7.3.4	Experimental Demonstration	243
7.4	Summary	244
8	Conclusion	247
	References	249

List of Tables

6.1	Commutation relations for $SU(4)$ algebra.	180
-----	--	-----

List of Figures

2.1	The main mathematical representations for completely positive maps and how one may transform between them. Solid arrows represent linear operations which we prove can be done by “wire bending” transformations in our graphical calculus. Dashed arrows represent non-linear transformations. Reshuffling and Stinespring dilation are bijective transformations, vectorization and the Jamiołkowski isomorphism are surjective transformations, and the spectral decomposition is an injective transformation.	17
2.2	Graphical depiction of elementary tensors. We represent vectors (states) and dual-vectors (effects) as triangles, linear operators as boxes, and scalars as diamonds, with each index of the tensor depicted as an open wire on the diagram. The orientation of the wires determines the type of tensor, in our convention the open end of the wires point to the left for vectors, right for dual-vectors, and both left and right for linear operators.	20
5.1	System dynamics in the presence of an environment. (a) Special case: with no initial SE correlations the reduced dynamics of the system, which interacts unitarily (U) with an environment, can be completely reconstructed from tomographically complete sets of input states $\{\rho_i\}$ (resulting from preparation procedures $\{\mathcal{P}_i\}$), and measurements $\{M_j\}$. (b) General case: the joint SE state may be initially correlated before the state preparation procedure. The superchannel approach encompasses this situation by treating the preparation procedure P_{ki} as the input state to a more general description of the reduced system dynamics.	137

5.2	Comparison of the IC-Value and rank-normalized IC-norm for an initially correlated qubit state interacting with the environment by a swap-gate. The $\mathcal{M}_{\text{sw}}^c, \mathcal{M}_{\text{sw}}^q$ correspond to the classically correlated and entangled initial SE states in Eqs. (5.113) and (5.114) respectively. θ parameters the degree of correlation with $\theta = \pi/4$ corresponding to the maximally correlated configuration for each state. We see that the measure of initial correlations via the IC-norm and IC-value is strictly greater for the entangled state than the classically correlated state.	159
5.3	Experimental setup. System and environment photons are created in the state ρ_{SE} with controllable degree of entanglement, using the source of Ref. [FTP ⁺ 07]. Arbitrary preparations \mathcal{P}_{ij} on the system and measurements $\{M_j\}$ are implemented by means of polarizers (PBS), quarter- and half-wave plates (QWP, HWP) and single-photon detectors (APD). The joint SE evolution U is implemented as a CZ gate between a set of HWPs and QWPs. In the case of no initial correlations this setup implements the target system evolution U_s . The CZ gate is based on non-classical interference at a partially polarizing beam splitter (PPBS) with reflectivities of $r_H = 0$ ($r_V = 2/3$) for horizontally (vertically) polarized light [LWP ⁺ 05].	160
5.4	Real parts of $\Lambda_{\mathcal{M}}$ for (a) $U_s = H$ and (b) $U_s = H$ in the ideal, uncorrelated case and experimental results for increasing strength of initial correlations. The matrices $\Lambda_{\mathcal{M}}$ are shown in a polarization-Pauli basis, with the elements from left to right corresponding to $\{ H\rangle, V\rangle\} \otimes \{\mathcal{I}, X, Y, Z\}$ and from front to back corresponding to $\{\langle H , \langle V \} \otimes \{\mathcal{I}, X, Y, Z\}$. The emergence of a peak corresponding to the identity operation (shown in yellow) is characteristic for the simulated increased tendency of the single-qubit operation U_s (shown in green) to fail in the presence of stronger initial correlations. The negligible imaginary parts are not shown.	163
5.5	(a) Real and (b) imaginary parts of $\Lambda_{\mathcal{M}}$ for an intended $U = R_Y Z$ operation on the system in the ideal, uncorrelated case and experimentally for increasing strength of initial correlations. The matrices $\Lambda_{\mathcal{M}}$ are shown in the same basis as in Fig. 5.4.	164

5.6	Real parts of the χ -matrices $\chi_{\mathcal{U}}$ (shown in the Pauli basis) for U_s obtained via QPT for different choices of preparation procedure in the case of low initial correlation $\tau = 0.136$. Cases (i) and (ii) correspond to a fixed ρ_k in Fig. 5.1(b), (iii) corresponds to $\rho_k = \rho_i$, and (iv) is the case where $1 \leq k \leq 4$. The information contained in the superchannel \mathcal{M} can be used to identify the optimal preparation procedure.	165
5.7	Initial correlation norm of the IC superchannel \mathcal{M}_{IC} vs correlation strength of the initial state ρ_{SE} . The correlation strength is given by the tangle τ from Eq. (5.131) and \mathcal{M}_{IC} is formed from a controlled unitary SE interaction with target unitary $U = \sigma_z$ (blue circles), $U = H$ (yellow squares) and $U = R_Y$ (green diamonds). The values of τ were obtained from state tomography of ρ_{SE} for each experiment. The measured real parts of the states with weakest and strongest initial correlations are shown in the respective insets. The solid line corresponds to the IC-norm in the ideal case. Error bars from Poissonian counting statistics are on the order of the symbol size.	166
5.8	Optimization of the preparation procedure. The average preparation fidelity $F_{\text{prep}}(\mathcal{M}, \rho_1, Z)$ for (a) $U_s = Z$ and (b) $U_s = R_Y Z$ is shown as a density plot on the surface of the Bloch sphere of the initial-projection state ρ_1 . In both cases, we chose the lowest strength of initial correlation realized in the experiment to visualize the effect even for very weak SE correlations.	167
6.1	Illustration of the driven Tavis-Cummings Hamiltonian for an ensemble of spins identically coupled to a single mode cavity represented as a quantum harmonic oscillator.	188
6.2	Illustration of the effective interaction of the Tavis-Cummings Hamiltonian after the state space of the spin ensemble has been decomposed into irreducible representations of $SU(2)$	189
6.3	Energy level diagram of the joint spin-cavity system with coherent transitions denoted by a solid line and cavity dissipation rates denoted by a curved line. States are labelled as $ n\rangle_c -J_x + m\rangle_s$, where m is the number of spin excitations and n is the number of cavity excitations.	191
6.4	Simulated evolution of the normalized expectation value of $-\langle J_x(t) \rangle / J$ for the Dicke subspace of a cavity-cooled spin ensemble containing N_s spins. The time axis is scaled by the effective dissipation rate, Γ_s , for the spin-ensemble given in Section 6.4.	201

6.5	Effective cooling time-constant, $T_{1,\text{eff}}$, of the spin ensemble as a function of the equilibrium excitation number of the cavity, \bar{n} , for $N_s = 10, 10^2, 10^3, 10^4$ spins in the ensemble.	203
6.6	Normalized spin expectation value $-\langle J_x \rangle / J$ of the spin ensemble as a function of time for various equilibrium temperatures of the cavity. We consider the case of $N_s = 10, 100, 1000$ and $10,000$ spins in the ensemble, and a cavity with resonant frequency $\omega_c = 10$ GHz, and a Rabi drive on resonance with the detuning $\Delta = \delta\omega - \Omega = 0$	204
6.7	Comparison of the cooling dynamics by the Markovian master equation (pink dotted curve) and a full spin-ensemble cavity simulation with RWA (blue dashed curve). The normalized expectation value of $-\langle J_x(t) \rangle / J$ is plotted for $N_s = 10$ spins with $\kappa = 0.5, 1, 5, 10g\sqrt{N_s}$, a cavity temperature $T = 0K$, and Rabi-drive matched to the detuning $\Delta = \delta\omega - \Omega = 0$. When $\kappa \geq 10g\sqrt{N_s}$ the Markovian master equation calculation agrees very well with the full simulation. Also, as predicted by eqn. (29), the cooling rate increases for larger κ , until the point where strong coupling effects take over.	206
6.8	Simulation of $\langle J_x(t) \rangle$ for cavity cooling of individual coupled angular momentum subspaces for $N = 10$ spins. The subspaces have J values ranging from 5 for the Dicke subspace, to 0 for the smallest singlet subspaces. We see that the cooling time decreases as the J value of the subspace increases. The black line shows the simulation of $\langle J_x \rangle_{eq}$ in the full Hilbert space by summing over all subspace values weighted by the multiplicity of each J value.	209
6.9	Comparison of the function $\Theta(\gamma)$ from Eq. (6.218) (solid blue line) with exponential approximations $e^{-\gamma\sqrt{\pi/2}}$ (dashed yellow line), and $e^{-\gamma}$ (dotted green line).	213
6.10	Simulation of cavity cooling under Eq. (6.224) for an ensemble of $N = 2$ spins, detuned to $\pm\Delta\omega$ from the cavity resonant frequency. We find that as the detuning increases from zero we couple the singlet and triplet subspaces allowing population transfer from the single to the triplet, and full cooling to the ground state of the triplet. However, as the detuning increases further the cooling rate is reduced, as a large detuning reduces the coupling strength between the spins and cavity thus preventing efficient transfer of energy.	215

6.11	Simulations of the expectation value $\langle J_z(t) \rangle$ for a maximally mixed initial state of an ensemble of $N = 10$ spins (left), and $N = 100$ spins (right) for master equation described by first order average dissipator in Eq. (6.267) (dotted black line), and of the cavity cooling with local dephasing master equation in Eq. (6.292) for $\gamma = \lambda N \Gamma$, and $\lambda = 0, 0.1, 1, 10$. In both cases we see good agreement between the first order approximation and the full master equation for dephasing rate $\gamma = 10 N \Gamma$	227
6.12	Simulations of the expectation value $\langle J_z(t) \rangle$ for the full spin-cavity master equation in Eq. (6.293) with the addition of a local dephasing dissipator for $N = 10$ spins in a maximally mixed initial state, and a cavity truncated to 4 levels initialized in the ground state. Evolution under the 1st order average dissipator in Eq. (6.267) is shown as the dotted black line in both figures. In the left figure we see good agreement with the cavity cooling master equation simulation in Fig. 6.11, however for stronger dephasing rates (right) we see that the cooling rate of the spin ensemble decreases.	228
7.1	Experimental setup for the three blade neutron interferometer (top) and the corresponding quantum circuit for the ideal model (bottom). The red (blue) paths in NI schematic are defined as the $ 0\rangle(1\rangle)$ path states, H is a Hadamard gate, $R_x(\alpha)$ is a rotation of the neutron spin in the $ 0\rangle$ path of α radians about x -axis, X is a bit-flip, $R_z(\phi)$ is a relative phase shift of ϕ radians between the beam paths, Π is a projective measurement performed on the spin-state (spin-analyser) in the basis $\cos(\theta) \uparrow\rangle \pm e^{i\phi} \sin(\theta) \downarrow\rangle$, and Z is a projective measurement of the path intensities in the $ 0\rangle, 1\rangle$ basis.	232
7.2	Entanglement of formation (left) and quantum discord $D(A B)$ (right) between the spin and path degrees of freedom of neutrons exiting a three-blade NI as a function of the spin rotation angle of neutrons in the $ 0\rangle$ interferometer path, and noise strength σ . The NI schematic is described in Fig. 7.1, and the noise model considered introduces a normally distributed random phase, with mean 0 and standard deviation σ , between the NI paths. The dashed line corresponds to the maximum noise case of a uniform distribution of angles. While the entanglement approaches zero for all spin rotation angles as the noise strength increases, the quantum discord remain non-zero.	236
7.3	Measured intensity curves at detector D_0 as a function of phase-flag rotation for three NIs. The corresponding path-contrast values are $C_{P1} = (82.5 \pm 1.3)\%$, $C_{P2} = (23 \pm 1.5)\%$, $C_{P3} = (2 \pm 1.7)\%$ for interferometers N_1, N_2, N_3 respectively.	244

7.4	Measured intensity curves at detector D_0 as a function of spin-rotation for three NIs where we have applied a spin-filter on the output beam to select spin-down neutrons with respect to a static magnetic field in the z -direction. The corresponding spin-filtered contrast values are $C_{S1(\downarrow z)} = (78.0 \pm 3)\%$, $C_{S2(\downarrow z)} = (74.2 \pm 2.2)\%$, $C_{S3(\downarrow z)} = (84 \pm 4)\%$ for interferometers N_1, N_2, N_3 respectively.	245
-----	---	-----

Abbreviations

AAPT ancilla assisted process tomography.

BCH Baker-Campbell-Hausdorf.

CB completely-bounded.

CP completely-positive.

CPTP completely-positive trace-preserving.

EOF entanglement of formation.

HP Hermitian-preserving.

IC initial correlation.

MLE maximum likelihood estimation.

NI neutron interferometer.

POVM positive-operator valued measure.

PVM projection-valued measure.

QD quantum discord.

QED quantum eletrodynamics.

QIP quantum information processing.

QPT quantum process tomography.

RB randomized benchmarking.

RWA rotating-wave approximation.

SDP semi-definite program.

SE system-environment.

SPAM state preparation and measurement.

TC Tavis-Cummings.

TCL time-convolutionless.

TP trace-preserving.

Chapter 1

Introduction

1.1 Noise in Quantum Systems

In the last several decades the field of quantum information processing has grown from a largely theoretical endeavour to see the application of quantum enhanced technologies for tasks such as quantum key distribution [SMWF⁺07] and high sensitivity measurements of quantities ranging from small magnetic fields using single spins [MSH⁺08] through to the proposed detection of gravitational waves using quantum squeezed states of light [AAA⁺13]. Future developments promise to be one of the most revolutionary technological advancements of the 21st century, with the eventual goal being a large scale quantum processor that can implement quantum algorithms that offer a superpolynomial speedup over the equivalent classical algorithms for certain tasks [Mos09, NIS]. There are several promising experimental implementations of small scale quantum processors consisting of multiple two-dimensional quantum systems, or qubits, that have the potential to scale to 10s to 100s of qubits in the near future. These include superconducting qubits [BHW⁺04], quantum dots [LD98], defects in silicon [SST⁺12], nitrogen vacancy centres in diamond [WJ06] and trapped ions [KMW02].

All of these real-world quantum devices are inevitably *open quantum systems* in the sense that they experience unwanted interactions with their environment [PB02]. The environment can abstractly be thought of as representing all the degrees of freedom that interact with the system, but are either inaccessible to the experimenter, or outside their ability to control. Typically these unwanted interactions manifest as noise processes on the quantum system, and one of the major challenges facing quantum information processing is limiting the influence of noise processes which inhibit or destroy quantum effects

such as coherence and entanglement, which are required for the successful operation of most quantum devices. Accurately characterize noise sources in a quantum system is thus a critical task for the design and operation of quantum devices. Once these noise sources are known work can then be done to suppress their influence, whether through new design and engineering advancements, or through application of quantum error correction methods [NC00]. In some situations noise can also be harnessed as a resource for example in initializing a quantum system in a state that may be otherwise unobtainable [PHBK99, BKPV99, CBHP13, LGR+13].

1.2 Thesis Outline and Main Results

The focus of this thesis is describing and developing the theory of open quantum systems for modelling, characterizing, and exploiting noise processes in these systems. We break this up into three main topics:

1. The first topic concerns the characterization of open quantum systems in the quantum channel formalism and is contained in [Chapters 2 to 5](#). These chapters are focused on the representation and characterization of the evolution of open quantum systems in the quantum channel formalism, and on generalizations of the quantum channel formalism to characterize open quantum system dynamics which lie outside the scope of the standard formalism.
2. The second topic concerns the initialization of an ensemble quantum system into a high purity state and is contained in [Chapter 6](#). This chapter considers applying cavity cooling techniques to an ensemble spin system by using the collective coupling of the ensemble to a cold high-Q resonator to enable dissipation to the ground state of the ensemble.
3. The third topic concerns quantum correlations in a neutron interferometer and is contained in [Chapter 7](#). This chapter considers the role of correlations between the momentum and spin degrees of freedom of a neutron in an interferometer, and the robustness of these correlations in the presence of noise arising from randomized phase differences between the interferometer beams.

In [Chapter 2](#) we describe a graphical calculus for completely positive maps and in doing so review the theory of open quantum systems and other fundamental primitives of quantum information theory using the language of tensor networks. In [Section 2.3](#) we

demonstrate the construction of tensor networks to pictographically represent the Liouville-superoperator, Choi-matrix, process-matrix, Kraus, and system-environment representations for completely positive (CP) maps, or quantum channels, which describe the evolution of quantum states. In [Section 2.4](#) we review how these representations interrelate, and illustrate how graphical manipulations of the tensor networks may be used to concisely transform between them. To further demonstrate the utility of the presented graphical calculus in [Section 2.5](#) we include several examples for how graphical techniques can simplify the composition and contraction of multipartite channels, and in [Section 2.6](#) we demonstrate how linear matrix operations may be represented as channels themselves.

In [Chapter 3](#) we review several common techniques for characterizing open quantum systems using the quantum channel formalism. In [Section 3.2](#) we review the completely-bounded (CB) trace norm, entanglement fidelity and average gate fidelity, which may be used to compare noisy quantum channels to an ideal target channel. Here we prove a new result for the upper bound of the CB trace norm in terms of the Choi-matrix of a Hermitian preserving channel, and further prove a sufficient condition for this bound to be tight. We also use the graphical techniques developed in [Chapter 2](#) to provide arguably simpler graphical proofs of entanglement fidelity and average gate fidelity in terms of each of the channel representations discussed in [Section 2.3](#). In [Section 3.3](#) we review the use of Clifford twirling and randomized benchmarking for computing the average gate fidelity of an unknown channel, and in particular we demonstrate how unitary 2-designs, used to represent these operations, may be constructed and applied to channels using the graphical techniques from [Chapter 2](#). In [Section 3.4](#) we present an overview of ideal quantum state tomography and show how standard linear inversion and maximum likelihood estimation (MLE) can be related to a least squares fitting problem, with typical linear inversion corresponding to the unconstrained least squares fit and MLE to a constrained fit that enforces the positivity of the reconstructed state. In [Section 3.5](#) we show how the presented formalism for quantum state tomography can be applied to quantum process tomography and several other generalizations by considering process tomography as a special case of quantum state tomography of the Choi-matrix representation of a quantum channel. Using graphical techniques we also derive a new condition for whether an arbitrary bipartite state may be used for ancilla assisted process tomography which is equivalent to known results, but which we believe is conceptually simpler and easier to check.

In [Chapter 4](#) we introduce a generalization of quantum channels we call *quantum superchannels*, and discuss the properties and some applications of these constructions. These are effective quantum channels which can be thought of as linear maps from quantum channels to quantum channels defined by their action on the Choi-matrix representation of the input channel. In [Section 4.2](#) we define quantum superchannels as a special type

of quantum channel with a bipartite input and output space and discuss their properties in relation to the properties of standard quantum channels introduced in [Chapter 2](#). In [Section 4.3](#) we describe how to construct quantum superchannels from an underlying composition sequence of channels acting on a Hilbert space, and discuss their conditions for these effective channels to be completely-positive, trace preserving, and Hermitian preserving. In particular we show how to construct superchannels for a single input, and how to compose superchannels to form channels for multiple inputs. We also describe how to post-select the output of the superchannel on a fixed input state or measurement outcome on the underlying physical system. These constructions are achieved by manipulating the tensor networks for the graphical representation of the channels. In [Section 4.4](#) we describe how to generalize quantum process tomography to quantum superchannel tomography and discuss the limitations of performing tomography when the input channels do not include projective measurements. In [Section 4.5](#) we give a simple example of a superchannel in terms of the Clifford twirling operation introduced in [Section 3.3](#). Following this we show how one may use the twirling superchannel to consider the action of twirling a number of subsystems of a quantum channel acting on a composite system. These techniques enable a description of a strictly greater range of dynamics than the quantum channel formalism and in principle could be used to model and characterize effects such as non-Markovian evolution of a quantum system.

In [Chapter 5](#) we introduce a quantum superchannel for fully characterizing the dynamics of a quantum system initially correlated with its environment. In [Section 5.2](#) we describe the construction the initial correlation (IC) superchannel which takes a Choi-matrix corresponding to a state preparation procedure as its input, rather than the resulting state. The resulting superchannels describes the subsequent evolution of the system including the interaction between the joint system-environment dynamics and the initial environment state that may be altered in the presence of initial correlations conditional on the choice of preparation procedure. We discuss properties of the IC superchannel, and in particular show that is a CP-map even in the presence of initial correlations, and also how one may perform a tomographic experiment to fully reconstruct the superchannel description of the dynamics, and thus experimentally characterize an initially correlated quantum system. Critically this reconstruction does not rely on access to the environment and is done entirely with measurement of the system alone. In [Section 5.3](#) we introduce quantitative measures for determining the strength of initial correlations in terms of an initial correlation (IC) norm. We also introduce a fidelity measure called *preparation fidelity* which quantifies how an experiment may be optimized while taking into consideration the effect of initial correlations with the environment. In [Section 5.4](#) we consider several theoretical examples of an initially correlated system. In particular we consider controlled-unitary

and swap system-environment interactions with both classically correlated and entangled initial system-environment states. We find that for the case of a controlled-unitary SE interaction there is no quantitative difference between a classically correlated and entangled unital state with respect to the IC-norm, however for the case of the swap interaction the entangled case displays strictly strong correlations under this measure. In [Section 5.5](#) we present an experimental demonstration of these techniques using a photonic qubit coupled to a simulated environment. We tomographically reconstruct the IC superchannel and compute the resulting IC-norm as a function of the initial correlations in the input state and find good agreement with our theoretical predictions. We also demonstrate how preparation fidelity may in principle be used to improve the implementation of a target operation in the presence of initial correlations by exploiting the action of these correlations.

In [Chapter 6](#) we describe how sideband cooling techniques may be applied to large spin ensembles in magnetic resonance to enable active cooling to a high purity ground state at a rate potentially faster than thermal relaxation processes. In [Section 6.2](#) we review the mathematical description of the state space of collective two-level quantum systems in terms of the irreducible representations of $SU(2)$ and $SU(4)$ for the collective state vectors and density matrices respectively. In [Section 6.3](#) we introduce the Tavis-Cummings model for describe the dynamics of a spin-ensemble interacting with a single mode cavity in the presence of a Rabi drive and describe how to adiabatically eliminate the cavity in the Markovian regime to obtain a reduced Markovian master equation for the dynamics of the spin-ensemble. In [Section 6.4](#) we solve the spin-ensemble Markovian master equation to derive cooling rates for the coupled angular momentum subspace of the ensemble as a function of ensemble size. Our calculations indicate that the Dicke subspace of a spin ensemble containing roughly 10^{11} electron spins may be polarized in a time many orders of magnitude shorter than the typical thermal relaxation time. We also discuss cavity cooling dynamics in the full Hilbert space and derive approximate expressions for the final magnetization and cooling rate of the full Hilbert space which scale with the square root of the ensemble size. In [Section 6.5](#) we include the effect of local dephasing on the spin ensemble and show that this couples the collective subspaces to enable cooling to the ground state of the ensemble. To solve the dynamics of the ensemble system we introduce a method for dissipative perturbation theory that applies average Hamiltonian theory in an imaginary-time dissipative interaction frame to find an average effective dissipator for the system dynamics. We use $SU(4)$ algebra generators to analytically solve the first order perturbation for an arbitrary number of systems in the ensemble. We find that to first order the effective dissipator describes local thermal relaxation to the ground state of each qubit in the ensemble at a rate equal to the collective cavity cooling dissipation rate. The described techniques should permit the parallel initialization of high purity states in

large ensemble quantum systems based on solid state spins. The proposed application of a standard technique in quantum optics to magnetic resonance also serves to reinforce the connection between the two fields, which has recently begun to be explored in further detail due to the development of hybrid designs for manufacturing noise-resilient quantum devices.

In [Chapter 7](#) we investigate quantum coherences in the presence of noise by entangling the spin and path degrees of freedom of the output neutron beam from a noisy three-blade perfect crystal neutron interferometer. In [Section 7.2](#) we introduce a quantum information description of a three blade neutron interferometer treating it as a bipartite system where one subsystem is represented by the path, or momentum, degree of freedom of neutrons travelling through the interferometer crystal, and the other subsystem is the spin of the neutron. In [Section 7.3](#) we show that in the presence of dephasing noise on the path degree of freedom the entanglement of the output state reduces to zero, however the quantum discord remains non-zero for all noise values. Hence even in the presence of strong phase noise non-classical correlations persist between the spin and path of the neutron beam. This indicates that measurements performed on the spin of the neutron beam will induce a disturbance on the path state. We calculate the effect of the spin measurement by observing the changes in the observed contrast of the interferometer for an output beam post-selected on a given spin state. In doing so we demonstrate that these measurements allow us to implement a quantum eraser, and a which-way measurement of the path taken by the neutron through the interferometer. While strong phase noise removes the quantum eraser, the spin-filtered which-way measurement is robust to phase noise. We experimentally demonstrate this disturbance by comparing the contrasts of the output beam with and without spin measurements of three neutron interferometers with varying noise strengths. This demonstrates that even in the presence of noise that suppresses path coherence and spin-path entanglement, a neutron interferometer still exhibits uniquely quantum behaviour.

Acknowledgement of Contributions

- [Chapter 2](#) and [Chapter 3](#) contains results published in [\[WBC15\]](#) with additional unpublished results of the author. The work in [\[WBC15\]](#) was done in collaboration with Jacob D. Biamonte and David G. Cory.
- [Chapter 4](#) contains unpublished results of the author.
- [Chapter 5](#) contains results published in [\[RWM⁺15\]](#) with additional unpublished results of the author. The work in [\[RWM⁺15\]](#) was done in collaboration with Martin

Ringbauer, Kavan Modi, Alexei Gilchrist, Andrew G. White, Alessandro Fedrizzi. The theoretical work was done by Modi, Gilchrist and the author, and the experimental work was done by Ringbauer, Fedrizzi, and White, with the tomographic reconstruction of experimental data done by Ringbauer and the author.

- [Chapter 6](#) contains results published in [\[WBC14\]](#) and [\[WC15\]](#) with additional unpublished results of the author. The work in [\[WBC14\]](#) was done in collaboration with Troy W. Borneman and David G. Cory.
- [Chapter 7](#) contains results published in [\[WAH⁺14\]](#) which was done in collaboration with Mohamed O. Abutaleb, Michael G. Huber, Arif, David G. Cory Dmitry A. Pushin. In particular the experimental work was implemented by Abutaleb, Huber and Pushin.

1.3 Introduction to Open Quantum Systems

1.3.1 Mathematical Preliminaries

We now introduce the basic mathematical notations and conventions that will be used throughout this thesis. We use the notation that script letters $\mathcal{X}, \mathcal{Y}, \mathcal{Z}$ represent finite dimensional complex Euclidean vector spaces or Hilbert spaces \mathbb{C}^d .

We use Dirac notation for vectors $|v\rangle \in \mathcal{X} \cong \mathbb{C}^d$

$$|x\rangle \in \mathcal{X} \cong \mathbb{C}^d \rightarrow |x\rangle = \sum_{j=0}^{d-1} x_j |j\rangle = \begin{pmatrix} x_0 \\ x_1 \\ \vdots \\ x_{d-1} \end{pmatrix}, \quad (1.1)$$

where $x_j \in \mathbb{C}$ and $\{|j\rangle\}$ is the standard or computation basis. Each element of the computation basis $|j\rangle$ can be represented as a unit column vectors with a 1 in the $j^{\text{th}} + 1$ row, and 0's elsewhere.

Euclidean vector spaces come equipped with a Euclidean *norm* and *inner product*. Euclidean inner product (or dot product) is given by

$$\langle x|y\rangle = \sum_{i,j=0}^{d-1} \bar{x}_i y_j \langle i|j\rangle = (\bar{x}_0 \quad \bar{x}_1 \quad \dots \quad \bar{x}_{d-1}) \begin{pmatrix} y_0 \\ y_1 \\ \vdots \\ y_{d-1} \end{pmatrix} = \sum_{j=0}^{d-1} \bar{x}_j y_j \quad (1.2)$$

where \bar{x}_j denotes complex conjugation of $x_j \in \mathbb{C}$.

The Euclidean norm $\|\cdot\|$ is given by

$$\| |x\rangle \| = \sqrt{\langle x|x \rangle} = \sqrt{\sum_{j=0}^{d-1} x_j^2} \quad \text{for } |x\rangle \in \mathcal{X} \quad (1.3)$$

In terms of the Euclidean inner product, we may think of the object $\langle x|\cdot\rangle$ as a row vector $(\bar{x}_0 \ \bar{x}_1 \ \dots \ \bar{x}_{d-1})$ in the dual vector space $\mathcal{X}^\dagger \cong \mathbb{C}^d$, hence they are sometimes called *dual vectors* or *effects*.

The set of linear operators $A : \mathcal{X} \rightarrow \mathcal{Y}$ mapping between two vector spaces form their own vector space $L(\mathcal{X}, \mathcal{Y})$. In the case where the operator maps to the same vector space we use the notation $L(\mathcal{X}) \equiv L(\mathcal{X}, \mathcal{X})$. We can define several useful norms on operators. The *Schatten p -norm* of an operator A is defined as

$$\|A\|_p \equiv \left(\text{Tr} \left[(A^\dagger A)^{p/2} \right] \right)^{1/p} \quad (1.4)$$

where $1 \leq p$. The most common Schatten norms are $p = 1, 2$ and the limiting case of $p = \infty$. In these cases Eq. (1.4) reduce to the following expressions:

1. The *trace norm* or 1-norm

$$\|A\|_1 \equiv \text{Tr} |A| \quad \text{where } |A| \equiv \sqrt{A^\dagger A} \quad (1.5)$$

2. The *Frobenius norm* or 2-norm

$$\|A\|_2 \equiv \sqrt{\text{Tr}[A^\dagger A]} \quad (1.6)$$

3. The *Spectral norm, operator norm* or ∞ -norm

$$\|A\|_\infty \equiv \max \{ \|A|v\rangle\| : \| |v\rangle \| = 1 \} \quad (1.7)$$

An alternate expression for the Schatten norm is the result of the maximization

$$\|A\|_p = \max \left\{ |\text{Tr}[B^\dagger A]| : B \in L(\mathcal{X}, \mathcal{Y}), \|B\|_{p^*} \leq 1 \right\}, \quad p^* = \frac{p}{1-p} \quad (1.8)$$

The Schatten p -norms satisfy

$$\|A\|_p = \|\bar{A}\|_p = \|A^T\|_p = \|A^\dagger\|_p \quad (1.9)$$

where \bar{A} denotes element wise complex conjugation with respect to the standard basis, A^T denotes transposition, and A^\dagger denotes Hermitian-conjugation $A^\dagger \equiv \bar{A}^T$. For tensor product operators these norms reduce to norms on the corresponding subsystem operators:

$$\|A \otimes B\|_p = \|A\|_p \|B\|_p. \quad (1.10)$$

The operator norms also bound each other in the following manner:

$$\|A\|_\infty \leq \|A\|_2 \leq \|A\|_1. \quad (1.11)$$

For a square operator $A \in L(\mathcal{X})$ the *trace* of A is given by

$$\text{Tr}[A] = \sum_{j=0}^{d-1} \langle j|A|j\rangle. \quad (1.12)$$

For operators on a bipartite vector space $A \otimes B \in L(\mathcal{X} \otimes \mathcal{Y})$, the *partial trace* over \mathcal{Y} of $A \otimes B$ is given by

$$\text{Tr}_{\mathcal{Y}}[A \otimes B] = A \text{Tr}[B]. \quad (1.13)$$

Similarly $\text{Tr}_{\mathcal{X}}[A \otimes B] = \text{Tr}[A] B$. The 1-norm has the useful property that it is decreasing with respect to the partial trace. For any operator $A \in L(\mathcal{X} \otimes \mathcal{Y})$ we have

$$\|A\|_1 \geq \|\text{Tr}_{\mathcal{Y}}[A]\|_1. \quad (1.14)$$

An important class of operators are *positive semidefinite* operators. An operator $A \in L(\mathcal{X})$ is positive-semidefinite, denoted by $A \geq 0$, if and only if all the eigenvalues of A are non-negative. By the spectral decomposition such an operator can be written as

$$A = \sum_{i=1}^r \lambda_i |v_i\rangle\langle v_i| \quad (1.15)$$

where $\lambda_i \geq 0$, $\{|v_i\rangle : i = 1, \dots, \dim(\mathcal{X})\}$ is an orthonormal basis, and $1 \leq r \leq \dim(\mathcal{X})$ is called the *rank* of A . For such operators the previously mentioned operator norms are given by

$$\|A\|_1 = \sum_{i=1}^r |\lambda_i|, \quad \|A\|_2 = \sqrt{\sum_{i=1}^r |\lambda_i|^2}, \quad \|A\|_\infty = \max\{|\lambda_i| : i = 1, \dots, r\} \quad (1.16)$$

Two other common used classes of operators are *Hermitian* operators, and *unitary* operators. A linear operator $A \in L(\mathcal{X})$ is

- *Hermitian* if and only if $A^\dagger = A$.
- *Unitary* if and only if $A^\dagger A = AA^\dagger = \mathbb{1}$.

1.3.2 Closed Quantum Systems

Closed quantum systems are ideal quantum systems which are perfectly isolated from the environment. In quantum information theory we are often interested in systems where the base subsystem is a 2-dimensional system called a *qubit*. Larger systems are then formed as composite systems of multiple qubits. The behaviour of closed quantum systems are governed by the postulates of quantum mechanics [NC00]:

1. **States** are represented as unit vectors $|\psi\rangle \in \mathcal{X}$, $\| |\psi\rangle \| = 1$, where $\mathcal{X} \cong \mathbb{C}^d$ is called the *state space* of the system.
2. **Transformations** are represented by unitary matrices $U \in L(\mathcal{X})$ that act on states: $|\psi'\rangle = U|\psi\rangle$.
3. **Measurements** are represented by a *projective valued measures* (PVM). A PVM is an orthonormal set of projectors $\{|a_j\rangle\langle a_j| := 1, \dots, d\}$, $\sum_j |a_j\rangle\langle a_j| = \mathbb{1}$. The probability of a measurement outcome j is given by $p_j = |\langle a_j|\psi\rangle|^2$, and the state of the system after observation of outcome j is $|\psi\rangle \mapsto |a_j\rangle$.
4. **Composite** quantum systems may be formed by taking the tensor product of the state spaces of multiple quantum systems. A composite state is *separable* if it may be written in the form $|\Psi\rangle = |\psi_1\rangle \otimes |\psi_2\rangle$. If a state is not separable it is called *entangled*.

These postulates describe the discrete time, or circuit description, of closed quantum systems. They allow us to break any experiment involving quantum systems into three components: *State preparation*, where we initialize a quantum system in a known state $|\psi\rangle$; *Transformations*, where we apply a unitary to the state $|\psi\rangle$, and finally *measurement* of the final state. Measurements are how we extract classical information from a quantum system, for example to determine the output of an algorithm, or to determine properties of a quantum state.

For considering the specific dynamics of transformations of a quantum system, the time evolution of a closed quantum systems is governed by *Schrödinger's equation*:

$$i\hbar \frac{d}{dt} |\psi(t)\rangle = H(t) |\psi(t)\rangle. \quad (1.17)$$

The operator $H(t)$ is a Hermitian operator called the *Hamiltonian* of the system. For the remainder of this thesis we will work in energy units where $\hbar = 1$. The solution of Schrödinger's equation is unitary evolution

$$|\psi(t)\rangle = U(t) |\psi(0)\rangle \quad \text{where } U(t) = \mathcal{T} \exp \left(-i \int_0^t ds H(s) \right) \quad (1.18)$$

and \mathcal{T} is the *time-ordering* operator.

1.3.3 Open Quantum Systems

Open quantum systems describe real world quantum systems which interact with an environment. In such systems the postulates of quantum mechanics must be modified as the observable dynamics on the system may no longer appear unitary. One may always think of an open quantum system as a subsystem of a larger closed quantum system consisting of a *system* which represents all the degrees of freedom accessible to an experimenter, and an *environment* which represents all the degrees of freedom that are inaccessible. The postulates for quantum mechanics that describe closed quantum systems can be modified to describe open quantum systems as follows [NC00]:

1. **States** are represented by density matrices $\rho \in D(\mathcal{X})$, where $D(\mathcal{X}) = \{\rho, \rho \in L(\mathcal{X}), \text{Tr}[\rho] = 1\}$. We refer to density matrices of the form $\rho = |\psi\rangle\langle\psi|$ as *pure states*. While density matrices of the form

$$\rho = \sum_i p_i |\psi_i\rangle\langle\psi_i|, \quad p_i > 0, \quad \sum_i p_i = 1 \quad (1.19)$$

are called *mixed states*. Operationally mixed states represent our uncertainty about a given quantum state. We may interpret a mixed state of the form in Eq. (1.19) as describing a system that we believe may be in a state $|\psi_j\rangle$ with probability p_i . The *purity* of a quantum state is defined as $\text{Tr}[\rho^2]$ and is a measure of how close a density matrix is to a pure state. Pure states satisfy $\text{Tr}[\rho^2] = 1$, while the lower bound is given by the most mixed state, the *maximally mixed state* $\rho = \mathbb{1}/d$, which has purity $1/d$.

2. **Transformations** are represented by *completely positive trace preserving (CPTP) maps* $\mathcal{E} : D(\mathcal{X}) \mapsto D(\mathcal{X})$. Informally CPTP maps can be thought of as the most general operation which will always map a valid density matrix to a valid density matrix. These operations are also called *quantum operations*, *superoperators* or *quantum channels* and we will discuss these in detail in Chapter 2.
3. **Measurements** are represented by a *positive operator valued measure (POVM)*. A POVM is a set of positive operators $\{E_j : E_j > 0, \sum_i E_j^\dagger E_j = \mathbb{1}\}$. The probability of a measurement outcome j is given by $p_j = \text{Tr}[E_j \rho]$. To consider the state of a system after a measurement we must decompose each POVM element into *effects*

$E_j = \sum_k M_{j,k}^\dagger M_{j,k}$. In this case the state of the system after observation of outcome j is given by

$$\rho \mapsto \rho_j = \sum_k \frac{M_{j,k} \rho M_{j,k}^\dagger}{\text{Tr}[E_j \rho]}. \quad (1.20)$$

Note that the decomposition of a POVM into effects is not unique, and depends on the description of the measurement being performed.

4. **Composite** quantum systems may be formed by taking the tensor product of the state spaces of multiple quantum systems. A composite state is called *simply separable* or a *product state* if it may be written in the form $\rho = \rho_1 \otimes \rho_2$. It is called *separable* if it may be written as a convex sum of simply separable states $\rho = \sum_i p_i \rho_{1,i} \otimes \rho_{2,i}$. If a state is not separable it is called *entangled*.

The time evolution of a closed quantum system in the density matrix picture is governed by the von-Neuman equation

$$\frac{d}{dt} \rho(t) = -i[H(t), \rho(t)] \quad (1.21)$$

where $[A, B] \equiv AB - BA$ is the *commutator* of A and B . The solution of this equation is the usual unitary evolution:

$$\rho(t) = U(t) \rho(t) U(t)^\dagger, \quad U(t) = \mathcal{T} \exp \left(-i \int_t^t ds H(s) \right). \quad (1.22)$$

To include the effects of noise acting on the system one can modify the von-Neuman equation into a *quantum master equation* which governs the dynamics of open system evolution [PB02]. The simplest quantum master equation is a *Markovian* master equation describing interaction with a memoryless environment. Informally we can think of this as stating that any information transferred from the system to the environment is lost and can have no influence on the dynamics of the system at later times. The simplest markovian master equation is called the *Lindblad equation* which is a generalization of the von-Neuman equation by including an additional term called a *dissipator*.

$$\frac{d}{dt} \rho(t) = -i[H(t), \rho(t)] + \sum_j \gamma_j D[A_j] \rho(t) \quad (1.23)$$

where $D[A_j]$ is called the *Lindblad dissipator* and is given by

$$D[A] \rho(t) = A \rho(t) A^\dagger - \frac{1}{2} \{A^\dagger A, \rho(t)\}, \quad (1.24)$$

where $\{A, B\} \equiv AB + BA$ is the *anti-commutator* of A and B . The Lindblad dissipator generates dissipative evolution described by a *Lindblad operator* A_j at rate γ_j . Unlike unitary dynamics, the dynamics generated by a dissipator are not in general reversible, and hence are used to describe decoherence mechanisms in a system. Just as the solution to the von-Neuman equation is unitary evolution, the solution to a Lindblad master equation is a CP-map which describes the discrete time evolution of an open quantum system. These CP-maps that govern the evolution of an open quantum system are also referred to as quantum channels.

1.3.4 Measures to Compare Quantum States

There are two widely used operational measures for comparing two quantum states. These are the *trace distance* and the *fidelity* [GLN05].

The trace distance between two states $\rho_1, \rho_2 \in D(\mathcal{X})$ is defined as

$$D_{\text{tr}}(\rho_1, \rho_2) = \frac{1}{2} \|\rho_1 - \rho_2\|_1. \quad (1.25)$$

The operational interpretation of this quantity is related to the probability of successful distinguishing state ρ_1 from ρ_2 in a single shot measurement. If an experimenter is given a state that they know is either ρ_1 with probability λ , or ρ_2 with probability $1 - \lambda$ their optimal strategy for distinguishing the two states with a single measurement will succeed with probability

$$p = \frac{1}{2} + D_{\text{tr}}(\lambda\rho_1, (1 - \lambda)\rho_2) = \frac{1}{2} + \frac{1}{2} \|\lambda\rho_1 + (1 - \lambda)\rho_2\|_1. \quad (1.26)$$

The fidelity between two quantum states $\rho_1, \rho_2 \in D(\mathcal{X})$ is defined as

$$F(\rho_1, \rho_2) = \|\sqrt{\rho_1}\sqrt{\rho_2}\|_1^2 = \left(\text{Tr} \sqrt{\sqrt{\rho_1}\rho_2\sqrt{\rho_1}} \right)^2 \quad (1.27)$$

If one of the states is a pure state the fidelity reduces to

$$F(|\psi\rangle\langle\psi|, \rho) = \langle\psi|\rho|\psi\rangle \quad (1.28)$$

Hence for pure states fidelity is given by $F(|\psi\rangle, |\phi\rangle) = |\langle\psi|\phi\rangle|^2$. This has the operational interpretation of the probability of distinguishing the two states, where the probability of successfully distinguishing $|\psi\rangle$ for $|\phi\rangle$ succeeds with probability $1 - F(|\psi\rangle, |\phi\rangle)$. For mixed

states an operational interpretation is that the fidelity is as an upper bound of the overlap of two pure input states that were mapped to ρ_1 and ρ_2 by some quantum channel [DN02]. Note that in many places in the literature fidelity is defined as the square root of the definition in Eq. (1.27). We adopt the squared convention as advocated for in [GLN05] due to its generalization to fidelity measures on quantum channels which we discuss in Chapter 3.

The fidelity and trace distance are related via the Fuchs-van de Graff inequality [FvdG99]:

$$1 - \sqrt{F(\rho_1, \rho_2)} \leq D_{\text{tr}}(\rho_1, \rho_2) \leq \sqrt{1 - F(\rho_1, \rho_2)}. \quad (1.29)$$

1.3.5 Measures of Correlations

A unique property of quantum theory is that when two or more quantum systems are allowed to interact they may exhibit correlations that cannot be explained classically. In the field of quantum information science protocols harnessing these correlations can exceed classical efficiencies for certain metrology applications and information processing tasks [NC00]. One of the most studied classes of correlated quantum states are entangled states. There are numerous measures for quantifying entanglement in a quantum state (for a review of entanglement see [HHHH09]). A convenient measure for a two-qubit mixed-state is the *entanglement of formation* (EOF) [Woo98] which is given by

$$EOF(\rho_{AB}) = h\left(\frac{1 + \sqrt{1 - \mathcal{C}(\rho_{AB})^2}}{2}\right) \quad (1.30)$$

where $h(x) = -x \log x - (1 - x) \log(1 - x)$, and $\mathcal{C}(\rho_{AB})$ is the *concurrence* of a bipartite state ρ_{AB} :

$$\mathcal{C}(\rho_{AB}) = \max\{0, \lambda_1 - \lambda_2 - \lambda_3 - \lambda_4\} \quad (1.31)$$

where λ_j are the eigenvalues of the Hermitian matrix

$$\sqrt{\sqrt{\rho_{AB}}(Y \otimes Y)\rho_{AB}^\dagger(Y \otimes Y)\sqrt{\rho_{AB}}} \quad (1.32)$$

sorted such that $\lambda_1 \geq \lambda_2 \geq \lambda_3 \geq \lambda_4$, where Y is the Pauli- Y matrix.

The class of states of interest to quantum computation however is broader than purely entangled quantum states, as certain non-entangled quantum states may still possess correlations that cannot be accounted for classically. There are many proposed measures for quantifying the distinction between classical and quantum states which are based on

entropic quantities, and one of the most widely used measures is *quantum discord* (for a review of quantum discord and related measures see [MBC⁺12]). Quantum discord is a non-symmetrical quantity defined by the difference between quantum generalizations of two classically equivalent expressions for *mutual information*. Let ρ_{AB} be a bipartite density matrix over two quantum systems A and B . One expression for the mutual information of ρ_{AB} is given by

$$I(A : B) = S(\rho_A) + S(\rho_B) - S(\rho_{AB}) \quad (1.33)$$

where $S(\rho) = -\text{Tr}(\rho \log \rho)$ is the von-Neumann entropy of the density matrix ρ , $\rho_A = \text{tr}_B(\rho_{AB})$ is the reduced density matrix on subsystem A taken by performing the partial trace over system B , and similarly $\rho_B = \text{tr}_A(\rho_{AB})$. An alternative expression for mutual information is formed by considering a quantum generalization of conditional entropy which accounts for possible measurement induced disturbances. Consider performing a measurement on subsystem B , this is most generally described by POVM E consisting of a set of measurement operators $\{E_b\}$ satisfying $E_b \geq 0$, $\sum_b E_b = \mathbb{1}$. Measurement outcome b will occur with probability $p_b = \text{Tr}(E_b \rho_{AB})$, and the post-measurement state of subsystem A , conditioned on outcome b , is given by

$$\rho_{A|b} = \frac{1}{p_b} \text{Tr}_B(E_b \rho_{AB}). \quad (1.34)$$

We may define a generalization of conditional entropy for a given POVM E as

$$S(\rho_{A|E}) = \sum_b p_b S(\rho_{A|b}). \quad (1.35)$$

This gives us an alternative expression for mutual information by maximizing over all possible POVMs:

$$J(A|B) = \max_E [S(\rho_A) - S(\rho_{A|E})] \quad (1.36)$$

Quantum discord is defined to be the difference between expressions Eq. (1.33) and Eq. (1.36):

$$D(A|B) = I(A : B) - J(A|B) = \min_E [S(\rho_{A|E}) + S(\rho_B) - S(\rho_{AB})] \quad (1.37)$$

Similarly one may define the quantum discord $D(B|A)$ where one optimizes over POVMs on subsystem A . In general to compute the quantum discord of a state one must minimize Eq. (1.37) over all extremal rank-one POVMs, however it has been shown that for rank-two states orthogonal projective valued measurements (PVMs) are optimal [GGZ11].

Chapter 2

Quantum Channels

2.1 Introduction

A complete description of the evolution of quantum systems is an important tool in quantum information processing. In contrast to closed quantum systems, for open quantum systems the evolution need no longer be unitary. The linear operator describing the evolution over a fixed time period of the density operator of an open quantum system called a *quantum operation* or *quantum channel*. This is defined mathematically by a *completely positive map* (CP-map) [NC00]. Let \mathcal{X} , \mathcal{Y} be complex Euclidean spaces and let $T(\mathcal{X}, \mathcal{Y})$ be the Hilbert space of operator maps $\mathcal{E} : L(\mathcal{X}) \rightarrow L(\mathcal{Y})$. A map $\mathcal{E} \in T(\mathcal{X}, \mathcal{Y})$ is

1. *Positive* if and only if it preserves the positivity of an operators spectrum, that is for all $\rho \in L(\mathcal{X})$ with $\rho \geq 0$ we have $\mathcal{E}(\rho) \geq 0$.
2. *Completely positive* if and only if it further satisfies the condition that the composite map $\mathcal{I} \otimes \mathcal{E}$ is positive, where \mathcal{I} is the identity map on a space of density operators with dimension greater than or equal to the dimension of the space on which \mathcal{E} acts.
3. *Trace preserving* if and only if $\text{Tr}[\mathcal{E}(\rho)] = \text{Tr}[\rho]$ for all operators ρ .

Analogously to how unitary operators are the most general operation which will map a quantum state vector to another quantum state vector, it can be shown that a completely positive map is the most general mathematical object that will map a density operator to another density operator. Since density operators are positive operators ρ with unit trace, the requirement that \mathcal{E} be complete-positive and trace-preserving ensures that the output

state of the map will always be a valid density operator. Such maps are called completely positive trace preserving maps (CPTP-maps).

There are numerous representations for completely-positive maps [JS61, Cho75, Kra83], and in Section 2.3 we review and describe the properties of the main representations used in quantum information theory. These are the system-environment, and closely related Stinespring, representations, the Kraus or operator-sum representation, the Choi-matrix representation and its change of basis to the proves-matrix of χ -matrix, and the super-operator. In Section 2.4 we describe how to transform between these representations, in doing so we will use the graphical calculus methods which we present in Section 2.2. This graphical approach facilitates an intuitive unification and interoperability between these representations, the outcome of which is depicted in Fig. 2.1. In doing so we provide a compact review of the properties and transformations of CP-maps.

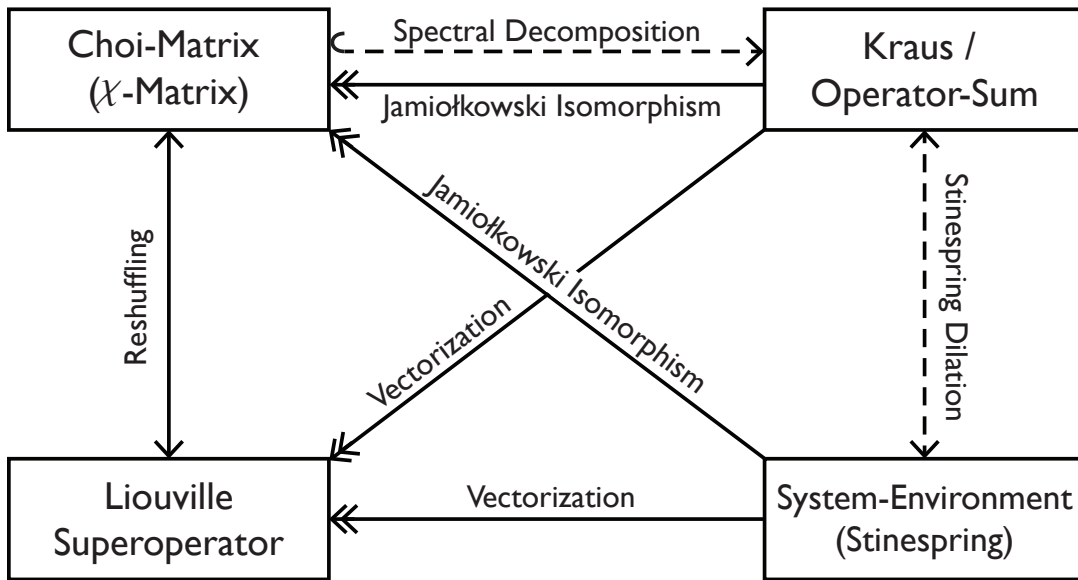


Figure 2.1: The main mathematical representations for completely positive maps and how one may transform between them. Solid arrows represent linear operations which we prove can be done by “wire bending” transformations in our graphical calculus. Dashed arrows represent non-linear transformations. Reshuffling and Stinespring dilation are bijective transformations, vectorization and the Jamiolkowski isomorphism are surjective transformations, and the spectral decomposition is an injective transformation.

Graphical calculi have been used to great benefit in several areas of modern physics with the most prolific example being the use of Feynman diagrams to calculate scattering amplitudes in quantum field theories [BL11]. In the context of quantum information theory there has been recent interest in employing graphical techniques more general than standard quantum circuits, with two popular approaches being based on *tensor networks* and *category theory*. The approach presented here casts the theory of open quantum systems into the framework of tensor networks, which comes equipped with a graphical means to represent and reason about the contraction of sequences of tensors [Pen71]. The use of tensor networks dates back to earlier work by Penrose, who’s graphical notation is a useful starting point [Pen71]. They have been used as computational tools for simulating certain many-body quantum systems efficiently [Vid08, EV09, VMC08], as a tool for manipulating tensor networks [BCJ11] and to generalize quantum circuits [BB11]. Although it is straightforward to translate equations into so-called tensor string diagrams, a missing piece has been a graphical calculus for open systems theory which provided new results, and hence enhanced the potential for diagrammatic reasoning.

The category theoretic approaches for quantum information theory are based on so-called *dagger-compact monoidal categories* which were used by Abramsky and Coecke to abstractly describe quantum teleportation [AC04]. This approach was then extended to include CP-maps by Selinger in the *CPM-construction* [Sel07]. The graphical language of these approaches are built upon well established graphical calculi for *compact closed categories* [KL80] and *symmetric monoidal categories* [JS91]. A key result is that Selinger’s calculus for CP-maps is *complete* for finite-dimensional Hilbert spaces [Sel11a]. This means that any identity which can be represented graphically is true if and only if it is algebraically true, which is important for diagrammatic proofs. Subsequent work based on these constructions has been used to graphically depict quantum protocols [CPP09, BH12] and Bayesian inference [CS12]; and for the axiomatic formulation of quantum theory [CH11]. For a review of graphical calculi for monoidal categories see [Sel11b]. Other alternative graphical approaches have also been used in the axiomatic formulation of quantum theory [CDP11, Har11].

There have been at least two graphical calculi previously presented for CP-maps: Selinger’s aforementioned category theoretic approach, and a graph-theory approach by Collins and Nechita which was used to compute ensemble averages of random quantum states [CN11, CN10]. Selinger’s CPM-construction bears some similarities to our approach, however there are important and practical differences between the two. The CPM-construction is most closely related to our superoperator representation in the row-vectorization convention which we present in Section 2.3.3. In the presented graphical calculus we tailor the tensor string diagrams of Penrose to unify several mathematical

representations used in open quantum systems and to transform freely between them. In accomplishing this, we express our graphical tensor calculus in the Dirac notation familiar in quantum information theory instead of the abstract index notation used by Penrose, or the category theoretic notation used by others. This provides a toolset which can be used for the manipulation, visual representation, and contraction of quantum circuits and general open quantum system equations.

To demonstrate the utility of the presented graphical calculus we provide examples where we use these tools to derive several common quantities used in quantum information theory. We emphasize that these are not new results, but rather it is the application of the graphical methods that we are aiming to highlight. In particular, in [Section 2.5](#) we demonstrate the construction of composite channels involving the composition of several subsystem channels, and also the contraction of multipartite channels to an effective subsystem channel. In [Section 2.6](#) we show how matrix operations may be represented graphically as superoperators, these expressions will be used in later chapters for several graphical proofs. In [Chapter 3](#) we also use graphical techniques to derive several quantities used for the channel characterization.

2.2 Tensor Networks and Graphical Calculus

Tensors can be thought of as indexed multi-dimensional arrays of complex numbers with respect to a fixed standard basis. The number of indices is called the *order* of a tensor, and the concurrent evaluation of all indices returns a complex number. For example, consider the Hilbert space $\mathcal{X} \cong \mathbb{C}^d$, where as is typical in QIP, we choose our standard basis to be the *computational basis* $\{|i\rangle : i = 0, \dots, d-1\}$. Then in Dirac notation a vector $|v\rangle \in \mathcal{X}$ is a 1st-order tensor which can be expressed in terms of its tensor components $v_i := \langle i|v\rangle$ with respect to the standard basis as $|v\rangle = \sum_{i=0}^{d-1} v_i |i\rangle$. Similarly one can represent linear operators on this Hilbert space, $A \in L(\mathcal{X})$, as 2nd-order tensors with components $A_{ij} := \langle i|A|j\rangle$ as $A = \sum_{i,j=0}^{d-1} A_{ij} |i\rangle\langle j|$.

Hence, in Dirac notation the number of indices of a tensor's components are what we refer to as the order of the tensor. Vectors $|v\rangle \in \mathcal{X}$ refer to tensors which only have *ket* “ $|i\rangle$ ” basis elements, vectors in the dual vector space $\langle u| \in \mathcal{X}^\dagger$ refer to those with only *bras* “ $\langle i|$ ”, and linear operators on $A \in \mathcal{L}(\mathcal{X})$ refer to tensors with a mixture of kets and bras in their component decomposition.

The idea of representing states, operators and maps (etc.) diagrammatically dates back to several works by Penrose and is often referred to as *Penrose graphical notation* or *string*

diagrams. We adopt Penrose's notation of representing states (vectors) and effects (dual-vectors) as triangles, linear operators as boxes, and scalars as diamonds, as illustrated in Fig. 2.2. Here each index corresponds to an open wire on the diagram and so we may define higher order tensors with increasingly more wires. The number of wires is then the order of the tensor, with each wire acting on a separate vector space \mathcal{X}_j .

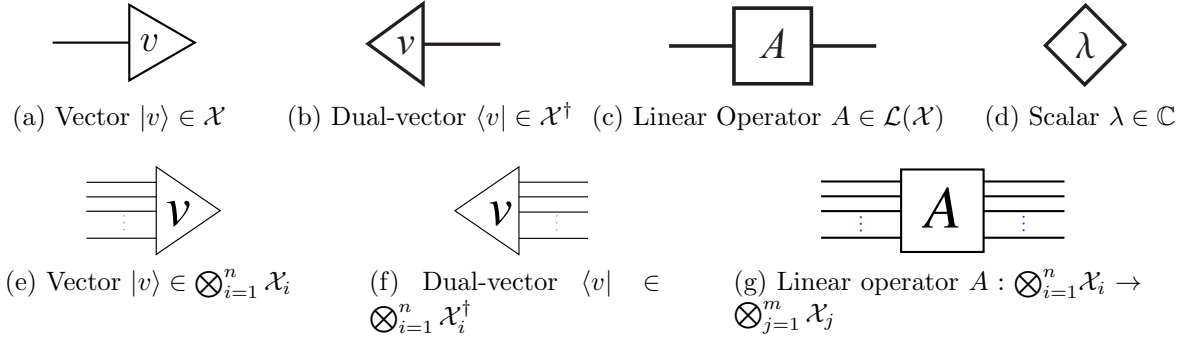


Figure 2.2: Graphical depiction of elementary tensors. We represent vectors (states) and dual-vectors (effects) as triangles, linear operators as boxes, and scalars as diamonds, with each index of the tensor depicted as an open wire on the diagram. The orientation of the wires determines the type of tensor, in our convention the open end of the wires point to the left for vectors, right for dual-vectors, and both left and right for linear operators.

We also insist that the orientation of these wires, rather than the number of wires, specifies whether they represent multi-partite vectors, dual-vectors, or linear operators. We have a freedom in choosing our orientation for the tensors, top-to-bottom, bottom-to-top, left-to-right or right-to-left. In this paper we will choose the *right-to-left* convention (the opposite of most orthodox quantum circuits) so that the graphical representation will most closely match the underlying equations. Thus we use the terms *vector*, *dual-vector* and *linear operator* to refer to tensors of any order, not just 1st-order and 2nd-order, based on the orientation of their wires as follows:

1. *Vectors* $|v\rangle \in \bigotimes_{j=1}^n \mathcal{X}_j$ are tensors with $n \geq 1$ wires oriented to the left.
2. *Vectors in the dual space* $\langle v| \in \bigotimes_{j=1}^n \mathcal{X}_j^\dagger$ are tensors with $n \geq 1$ wires are oriented to the right.
3. *Linear operators* $A : \bigotimes_{i=1}^n \mathcal{X}_i \rightarrow \bigotimes_{j=1}^m \mathcal{X}_j$ are tensors which have $n \geq 1$ wires going to the right and $m \geq 1$ wires to the left.

4. Tensors with no open wires are *scalars* $\lambda \in \mathbb{C}$.

The graphical depictions of these tensors are also illustrated in Fig. 2.2. In the present paper, we will generally be interested in the case where each wire indexes $\mathcal{X}_j \cong \mathbb{C}^d$ for fixed dimension d , though one may generalize most of what follows to situations where the dimensions of each wire are not equal. Note that we represent scalars as tensors with no open wires, this could either be a contracted tensor $\lambda = \langle v|u \rangle$, or a multiplicative factor λ acting on the tensor A as λA .

The mathematical rules of tensor network theory assert that the wires of tensors may be manipulated, with each manipulation corresponding to a specific contraction or transformation. We now introduce some tools which we have tailored for manipulations common in open quantum systems. Transposition of 1st-order vectors and dual-vectors, and 2nd-order linear operators is represented by a *bending* of a tensors wires as follows:

(a) Vector transposition:
 $|v\rangle^T = \langle \bar{v}|$

(b) Dual-vector transposition:
 $\langle v|^T = |\bar{v}\rangle$

(2.1)

(c) Linear operator transposition

(2.2)

Complex conjugation of a tensor's coefficients however is depicted by a bar over the tensor label in the diagram:

(a) Complex conjugation of $|v\rangle$

(b) Complex conjugation of $\langle v|$

(2.3)

(c) Complex conjugation of A

(2.4)

Hence we may represent the transformation of a vector to its dual vector, or the hermitian conjugation of a linear operator as the combination of these two operations:

(a) Hermitian conjugation of a vector : $|v\rangle^\dagger = \langle v|$ (b) Hermitian conjugation of a dual-vector : $\langle v|^\dagger = |v\rangle$

(2.5)

(c) Hermitian Conjugation of an operator A

(2.6)

We stress that under this convention a vector $|v\rangle = \sum_i v_i |i\rangle$ and its hermitian conjugate dual-vector $\langle v| = \sum_i \bar{v}_i \langle i|$ are represented as shown in [Figs. 2.2a](#) and [2.2b](#) respectively.

Tensor contraction is represented by joining the wires corresponding to the indices to be contracted. In the case of matrix multiplication $A \cdot B$ is represented by connecting the corresponding wires of the tensors representing the matrices:

(2.7)

To form multi-partite tensors we denote the tensor product of two tensors $A \otimes B$ by the vertical juxtaposition of their tensor networks:

(2.8)

The trace, $\text{Tr}[A]$, of an operator A is depicted by connecting the corresponding left and right wires of a linear operator:

(2.9)

We represent summation in our variant of graphical calculus adapted to open systems by introducing shading or coloring of the tensors being summed over. We call this the color summation convention. Tensors corresponding to the same summation index will be shaded the same color, and we use different colors for different summation indexes. For example, consider the spectral decomposition of a normal operator A with eigenvalues λ_i and eigenvectors $|a_i\rangle$. The graphical depiction of the spectral-decomposition $A = \sum_i \lambda_i |a_i\rangle$ using the color summation convention is given by:

$$\text{---} \boxed{A} \text{---} = \text{---} \begin{array}{c} \triangleleft a \\ \lambda \\ \triangleright a \end{array} \text{---} := \sum_i \text{---} \begin{array}{c} \triangleleft a_i \\ \lambda_i \\ \triangleright a_i \end{array} \text{---} \quad (2.10)$$

In our color summation convention, we will represent the sum over the standard basis as a shaded vector (or dual-vector) tensor with an empty label. This is demonstrated for the graphical resolution of the identity $\mathbb{1} = \sum_{i=0}^{d-1} |i\rangle\langle i|$ as follows:

$$\text{---} \begin{array}{c} \triangleleft \\ \triangleright \end{array} \text{---} := \sum_i \text{---} \begin{array}{c} \triangleleft i \\ \triangleright i \end{array} \text{---} = \text{---} \quad (2.11)$$

Using this convention we represent summation over a Kronecker delta, $\sum_{i,j} \delta_{ij} = \sum_{i,j} \langle i|j\rangle$ as shown:

$$\begin{array}{c} \triangleleft \\ \triangleright \end{array} := \begin{array}{c} \triangleleft \\ \triangleright \end{array} \text{---} \begin{array}{c} \triangleright \\ \triangleleft \end{array} = \sum_{ij} \delta_{ij} \quad (2.12)$$

This expression will be useful in graphical proofs.

The unnormalized maximally entangled Bell-state $|\Phi^+\rangle = \sum_{i=0}^{d-1} |i\rangle \otimes |i\rangle \in \mathcal{X} \otimes \mathcal{X}$ is represented graphically as the curve:

$$\text{)} := \text{---} \begin{array}{c} \triangleright \\ \Phi^+ \\ \triangleright \end{array} \text{---} = \text{---} \begin{array}{c} \triangleleft \\ \triangleleft \end{array} \quad (2.13)$$

Similarly the unnormalized Bell-effect $\langle \Phi^+|$ is represented as:

$$\text{(} := \text{---} \begin{array}{c} \triangleleft \\ \Phi^+ \\ \triangleleft \end{array} \text{---} = \text{---} \begin{array}{c} \triangleright \\ \triangleright \end{array} \quad (2.14)$$

As will be shown in [Section 2.2.2](#), our choice of graphical notation for $|\Phi^+\rangle$ is due to its equivalence to the vectorization of the identity operator.

As a simple example of graphical proof techniques we may use this to prove the tensor network in [Eq. \(2.9\)](#) for the trace of an operator A :

$$\begin{aligned}
 \text{Diagram of } A \text{ with a loop} &= \text{Diagram of } A \text{ with blue and green triangles} \\
 &= \text{Diagram of } A \text{ with blue triangles} \\
 &= \sum_i A_{ii}
 \end{aligned}
 \tag{2.15}$$

For illustrative purposes, to prove this algebraically we note that the tensor networks for trace correspond to the algebraic expressions $\langle \Phi^+ | A \otimes \mathbb{1} | \Phi^+ \rangle$ and $\langle \Phi^+ | \mathbb{1} \otimes A | \Phi^+ \rangle$, and that

$$\begin{aligned}
 \langle \Phi^+ | \mathbb{1} \otimes A | \Phi^+ \rangle &= \sum_{i,j} \langle i|j \rangle \langle i|A|j \rangle = \sum_{i,j} \delta_{ij} A_{ij} \\
 &= \sum_i A_{ii} \\
 &= \text{Tr}[A].
 \end{aligned}
 \tag{2.16}$$

Similarly we get $\langle \Phi^+ | A \otimes \mathbb{1} | \Phi^+ \rangle = \text{Tr}[A]$.

Using the graphical definition for $|\Phi^+\rangle$ we can compose the unnormalized Bell-state and its dual to form an identity element [[Pen71](#)]. This is known as the *snake equation* or *zig-zag equation* and is given by:

$$\text{Line} = \text{Snake} = \text{Zig-zag}
 \tag{2.17}$$

To prove the snake equation we must first make the following equivalence for tensor products of the elements $|i\rangle$ and $\langle j|$:

$$\langle j| \otimes |i\rangle \equiv |i\rangle \otimes \langle j| \equiv |i\rangle \langle j|
 \tag{2.18}$$

This is illustrated diagrammatically as

$$\text{Diagrammatic equivalence of } |i\rangle \otimes \langle j| \text{ and } |i\rangle \langle j|
 \tag{2.19}$$

With this equivalence made, the proof of the snake-equation for the “S” bend is given by

$$\begin{aligned}
 & \text{S-bend} = \text{Two blue triangles (left) and two green triangles (right)} \\
 & = \text{Blue triangle (left) and green triangle (right)} \\
 & = \text{Single straight wire}
 \end{aligned}
 \tag{2.20}$$

The proof for the reflected “S” snake-equation follows naturally from the equivalence defined in Eq. (2.19). The snake equations have several uses and provide an equivalence class of diagrams. Anytime we have a curved wire with two bends we can “pull the wire” to straighten it out into an identity. Anytime we bend a wire, transforming between say a bra and a ket, we can bend the wire to transform back again.

By combining the snake-equation with the wire-bending operation for transposition, we find that “sliding” a linear operator around an unnormalized Bell-state is also equivalent to transposition of the operator:

$$\text{Box } A^T \text{ with wire loop} = \text{Box } A \text{ with wire loop}
 \tag{2.21}$$

Note that due to the orientation of the wires this graphical representation of the operator A is actually a vector. This is called the vectorization of a matrix and we discuss this in more detail in Section 2.2.2.

The graphical proof of Eq. (2.21) follows from the tensor network for transposition in

Eq. (2.1):

$$\begin{aligned}
 \text{Diagram of } A^T \text{ contraction} &= \text{Diagram of } A \text{ contraction} \\
 &= \text{Diagram of } A \text{ contraction}
 \end{aligned}
 \tag{2.22}$$

In order to prove Eq. (2.1) itself we have

$$\begin{aligned}
 \text{Diagram of } A \text{ contraction} &= \text{Diagram of } A \text{ contraction with triangles} \\
 &= \text{Diagram of } A \text{ contraction with triangles} \\
 &= \text{Diagram of } A^T \text{ contraction}
 \end{aligned}
 \tag{2.23}$$

We can also prove this algebraically to demonstrate the correspondence between a tensor network and the underlying algebraic equation. This is given by

$$\text{Diagram of } A \text{ contraction} = \mathbb{1} \otimes \langle \Phi^+ | (\mathbb{1} \otimes A \otimes \mathbb{1}) | \Phi^+ \rangle \otimes \mathbb{1}
 \tag{2.24}$$

$$= \sum_{i,j} \langle j | A | i \rangle | i \rangle \otimes \langle j |
 \tag{2.25}$$

$$= \sum_{i,j} \langle j | A | i \rangle | i \rangle \langle j |
 \tag{2.26}$$

$$= \sum_{i,j} \langle i | A^T | j \rangle | i \rangle \langle j |
 \tag{2.27}$$

$$= \sum_{i,j} | i \rangle \langle i | A^T | j \rangle \langle j |
 \tag{2.28}$$

$$= A^T.
 \tag{2.29}$$

The proof for transposition by counter-clockwise wire bending follows from the equivalence relation in Eq. (2.18) and Eq. (2.19).

Another important operation is the graphical SWAP which exchanges the position of two Hilbert spaces in a composite system. Let \mathcal{X} and \mathcal{Y} be complex Hilbert spaces of dimensions d_1 and d_2 respectively, then the SWAP operation is the map

$$\text{SWAP} : \mathcal{X} \otimes \mathcal{Y} \rightarrow \mathcal{Y} \otimes \mathcal{X} \quad (2.30)$$

$$\text{SWAP} : |x\rangle \otimes |y\rangle \mapsto |y\rangle \otimes |x\rangle, \quad (2.31)$$

for all $|x\rangle \in \mathcal{X}, |y\rangle \in \mathcal{Y}$.

Given any two orthonormal basis $\{|x_i\rangle : i = 0, \dots, d_1 - 1\}$ and $\{|y_j\rangle : j = 0, \dots, d_2 - 1\}$ for \mathcal{X} and \mathcal{Y} respectively, we can give an explicit construction for the SWAP operation as

$$\text{SWAP} = \sum_{i_1=0}^{d_1-1} \sum_{j_2=0}^{d_2-1} |y_j\rangle\langle x_i| \otimes |x_i\rangle\langle y_j|. \quad (2.32)$$

The SWAP operation is represented graphically by two crossing wires as shown:

$$\begin{array}{c} \diagup \quad \diagdown \\ \diagdown \quad \diagup \end{array} := \begin{array}{c} \text{---} \triangleleft \triangleleft \text{---} \\ \text{---} \triangleright \triangleright \text{---} \end{array} \quad (2.33)$$

The basis decomposition in Eq. (2.32) is then an application of the resolution of the identity to each wire. In Section 2.2.2 we will see that the SWAP operation is the natural transformation between the row-stacking and column-stacking vectorization conventions.

Using the above primitives one may represent any linear equation involving the composition and contraction of tensors by a tensor network diagram. In the reverse case, given a tensor network diagram one may always write down an equivalent equation by labelling each wire by an index, and then writing down the corresponding tensor components and summing over contracted indices. By manipulating the tensor diagrams and making use of the primitives introduced one may obtain equivalent expressions, however these manipulated forms may have a more convenient equational form that is not inherently obvious from simply looking at the original equation. We gave several simple examples of this in our explicit proofs of the tensor networks for trace, transposition and the snake equation.

2.2.1 Bipartite Matrix Operations

Bipartite matrices are used in several representations of CP-maps, and manipulations of these matrices will be important in the following discussion. Consider two complex Hilbert spaces \mathcal{X} , and \mathcal{Y} with dimensions d_x and d_y respectively. The bipartite matrices we are interested in are then $d_x^2 \times d_y^2$ matrices $M \in L(\mathcal{X} \otimes \mathcal{Y})$ which we can represent as 4th-order tensors with tensor components

$$M_{m\mu, n\nu} := \langle m, \mu | M | n, \nu \rangle \quad (2.34)$$

where $m, n \in \{0, \dots, d_x - 1\}$, $\mu, \nu \in \{0, \dots, d_y - 1\}$ and $|n, \nu\rangle := |n\rangle \otimes |\nu\rangle \in \mathcal{X} \otimes \mathcal{Y}$ is the tensor product of the standard bases for \mathcal{X} and \mathcal{Y} . Graphically this is given by

$$M_{m\mu, n\nu} = \begin{array}{c} \triangleleft m \\ \mu \triangleleft \\ \square M \\ \triangleright n \\ \nu \triangleright \end{array} \quad (2.35)$$

We can also express the matrix M as a 2nd-order tensor in terms of the standard basis $\{|\alpha\rangle : \alpha = 0, \dots, D - 1\}$ for $\mathcal{X} \otimes \mathcal{Y}$ where $D = d_x d_y$. In this case M has tensor components

$$M_{\alpha\beta} = \langle \alpha | M | \beta \rangle \quad (2.36)$$

This is represented graphically as

$$M_{\alpha\beta} = \begin{array}{c} \triangleleft \alpha \\ \square M \\ \triangleright \beta \end{array} = \begin{array}{c} \triangleleft \alpha \\ \square M \\ \triangleright \beta \end{array} \quad (2.37)$$

We can specify the equivalence between the tensor components $M_{\alpha\beta}$ and $M_{m\mu, n\nu}$ by making the assignment

$$\alpha = d_y m + \mu \quad (2.38)$$

$$\beta = d_y n + \nu, \quad (2.39)$$

where d_y is the dimension of the Hilbert space \mathcal{Y} .

The bipartite matrix operations which are the most relevant for open quantum systems (see Fig. 2.1) are the *partial trace over \mathcal{X}* ($\text{Tr}_{\mathcal{X}}$) (and $\text{Tr}_{\mathcal{Y}}$ over \mathcal{Y}), *transposition* (T),

bipartite-SWAP (S), *col-reshuffling* (R_c), and *row-reshuffling* (R_r). The corresponding graphical manipulations are:

(a) Partial Trace $\text{Tr}_{\mathcal{X}}[M]$ (b) Partial Trace $\text{Tr}_{\mathcal{Y}}[M]$ (c) Transpose M^T (2.40)

(d) Bipartite-Swap M^S (e) Row-Reshuffle M^{R_r} (f) Col-Reshuffle M^{R_c}

In terms of the tensor components of M these operations are respectively given by:

Partial trace over \mathcal{X}	$\text{Tr}_{\mathcal{X}} : L(\mathcal{X} \otimes \mathcal{Y}) \rightarrow L(\mathcal{Y}),$	$M_{m\mu, n\nu} \mapsto \sum_m M_{m\mu, m\nu}$
Partial trace over \mathcal{Y}	$\text{Tr}_{\mathcal{Y}} : L(\mathcal{X} \otimes \mathcal{Y}) \rightarrow L(\mathcal{X})$	$M_{m\mu, n\nu} \mapsto \sum_{\mu} M_{m\mu, n\mu}$
Transpose	$T : L(\mathcal{X} \otimes \mathcal{Y}) \rightarrow L(\mathcal{X} \otimes \mathcal{Y}),$	$M_{m\mu, n\nu} \mapsto M_{n\nu, m\mu}$
Bipartite-SWAP	$S : L(\mathcal{X} \otimes \mathcal{Y}) \rightarrow L(\mathcal{Y} \otimes \mathcal{X}),$	$M_{m\mu, n\nu} \mapsto M_{\mu m, \nu n}$
Row-reshuffling	$R_r : L(\mathcal{X} \otimes \mathcal{Y}) \rightarrow L(\mathcal{Y} \otimes \mathcal{Y}, \mathcal{X} \otimes \mathcal{X}),$	$M_{m\mu, n\nu} \mapsto M_{mn, \mu\nu}$
Col-reshuffling	$R_c : L(\mathcal{X} \otimes \mathcal{Y}) \rightarrow L(\mathcal{X} \otimes \mathcal{X}, \mathcal{Y} \otimes \mathcal{Y}),$	$M_{m\mu, n\nu} \mapsto M_{\nu\mu, nm}$

The reshuffling operations are used for transforming between the Choi-matrix and superoperator representations of CP-maps which is described in [Section 2.4.1](#). Note that we will generally use reshuffling R to refer to col-reshuffling R_c . Similarly we can represent the partial transpose operation by only transposing the wires for \mathcal{X} (or \mathcal{Y}), and the partial-SWAP operations by only swapping the left (or right) wires of M .

2.2.2 Vectorization of Matrices

We now recall the concept of *vectorization* which is a reshaping operation, transforming a $(m \times n)$ -matrix into a $(1 \times mn)$ -vector [[HJ85](#)]. This is necessary for the description of open quantum systems in the superoperator formalism, which we will consider in [Section 2.3.3](#). Vectorization can be done with using one of two standard conventions: *column-stacking*

(col-vec) or *row-stacking* (row-vec). Consider two complex Hilbert spaces $\mathcal{X} \cong \mathbb{C}^m$, $\mathcal{Y} \cong \mathbb{C}^n$, and linear operators $A \in L(\mathcal{X}, \mathcal{Y})$ from \mathcal{X} to \mathcal{Y} . Column and row vectorization are the mappings

$$\text{col-vec: } L(\mathcal{X}, \mathcal{Y}) \rightarrow \mathcal{X} \otimes \mathcal{Y} : A \mapsto |A\rangle\rangle_c \quad (2.41)$$

$$\text{row-vec: } L(\mathcal{X}, \mathcal{Y}) \rightarrow \mathcal{Y} \otimes \mathcal{X} : A \mapsto |A\rangle\rangle_r \quad (2.42)$$

respectively, where the operation col(row)-vec when applied to a matrix, outputs a vector with the columns (rows) of the matrix stacked on top of each other. Graphical representations for the row-vec and col-vec operations are found from bending a wire to the left either clockwise or counterclockwise respectively:

(a) Row-vec

(b) Col-vec

$$(2.43)$$

Vectorized matrices in the col-vec and row-vec conventions are naturally equivalent under wire exchange (the SWAP operation)

$$(2.44)$$

In particular we can see that the unnormalized Bell-state $|\Phi^+\rangle \in \mathcal{X} \otimes \mathcal{X}$ is in fact the vectorized identity operator $\mathbb{1} \in L(\mathcal{X})$

$$|\Phi^+\rangle = |\mathbb{1}\rangle\rangle_r = |\mathbb{1}\rangle\rangle_c. \quad (2.45)$$

We may also define a vectorization operation with respect to an arbitrary operator basis for $L(\mathcal{X}, \mathcal{Y})$. Let $\mathcal{X} \cong \mathbb{C}^{d_x}$, $\mathcal{Y} \cong \mathbb{C}^{d_y}$, and $\mathcal{Z} \cong \mathbb{C}^D$ where $D = d_x d_y$. Vectorization with respect to an orthonormal operator basis $\{\sigma_\alpha : \alpha = 0, \dots, D-1\}$ for $L(\mathcal{X}, \mathcal{Y})$ is given by

$$\sigma\text{-vec: } L(\mathcal{X}, \mathcal{Y}) \rightarrow \mathcal{Z} : A \mapsto |A\rangle\rangle_\sigma. \quad (2.46)$$

This operation extracts the coefficients of the basis elements returning the vector

$$|A\rangle\rangle_\sigma := \sum_{\alpha=0}^{D-1} \text{Tr}[\sigma_\alpha^\dagger A] |\alpha\rangle \quad (2.47)$$

where $\{|\alpha\rangle : \alpha = 0, \dots, D - 1\}$ is the standard basis for $\mathcal{Z} \cong \mathbb{C}^D$. This is depicted in our graphical calculus as

$$\text{Diagram of } A_\sigma = \text{Diagram of } \sigma^\dagger \circ A \quad (2.48)$$

To distinguish between these different conventions we use the notation $|A\rangle\rangle_x$ to denote the vectorization of a matrix A , where the subscript $x = c, r, \sigma$ labels which convention we use; either c for col-vec, r for row-vec, or σ for an arbitrary operator basis.

For the case $\mathcal{X} \cong \mathcal{Y} \cong \mathbb{C}^d$, we can define row-vec and col-vec in terms of Eq. (2.47) by taking our basis to be the elementary matrix basis $\{E_{i,j} = |i\rangle\langle j| : i, j = 0, \dots, d^2 - 1\}$, and making the assignment $\alpha = di + j$ and $\alpha = i + dj$ respectively. Hence we have

$$|A\rangle\rangle_r = \sum_{i,j=0}^{d-1} A_{ij} |i\rangle \otimes |j\rangle \quad (2.49)$$

$$|A\rangle\rangle_c = \sum_{i,j=0}^{d-1} A_{ij} |j\rangle \otimes |i\rangle. \quad (2.50)$$

Using the definition of the unnormalized Bell-state $|\Phi^+\rangle$ and summing over i and j one can rewrite Eqs. (2.49) and (2.50) as

$$|A\rangle\rangle_r = (A \otimes \mathbb{1})|\Phi^+\rangle \quad (2.51)$$

$$|A\rangle\rangle_c = (\mathbb{1} \otimes A)|\Phi^+\rangle \quad (2.52)$$

which are the equational versions of our graphical definition of row and col vectorization shown in Eq. (2.43).

When working in the superoperator formalism for open quantum systems, it is sometimes convenient to transform between vectorization conventions in different bases. Given two orthonormal operator bases $\{\sigma_\alpha\}$ and $\{\omega_\alpha\}$ for $L(\mathcal{X}, \mathcal{Y})$, the basis transformation operator

$$T_{\sigma \rightarrow \omega} : \mathcal{Z} \rightarrow \mathcal{Z} : |A\rangle\rangle_\sigma \mapsto |A\rangle\rangle_\omega \quad (2.53)$$

transforms vectorized operators in the σ -vec convention to the ω -vec convention. Graphically this is given by

$$\text{Diagram of } A_\omega = \text{Diagram of } T_{\sigma \rightarrow \omega} \circ A_\sigma \quad (2.54)$$

The basis transformation operator $T_{\sigma \rightarrow \omega}$ is given by the equivalent expressions

$$T_{\sigma \rightarrow \omega} = \sum_{\alpha} |\alpha\rangle \langle\langle \omega_{\alpha} |_{\sigma} = \sum_{\alpha} |\sigma_{\alpha}\rangle_{\omega} \langle\alpha|, \quad (2.55)$$

and the corresponding graphical representations are:

$$\boxed{T_{\sigma \rightarrow \omega}} = \begin{array}{c} \text{---} \\ \text{---} \end{array} \begin{array}{c} \text{---} \\ \text{---} \end{array} = \begin{array}{c} \text{---} \\ \text{---} \end{array} \begin{array}{c} \text{---} \\ \text{---} \end{array} \begin{array}{c} \text{---} \\ \text{---} \end{array} = \begin{array}{c} \text{---} \\ \text{---} \end{array} \begin{array}{c} \text{---} \\ \text{---} \end{array} \begin{array}{c} \text{---} \\ \text{---} \end{array} \quad (2.56)$$

The proof of Eq. (2.55) is as follows:

$$T_{\sigma \rightarrow \omega} |A\rangle_{\sigma} = \left(\sum_{\alpha} |\alpha\rangle \langle\langle \omega_{\alpha} |_{\sigma} \right) |A\rangle_{\sigma} \quad (2.57)$$

$$= \sum_{\alpha} |\alpha\rangle \langle\langle \omega_{\alpha} |A\rangle_{\sigma} \quad (2.58)$$

$$= \sum_{\alpha} |\alpha\rangle \text{Tr}[\omega_{\alpha}^{\dagger} A] \quad (2.59)$$

$$= |A\rangle_{\omega}. \quad (2.60)$$

The inverse of $T_{\sigma \rightarrow \omega}$ is given by

$$T_{\sigma \rightarrow \omega}^{-1} = T_{\sigma \rightarrow \omega}^{\dagger} = T_{\omega \rightarrow \sigma} \quad (2.61)$$

and hence $T_{\sigma \rightarrow \omega}$ is unitary.

For the remainder of this paper we will use the col-vec convention by default, and drop the vectorization label subscripts unless referring to a general σ -basis. The main transformation we will be interested in is then from col-vec to another arbitrary orthonormal operator basis $\{\sigma_{\alpha}\}$. Tensor networks for the change of basis $T_{c \rightarrow \sigma}$ and its inverse $T_{\sigma \rightarrow c}$ are

$$\begin{array}{c} \text{---} \\ \text{---} \end{array} \boxed{T_{c \rightarrow \sigma}} = \begin{array}{c} \text{---} \\ \text{---} \end{array} \begin{array}{c} \text{---} \\ \text{---} \end{array} \begin{array}{c} \text{---} \\ \text{---} \end{array} \begin{array}{c} \text{---} \\ \text{---} \end{array} \begin{array}{c} \text{---} \\ \text{---} \end{array} \begin{array}{c} \text{---} \\ \text{---} \end{array} \begin{array}{c} \text{---} \\ \text{---} \end{array} \quad (a) \text{ Col-vec to } \sigma\text{-basis} \quad \Bigg| \quad \begin{array}{c} \text{---} \\ \text{---} \end{array} \boxed{T_{\sigma \rightarrow c}} = \begin{array}{c} \text{---} \\ \text{---} \end{array} \begin{array}{c} \text{---} \\ \text{---} \end{array} \begin{array}{c} \text{---} \\ \text{---} \end{array} \begin{array}{c} \text{---} \\ \text{---} \end{array} \begin{array}{c} \text{---} \\ \text{---} \end{array} \begin{array}{c} \text{---} \\ \text{---} \end{array} \begin{array}{c} \text{---} \\ \text{---} \end{array} \quad (b) \text{ Row-vec to } \sigma\text{-basis} \quad (2.62)$$

In the case where one wants to convert to row-vec convention, as previously shown the transformation is given by

$$T_{c \rightarrow r} = T_{r \rightarrow c} = \text{SWAP}. \quad (2.63)$$

One final important result that often arises when dealing with vectorized matrices is Roth's Lemma for the vectorization of the matrix product ABC [HJ85]. Given matrices $A, B, C \in L(\mathcal{X})$ we have

$$|ABC\rangle\rangle = (C^T \otimes A)|B\rangle\rangle \quad (2.64)$$

The graphical tensor network proof of this lemma is as follows:

$$\quad (2.65)$$

2.3 Representations of Completely Positive Maps

In this section we review several common mathematical descriptions for completely-positive trace-preserving maps, and show how several key properties may be captured graphically using the diagrammatic notation introduced in Section 2.2. The representations we will consider are the Kraus (or operator-sum) representation, the system-environment (or Stinespring) representation, the Liouville superoperator description based on vectorization of matrices, and the Choi-matrix or dynamical matrix description based on the Choi-Jamiołkowski isomorphism. We will also describe the often used process matrix (or χ -matrix) representation and show how this can be considered as a change of basis of the Choi-matrix. Following this, in Section 2.4 we will show how the graphical framework facilitates transformations between these representations as illustrated in Fig. 2.1.

2.3.1 Kraus Representation

One of the most commonly used representations of CPTP-maps is the *Kraus* [Kra83] or *operator-sum* [NC00] representation. This representation is particularly useful in phenomenological models of noise in quantum systems. Kraus's theorem states that a linear map $\mathcal{E} \in T(\mathcal{X}, \mathcal{Y})$ is CPTP if and only if it may be written in the form

$$\mathcal{E}(\rho) = \sum_{\alpha=1}^D K_{\alpha} \rho K_{\alpha}^{\dagger} \quad \text{where} \quad \sum_{\alpha=1}^D K_{\alpha}^{\dagger} K_{\alpha} = \mathbb{1}_{\mathcal{X}}. \quad (2.66)$$

The operators $\{K_\alpha : \alpha = 1, \dots, D\}$, $K_\alpha \in L(\mathcal{X}, \mathcal{Y})$ are called the *Kraus operators*. The Kraus representation of \mathcal{E} in Eq. (2.66) has the graphical representation

$$\text{---} \boxed{\mathcal{E}(\rho)} \text{---} = \text{---} \boxed{K} \text{---} \boxed{\rho} \text{---} \boxed{K^\dagger} \text{---} \quad (2.67)$$

The maximum number of Kraus operators needed for a Kraus description of \mathcal{E} is equal to the dimension of $L(\mathcal{X}, \mathcal{Y})$. For the case where $\mathcal{X} \cong \mathcal{Y} \cong \mathbb{C}^d$ the maximum number of Kraus operators is d^2 , and the minimum number case corresponds to unitary evolution where there is only a single Kraus operator.

It is important to note that the Kraus representation of \mathcal{E} is not unique as there is unitary freedom in choosing the Kraus operators. We can give preference to a particular representation called the *Canonical Kraus Representation* [BZ06] which is the unique set of Kraus operators satisfying the orthogonality relation $\text{Tr}[K_\alpha^\dagger K_\beta] = \lambda_\alpha \delta_{\alpha\beta}$. The canonical Kraus representation will be important when transforming between representations in Section 2.4.

2.3.2 System-Environment Representation

The second representation of CPTP-maps we consider is the system-environment model [NC00], which is typically considered the most physically intuitive description of open system evolution. This representation is closely related to (and sometimes referred to as) the *Stinespring* representation as it can be thought of as an application of the Stinespring dilation theorem [Sti55], which we also describe in this section. In this model we consider a system of interest \mathcal{X} , called the *principal system*, coupled to an additional system \mathcal{Z} called the *environment*. The composite system of the principal system and environment is then treated as a closed quantum system which evolves unitarily. We recover the reduced dynamics on the principal system by performing a partial trace over the environment. Suppose the initial state of our composite system is given by $\rho \otimes |v_0\rangle\langle v_0| \in L(\mathcal{X} \otimes \mathcal{Z})$, where $|v_0\rangle \in \mathcal{Z}$ is the initial (pure) state of the environment. The joint evolution is described by a unitary operator $U \in L(\mathcal{X} \otimes \mathcal{Z})$ and the reduced evolution of the principal system's state ρ is given by

$$\mathcal{E}(\rho) = \text{Tr}_{\mathcal{Z}}[U(\rho \otimes |v_0\rangle\langle v_0|)U^\dagger] \quad (2.68)$$

For convenience we can assume that the environment starts in a pure state $|v_0\rangle$ as one can always purify an initial mixed state, and in practice one only need consider the case where the Hilbert space describing the environment has at most dimension d^2 for $\mathcal{X} \cong \mathbb{C}^d$ [NC00].

The system-environment representation of the CP-map \mathcal{E} may then be represented graphically as

$$(2.69)$$

The system-environment model is advantageous when modelling the environment as a physical system. However, care must be taken when ascribing physical reality to any particular model as the system-environment description is not unique. This is not surprising as many different physical interactions could give rise to the same reduced dynamics on the principal system. This freedom manifests in an ability to choose the initial state of the environment in the representation and then adjust the unitary operator accordingly. In practice, the system-environment model can be cumbersome for performing many calculations where the explicit dynamics of the environment system are irrelevant. The remaining descriptions, which we cast into diagrammatic form, may be more convenient in these contexts.

Note that the system-environment evolution for the most general case will be an isometry and this is captured in Stinespring's representation [Sti55]. Stinespring's dilation theorem states that a CP-map $\mathcal{E} \in C(\mathcal{X}, \mathcal{Y})$ can be written in the form

$$\mathcal{E}(\rho) = \text{Tr}_{\mathcal{Z}} [A\rho A^\dagger] \quad (2.70)$$

where $A \in L(\mathcal{X}, \mathcal{Y} \otimes \mathcal{Z})$ and the Hilbert space \mathcal{Z} has dimension at most equal to $L(\mathcal{X}, \mathcal{Y})$. The map \mathcal{E} is trace preserving if and only if $A^\dagger A = \mathbb{1}_{\mathcal{X}}$ [Sti55].

In the case where $\mathcal{Y} \cong \mathcal{X}$, the Hilbert space $\mathcal{X} \otimes \mathcal{Z}$ mapped into by the Stinespring operator A is equivalent to the joint system-environment space in the system-environment representation. Hence one may move from the system-environment description to the

Stinespring representation as follows:

where $|v_0\rangle \in \mathcal{Z}$ is the initial state of the environment, and we have defined the Stinespring operator

$$A = U \cdot (\mathbb{1}_{\mathcal{X}} \otimes |v_0\rangle), \quad (2.73)$$

This close relationship is why these two representations are often referred to by the same name, and as we will show in [Section 2.4.5](#), it is straight forward to construct a Stinespring representation from the Kraus representation. However, generating a full description of the joint system-environment unitary operator U from a Stinespring operator A is cumbersome. It involves an algorithmic completion of the matrix elements in the unitary U not contained within the subspace of the initial state of the environment [[BZ06](#)]. Since it usually suffices to define the action of U when restricted to the initial state of the environment, which by [Eq. \(2.73\)](#) is the Stinespring representation, this is often the only transformation one need consider.

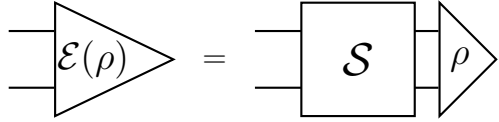
A further important point is that the evolution of the principal system $\mathcal{E}(\rho)$ is guaranteed to be CP if and only if the initial state of the system and environment is separable; $\rho_{\mathcal{X}\mathcal{Z}} = \rho_{\mathcal{X}} \otimes \rho_{\mathcal{Z}}$. In the case where the physical system is initially correlated with the environment, it is possible to have reduced dynamics which are non-completely positive [[WHE⁺04](#), [CTZ08](#)], we discuss this situation in greater detail in [Chapter 5](#).

2.3.3 Superoperator Representation

We now move to the *superoperator representation* of a CP-map $\mathcal{E} \in C(\mathcal{X}, \mathcal{Y})$. This representation is also referred to as the *Liouville representation* and the *linear representation*. The superoperator representation is based on the vectorization of the density matrix $\rho \mapsto |\rho\rangle\rangle_\sigma$ with respect to some orthonormal operator basis $\{\sigma_\alpha : \alpha = 0, \dots, d^2 - 1\}$ as introduced in [Section 2.2.2](#). Once we have chosen a vectorization basis (col-vec in our case) we define the superoperator for a map $\mathcal{E} \in T(\mathcal{X}, \mathcal{Y})$ to be the linear map

$$\mathcal{S} : \mathcal{X} \otimes \mathcal{X} \rightarrow \mathcal{Y} \otimes \mathcal{Y} : |\rho\rangle\rangle \mapsto |\mathcal{E}(\rho)\rangle\rangle \quad (2.74)$$

This is depicted graphically as



$$\text{Diagram: } \triangleleft \mathcal{E}(\rho) \triangleright = \square \mathcal{S} \square \triangleleft \rho \triangleright \quad (2.75)$$

We note that in situations where there may be ambiguity in what map the superoperator represents we may use the notation $\mathcal{S}_\mathcal{E}$ to represent the superoperator for a specific channel \mathcal{E} . This will be useful when we are dealing with expressions involving multiple channels.

In the col-vec basis we can express the evolution of a state ρ in terms of tensor components of \mathcal{S} as

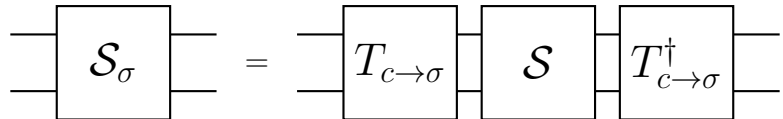
$$\mathcal{E}(\rho)_{mn} = \sum_{\mu\nu} \mathcal{S}_{nm,\nu\mu} \rho_{\mu\nu}. \quad (2.76)$$

For the case where $\mathcal{E} \in T(\mathcal{X})$, it is sometimes useful to change the basis of our superoperators from the col-vec basis to an orthonormal operator basis $\{\sigma_\alpha\}$ for $L(\mathcal{X})$. This is done using the basis transformation operator $T_{c \rightarrow \sigma}$ introduced in [Section 2.2.2](#). We have

$$\mathcal{S}_\sigma = T_{c \rightarrow \sigma} \cdot \mathcal{S} \cdot T_{c \rightarrow \sigma}^\dagger \quad (2.77)$$

$$= \sum_{\alpha\beta} \mathcal{S}_{\alpha\beta} |\sigma_\alpha\rangle\rangle \langle\langle \sigma_\beta|. \quad (2.78)$$

where the subscript σ indicates that \mathcal{S}_σ is the superoperator in the σ -vec convention. The tensor networks for this transformation is given by



$$\square \mathcal{S}_\sigma \square = \square T_{c \rightarrow \sigma} \square \square \mathcal{S} \square \square T_{c \rightarrow \sigma}^\dagger \square \quad (2.79)$$

Note that for a general map $\mathcal{E} \in T(\mathcal{X}, \mathcal{Y})$ we could do a similar construction but would need different bases for the initial and final Hilbert spaces $L(\mathcal{X})$ and $L(\mathcal{Y})$.

The structural properties the superoperator \mathcal{S} must satisfy for the linear map \mathcal{E} to be hermitian-preserving (HP), trace-preserving (TP), and completely positive (CP) are [BZ06]:

$$\mathcal{E} \text{ is HP} \iff \bar{\mathcal{S}} = \mathcal{S}^S \quad (2.80)$$

$$\iff \begin{array}{c} \text{---} \\ \text{---} \end{array} \boxed{\bar{\mathcal{S}}} \begin{array}{c} \text{---} \\ \text{---} \end{array} = \begin{array}{c} \text{---} \\ \text{---} \end{array} \boxed{\bar{\mathcal{S}}} \begin{array}{c} \text{---} \\ \text{---} \end{array} \quad (2.81)$$

$$\mathcal{E} \text{ is TP} \iff \mathcal{S}_{mm,nn} = \delta_{nn} \quad (2.82)$$

$$\iff \begin{array}{c} \text{---} \\ \text{---} \end{array} \boxed{\bar{\mathcal{S}}} \begin{array}{c} \text{---} \\ \text{---} \end{array} = \begin{array}{c} \text{---} \\ \text{---} \end{array} \quad (2.83)$$

$$\mathcal{E} \text{ is CP} \iff \mathcal{S}_{\mathcal{I} \otimes \mathcal{E}} |\rho_{AB}\rangle\rangle \geq 0 \quad \forall \rho_{AB} \geq 0 \quad (2.84)$$

Note that there is not a convenient structural criteria on the superoperator \mathcal{S} which specifies if \mathcal{E} is a CP-map. To test for positivity or complete positivity one generally uses the closely related *Choi-matrix* representation.

Superoperators are convenient to use for many practical calculations. Unlike the system-environment model the superoperator S is unique with respect to the choice of vectorization basis. Choosing an appropriate basis to express the superoperator in can often expose certain information about a quantum system. For example, if we want to model correlated noise for a mutli-partite system we can vectorize with respect to the mutli-qubit Pauli basis. Correlated noise would then manifest as non-zero entries in the superoperator corresponding to terms such as $\sigma_x \otimes \sigma_x$. We discuss in more detail how this may be done in Section 2.5.2.

2.3.4 Choi-Matrix Representation

The final representation shown in Fig. 2.1 is the *Choi matrix* [Cho75], or dynamical matrix [BZ06]. This is an application of the Choi-Jamiolkowski isomorphism which gives a bijection between linear maps and linear operators [Jam72]. Similarly to how vectorization mapped linear operators in $L(\mathcal{X}, \mathcal{Y})$ to vectors in $\mathcal{X} \otimes \mathcal{Y}$ or $\mathcal{Y} \otimes \mathcal{X}$, the Choi-Jamiolkowski isomorphism maps linear operators in $T(\mathcal{X}, \mathcal{Y})$ to linear operators in $L(\mathcal{X} \otimes \mathcal{Y})$ or $L(\mathcal{Y} \otimes \mathcal{X})$. The two conventions are

$$\text{col-}\Lambda : T(\mathcal{X}, \mathcal{Y}) \rightarrow L(\mathcal{X} \otimes \mathcal{Y}) : \mathcal{E} \mapsto \Lambda_c \quad (2.85)$$

$$\text{row-}\Lambda : T(\mathcal{X}, \mathcal{Y}) \rightarrow L(\mathcal{Y} \otimes \mathcal{X}) : \mathcal{E} \mapsto \Lambda_r. \quad (2.86)$$

For $\mathcal{X} \cong \mathbb{C}^d$, the explicit construction of the Choi-matrix is given by

$$\Lambda_c = \sum_{i,j=0}^{d-1} |i\rangle\langle j| \otimes \mathcal{E}(|i\rangle\langle j|) \quad (2.87)$$

$$\Lambda_r = \sum_{i,j=0}^{d-1} \mathcal{E}(|i\rangle\langle j|) \otimes |i\rangle\langle j| \quad (2.88)$$

where $\{|i\rangle : i = 0, \dots, d-1\}$ is an orthonormal basis for \mathcal{X} .

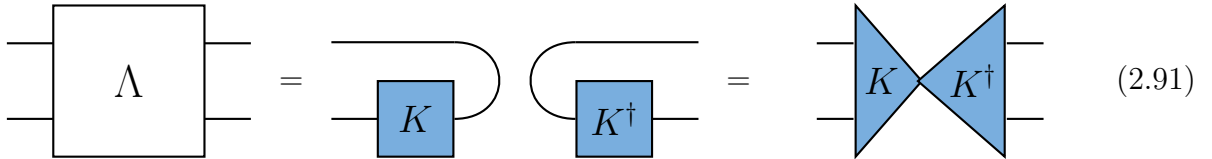
We call the two conventions col- Λ and row- Λ due to their relationship with the vectorization conventions introduced in [Section 2.2.2](#). The Choi-Jamiołkowski isomorphism can also be thought of as having a map $\mathcal{E} \in T(\mathcal{X}, \mathcal{Y})$ act on one half of an unnormalized Bell-state $|\Phi^+\rangle = \sum_i |i\rangle \otimes |i\rangle \in \mathcal{X} \otimes \mathcal{X}$, and hence these conventions corresponding to which half of the Bell state it acts on:

$$\Lambda_c = (\mathcal{I} \otimes \mathcal{E})|\Phi^+\rangle\langle\Phi^+| \quad (2.89)$$

$$\Lambda_r = (\mathcal{E} \otimes \mathcal{I})|\Phi^+\rangle\langle\Phi^+| \quad (2.90)$$

where $\mathcal{I} \in T(\mathcal{X})$ is the identity map. In what follows we will use the col- Λ convention and drop the subscript from Λ_c . We note that the alternative row- Λ Choi-matrix is naturally obtained by applying the bipartite-SWAP operation to Λ_c . In addition, as with the super-operator, we will sometimes use the notation that $\Lambda_{\mathcal{E}}$ to represent the Choi-Matrix for a specific channel \mathcal{E} .

As will be considered in [Section 2.4.3](#), if the evolution of the CP map \mathcal{E} is described by a Kraus representation $\{K_i\}$, then the Choi-Jamiołkowski isomorphism states that we construct the Choi-matrix by acting on one half of a bell state with the Kraus map as shown:



$$\Lambda = \text{Kraus } K = \text{Kraus } K^\dagger = \text{Kraus } K \text{ and } K^\dagger \quad (2.91)$$

Note that in general any tensor network describing a linear map \mathcal{E} , not just the Kraus description, may be contracted with one-half of the maximally entangled state $|\Phi^+\rangle\langle\Phi^+|$ to construct the Choi-matrix.

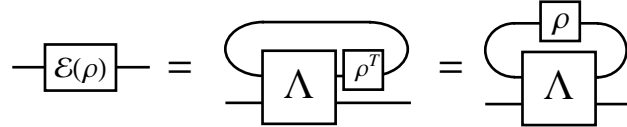
With the Choi-Jamiołkowski isomorphism defined, the evolution of a quantum state in terms of the Choi-matrix is then given by

$$\mathcal{E}(\rho) = \text{Tr}_{\mathcal{X}} [(\rho^T \otimes \mathbb{1}_{\mathcal{Y}})\Lambda] \quad (2.92)$$

or in terms of tensor components

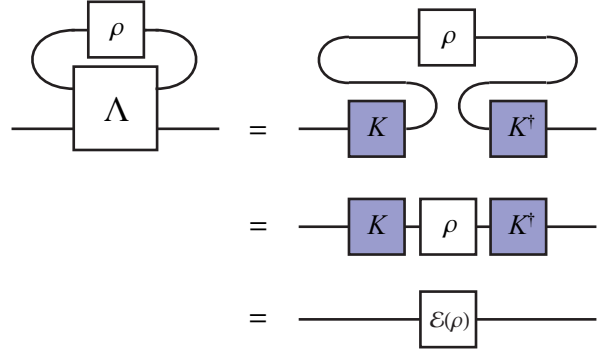
$$\mathcal{E}(\rho)_{mn} = \sum_{\mu,\nu} \Lambda_{\mu m, \nu n} \rho_{\mu\nu}. \quad (2.93)$$

The tensor network for Eq. (2.92) is given by



$$\text{---} \mathcal{E}(\rho) \text{---} = \text{---} \Lambda \text{---} \begin{array}{c} \text{---} \rho^T \text{---} \\ \text{---} \end{array} = \text{---} \Lambda \text{---} \begin{array}{c} \text{---} \rho \text{---} \\ \text{---} \end{array} \quad (2.94)$$

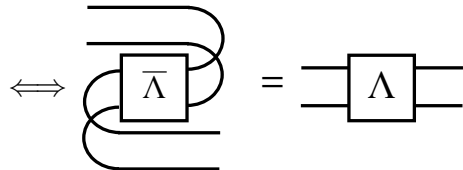
The graphical proof of Eq. (2.94) for the case where \mathcal{E} is described by a Kraus representation is as follows:



$$\begin{aligned} \text{---} \Lambda \text{---} \begin{array}{c} \text{---} \rho \text{---} \\ \text{---} \end{array} &= \text{---} K \text{---} \begin{array}{c} \text{---} \rho \text{---} \\ \text{---} \end{array} \text{---} K^\dagger \text{---} \\ &= \text{---} K \text{---} \rho \text{---} K^\dagger \text{---} \\ &= \text{---} \mathcal{E}(\rho) \text{---} \end{aligned} \quad (2.95)$$

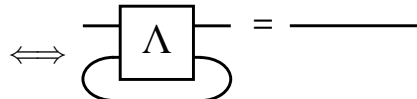
The structural properties the Choi-matrix Λ must satisfy for the linear map \mathcal{E} to be hermitian-preserving (HP), trace-preserving (TP), and completely positive (CP) are [BZ06]:

$$\mathcal{E} \text{ is HP} \iff \Lambda^\dagger = \Lambda \quad (2.96)$$



$$\iff \text{---} \bar{\Lambda} \text{---} \begin{array}{c} \text{---} \\ \text{---} \end{array} \begin{array}{c} \text{---} \\ \text{---} \end{array} = \text{---} \Lambda \text{---} \quad (2.97)$$

$$\mathcal{E} \text{ is TP} \iff \text{Tr}_Y[\Lambda] = \mathbb{1}_X \quad (2.98)$$



$$\iff \text{---} \Lambda \text{---} \begin{array}{c} \text{---} \\ \text{---} \end{array} = \text{---} \quad (2.99)$$

$$\mathcal{E} \text{ is CP} \iff \mathcal{E} \text{ is CP} \iff \Lambda \geq 0. \quad (2.100)$$

The Choi-matrix for a given map \mathcal{E} is unique with respect to the isomorphism convention chosen. We will provide tensor networks to illustrate a close relationship to the superoperator formed with the corresponding vectorization convention in [Section 2.4.1](#). The Choi-matrix finds practical utility as one can check the complete-positivity of the map \mathcal{E} by computing the eigenvalues of Λ . It is also necessary to construct the Choi-matrix for a given superoperator to transform to the other representations.

Due to the similarity of vectorization and the Choi-Jamiołkowski isomorphism, one could then ask what happens if we vectorize in a different basis. This change of basis of the Choi-matrix is more commonly known as the χ -matrix which we will discuss next. However, such a change of basis does not change the eigen-spectrum of a matrix, so the positivity criteria in [Eq. \(2.100\)](#) holds for any basis. Another desirable property of Choi matrices is that they can be directly determined for a given system experimentally by *ancilla assisted process tomography (AAPT)* [[DLP03](#), [ABJ⁺03](#)]. This is an experimental realization of the Choi-Jamiołkowski isomorphism which we discuss in detail in [Section 3.5.1](#).

2.3.5 Process Matrix Representation

As previously mentioned, one could consider a change of basis of the Choi-matrix analogous to that for the superoperator. The resulting operator is more commonly known as the χ -matrix or *process matrix* [[NC00](#)]. Consider Hilbert spaces $\mathcal{X} \cong \mathbb{C}^{d_x}$, $\mathcal{Y} \cong \mathbb{C}^{d_y}$ and let $D = d_x d_y$, and $\mathcal{Z} \cong \mathbb{C}^D$. If one chooses an orthonormal operator basis $\{\sigma_\alpha : \alpha = 0, \dots, D-1\}$ for $L(\mathcal{X}, \mathcal{Y})$, then a CPTP map $\mathcal{E} \in C(\mathcal{X}, \mathcal{Y})$ may be expressed in terms of a matrix $\chi \in L(\mathcal{Z})$ as

$$\mathcal{E}(\rho) = \sum_{\alpha, \beta=0}^{D-1} \chi_{\alpha\beta} \sigma_\alpha \rho \sigma_\beta^\dagger \quad (2.101)$$

where the process matrix χ is unique with respect to the choice of basis $\{\sigma_\alpha\}$.

The process matrix with respect to an orthonormal operator basis $\{\sigma_\alpha\}$ is related to the Choi matrix by the change of basis

$$\chi = T_{c \rightarrow \sigma} \cdot \Lambda \cdot T_{c \rightarrow \sigma}^\dagger \quad (2.102)$$

$$\Rightarrow \Lambda = \sum_{\alpha, \beta} \chi_{\alpha\beta} |\sigma_\alpha\rangle\rangle \langle\langle \sigma_\beta| \quad (2.103)$$

where $T_{c \rightarrow \sigma}$ is the vectorization change of basis operator introduced in [Section 2.2.2](#). Thus

evolution in terms of the χ -matrix is analogous to our Choi evolution as shown below:

$$\mathcal{E}(\rho) = \text{---} \left[T_{c \rightarrow \sigma}^\dagger \quad \chi \quad T_{c \rightarrow \sigma} \right] \text{---} \quad (2.104)$$

Starting with the expression for process matrix evolution in Eq. (2.101), the graphical proof asserting the validity of Eq. (2.102) is as follows

$$\begin{aligned} \mathcal{E}(\rho) &= \text{---} \left[\sigma \quad \rho \quad \sigma^\dagger \right] \text{---} \\ &\quad \left[\chi \right] \\ &= \text{---} \left[\sigma \quad \chi \quad \sigma^\dagger \right] \text{---} \quad (2.105) \\ &= \text{---} \left[T_{c \rightarrow \sigma}^\dagger \quad \chi \quad T_{c \rightarrow \sigma} \right] \text{---} \\ &= \text{---} \left[\Lambda \right] \text{---} \end{aligned}$$

We also see that if one forms the process matrix with respect to the col-vec basis $\sigma_\alpha = E_{j,i}$ where $\alpha = i + dj$ and d is the dimension of \mathcal{H} , then we have $\chi = \Lambda$.

Since the process matrix is a unitary transformation of the Choi-matrix, it shares the same structural conditions for hermitian preservation and complete-positivity as for the Choi-matrix given in Eq. (2.96) and Eq. (2.100) respectively. The condition for it to be trace preserving may be written in terms of the matrix elements and basis however. These conditions are

$$\mathcal{E} \text{ is TP} \iff \text{Tr}_y [T_{c \rightarrow \sigma}^\dagger \chi T_{c \rightarrow \sigma}] = \mathbb{1}_X \quad (2.106)$$

$$\iff \sum_{\alpha, \beta} \chi_{\alpha, \beta} \sigma_\alpha^T \bar{\sigma}_\beta = \mathbb{1}_X \quad (2.107)$$

$$\mathcal{E} \text{ is HP} \iff \chi^\dagger = \chi \quad (2.108)$$

$$\mathcal{E} \text{ is CP} \iff \chi \geq 0. \quad (2.109)$$

To convert a process-matrix χ in a basis $\{\sigma_\alpha\}$ to another orthonormal operator basis $\{\omega_\alpha\}$, we may use the inverse transformation to the superoperator change of basis from Section 2.3.3. That is

$$\chi^\omega = T_{\sigma \rightarrow \omega} \cdot \chi^\sigma \cdot T_{\sigma \rightarrow \omega}^\dagger \quad (2.110)$$

$$= \sum_{\alpha, \beta} \chi_{\alpha, \beta}^\sigma |\sigma_\alpha\rangle\rangle_\omega \langle\langle \sigma_\beta |_\omega \quad (2.111)$$

where the superscripts σ, ω denote the basis of the χ -matrices. This is illustrated as

$$\boxed{\chi^\omega} = \boxed{T_{\sigma \rightarrow \omega}} \boxed{\chi^\sigma} \boxed{T_{\sigma \rightarrow \omega}^\dagger} \quad (2.112)$$

2.4 Transforming Between Representations

We now proceed to the task of describing how one may transform between the representations of completely-positive trace-preserving maps depicted in Fig. 2.1. In particular, the transformations depicted as solid arrows in Fig. 2.1 have succinct descriptions in the graphical calculus we introduced in Section 2.2. These transformations are based on the wire bending dualities for reshuffling, vectorization, and the Choi-Jamiołkowski isomorphism. While the remaining transformations depicted as dashed lines are not based on dualities, but rather non-linear decompositions, or constructions, they also have diagrammatic representations in our graphical calculus for completely positive maps.

2.4.1 Transformations Between the Choi-Matrix and Superoperator

The Choi-matrix and superoperator are naturally equivalent under the reshuffling wire bending duality introduced in Section 2.2.1. In the col (row) convention we may transform between the two by applying the bipartite col (row)-reshuffling operation R introduced in Section 2.2.1. Let $\mathcal{E} \in T(\mathcal{X}, \mathcal{Y})$ be an operator map, and let $\Lambda \in L(\mathcal{X} \otimes \mathcal{Y})$, and $\mathcal{S} \in L(\mathcal{X} \otimes \mathcal{X}, \mathcal{Y} \otimes \mathcal{Y})$ respectively be the corresponding Choi-matrix and superoperator for \mathcal{E} . Then we have

$$\Lambda = \mathcal{S}^R \quad \mathcal{S} = \Lambda^R \quad (2.113)$$

The tensor networks for these transformations using the col convention are

$$\begin{array}{c} \text{Diagram 1: } \Lambda \text{ box with loops} \\ \text{Diagram 2: } S \text{ box} \end{array} = \begin{array}{c} \text{Diagram 3: } S \text{ box with loops} \\ \text{Diagram 4: } \Lambda \text{ box} \end{array} \quad (2.114)$$

In terms of tensor components we have $\Lambda_{mn, \mu\nu} = \mathcal{S}_{\nu n, \mu m}$, where m, n and μ, ν index the standard bases for \mathcal{X} and \mathcal{Y} respectively. Graphical proofs of the relations $\Lambda^{R_c} = \mathcal{S}$ and

$\mathcal{S}^{R_c} = \Lambda$ are given below

(2.115)

To transfer between a χ -matrix with respect to an arbitrary operator basis, and a superoperator with respect to an arbitrary vectorization basis, we must first convert both to col-vec (or row-vec) convention and then proceed by reshuffling.

Note that reshuffling is its own inverse, ie $(\Lambda^R)^R = \Lambda$, hence the solid bi-directional arrow connecting the Choi-matrix and superoperator representations in Fig. 2.1. This is the only transformation between the representations we consider which is linear, bijective, and self-inverse.

2.4.2 Transformations to the Superoperator

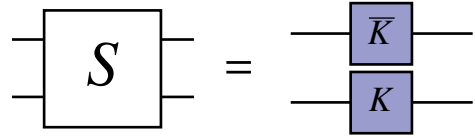
Transformations to the superoperator from the Kraus and system-environment representations of a CP-map are also accomplished by a wire-bending duality, in this case vector-

ization. However, unlike the bijective equivalence of the Choi-matrix and superoperator under the reshuffling duality, the vectorization duality is only surjective.

If we start with a Kraus representation for a CPTP map $\mathcal{E} \in C(\mathcal{X}, \mathcal{Y})$ given by $\{K_\alpha : \alpha = 0, \dots, D - 1\}$, with $K_\alpha \in L(\mathcal{X}, \mathcal{Y})$, we can construct the superoperator $\mathcal{S} \in L(\mathcal{X} \otimes \mathcal{X}, \mathcal{Y} \otimes \mathcal{Y})$ by

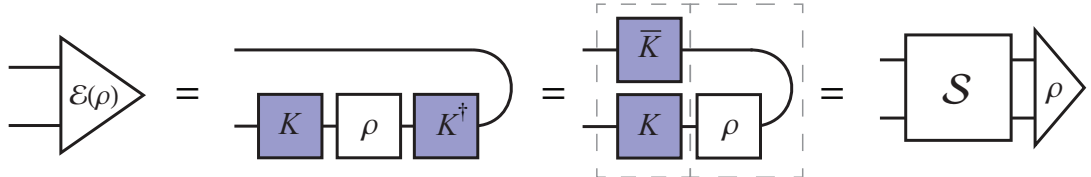
$$\mathcal{S} = \sum_{\alpha=0}^{D-1} \bar{K}_\alpha \otimes K_\alpha. \quad (2.116)$$

The corresponding tensor network is



$$\mathcal{S} = \begin{array}{c} \bar{K} \\ K \end{array} \quad (2.117)$$

and the graphical proof of this relationship follows directly from Roth's lemma:



$$\mathcal{E}(\rho) = \begin{array}{c} \bar{K} \\ K \end{array} \rho = \begin{array}{c} \bar{K} \\ K \end{array} \rho = \mathcal{S} \rho \quad (2.118)$$

Starting with a system-environment (or Stinespring) representation of a map $\mathcal{E} \in C(\mathcal{X}, \mathcal{Y})$ with input and output system Hilbert spaces $\mathcal{X} \cong \mathbb{C}^{d_x}$ and $\mathcal{Y} \cong \mathbb{C}^{d_y}$, respectively, and environment Hilbert space $\mathcal{Z} \cong \mathbb{C}^D$ with $1 \leq D \leq d_x d_y$, we may construct the superoperator for this map from the joint system-environment unitary U and initial environment state $|v_0\rangle$ by

$$\mathcal{S} = \sum_{\alpha} \langle \alpha | \bar{U} | v_0 \rangle \otimes \langle \alpha | U | v_0 \rangle, \quad (2.119)$$

where $\{|\alpha\rangle : \alpha = 0, \dots, D - 1\}$ is an orthonormal basis for \mathcal{Z} . The corresponding tensor

network is

(2.120)

As with the Kraus to superoperator transformation, the proof of [Eq. \(2.119\)](#) follows from Roth’s lemma.

Note that while the vectorization wire bending duality is invertible, these transformations to the superoperator from the Kraus and system-environment representations are single directional. In both cases injectivity fails as the superoperator is unique, while both the Kraus and system-environment representations are not. Hence we have solid single directional arrows in [Fig. 2.1](#) connecting both the Kraus and system-environment representations to the superoperator. The inverse transformation from a superoperator to the Kraus or system-environment representation requires a canonical decomposition of the operator \mathcal{S} (via first reshuffling to the Choi-matrix), which is detailed in [Sections 2.4.4](#) and [2.4.5](#).

2.4.3 Transformations to the Choi-Matrix Representation

Transforming to the Choi-matrix from the Kraus and system-environment representations is accomplished via a wire-bending duality which captures the Choi-Jamiołkowski isomorphism. As with the case of transforming to the superoperator, this duality transformation is surjective but not injective.

Given a set of Kraus matrices $\{K_\alpha : \alpha = 0, \dots, D - 1\}$ where $K_\alpha \in L(\mathcal{X}, \mathcal{Y})$ for a CPTP-map $\mathcal{E} \in C(\mathcal{X}, \mathcal{Y})$, one may form the Choi-Matrix Λ as was previously illustrated in [Eq. \(2.91\)](#) in [Section 2.3.4](#). In terms of both Dirac notation and tensor components we have:

$$\Lambda = \sum_{i,j} \left(|i\rangle\langle j| \otimes \sum_{\alpha} K_{\alpha} |i\rangle\langle j| K_{\alpha}^{\dagger} \right) \quad (2.121)$$

$$= \sum_{\alpha} |K_{\alpha}\rangle\rangle\langle\langle K_{\alpha}| \quad (2.122)$$

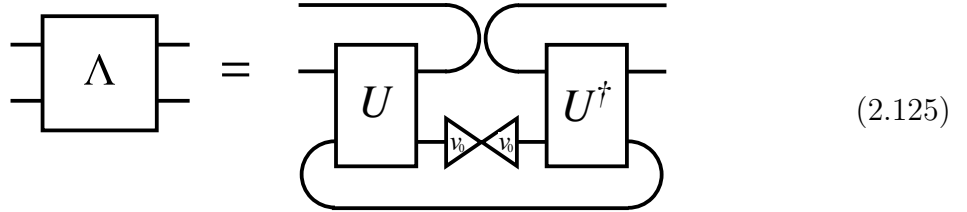
$$\Lambda_{mn,\mu\nu} = \sum_{\alpha} (K_{\alpha})_{\mu m} (\bar{K}_{\alpha})_{\nu n}. \quad (2.123)$$

where $\{|i\rangle\}$ is an orthonormal basis for \mathcal{X} , m, n index the standard basis for \mathcal{X} , and μ, ν index the standard basis for \mathcal{Y} .

Given a system-environment representation with joint unitary $U \in L(\mathcal{X} \otimes \mathcal{Z})$ and initial environment state $|v_0\rangle \in \mathcal{Z}$ we have

$$\Lambda = \sum_{i,j} (|i\rangle\langle j| \otimes \text{Tr}_{\mathcal{Z}} [U |i\rangle\langle j| \otimes |v_0\rangle\langle v_0| U^{\dagger}]) \quad (2.124)$$

Graphically this is given by



The proof of these transformations follow directly from the definition of the Choi-matrix in Eq. (2.87), and the tensor networks for the evolution via the Kraus or system-environment representations given in Eq. (2.67) and Eq. (2.69) respectively. As with the vectorization transformation to the superoperator discussed in Section 2.4.2, even though the Choi-Jamiołkowski isomorphism is linear these transformations are single directional as injectivity fails due to the non-uniqueness of both the Kraus and system-environment representations. Hence we have the solid single-directional arrows connecting both the Kraus and system-environment representations to the Choi-matrix in Fig. 2.1.

This completes our description of the linear transformations between the representations of CP-maps in Fig. 2.1. We will now detail the non-linear transformations to the Kraus and system environment representations.

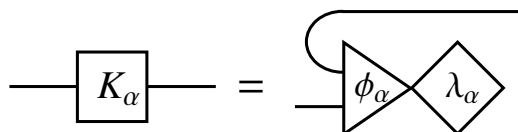
2.4.4 Transformations to the Kraus Representation

We may construct a Kraus representation from the Choi-matrix or system-environment representation by the non-linear operations of spectral-decomposition and partial trace decomposition respectively. To construct a Kraus representation from the Superoperator however, we must first reshuffle to the Choi-matrix.

To construct Kraus matrices from a Choi matrix we first recall the graphical Spectral decomposition we introduced as an example of our color summation convention in Eq. (2.10). If \mathcal{E} is CP, by Eq. (2.100) we have $\Lambda \geq 0$ and hence the spectral decomposition of the Choi-matrix is given by

$$\Lambda = \sum_{\alpha} \mu_{\alpha} |\phi_{\alpha}\rangle\langle\phi_{\alpha}|, \quad (2.126)$$

where $\mu_{\alpha} \geq 0$ are the eigenvalues, and $|\phi_{\alpha}\rangle$ the eigenvectors of Λ . Hence we can define Kraus operators $K_{\alpha} = \lambda_{\alpha} A_{\alpha}$ where $\lambda_{\alpha} = \sqrt{\mu_{\alpha}}$ and A_{α} is the unique operator satisfying $|A_{\alpha}\rangle\rangle = |\phi_{\alpha}\rangle$ as illustrated:



$$\text{---} \boxed{K_{\alpha}} \text{---} = \text{---} \begin{array}{c} \text{---} \\ \triangleleft \phi_{\alpha} \\ \diamond \lambda_{\alpha} \\ \text{---} \end{array} \text{---} \quad (2.127)$$

The number of Kraus operators will be equal to the rank r of the Choi-matrix, where

$1 \leq r \leq \dim(L(\mathcal{X}, \mathcal{Y}))$. The graphical proof of Eq. (2.126) is as follows:

The diagram illustrates the graphical proof of Eq. (2.126) through four stages of equivalence:

- Stage 1:** A box labeled $\mathcal{E}(\rho)$ with two horizontal lines passing through it.
- Stage 2:** A box labeled Λ with two horizontal lines passing through it. A box labeled ρ is positioned above Λ , with two curved lines connecting the top of Λ to the top of ρ .
- Stage 3:** A dashed box encloses the ρ box and the two curved lines. Below this dashed box, a diamond labeled λ^2 is connected to the horizontal lines. Two triangles labeled ϕ are placed on either side of the λ^2 diamond, with their vertices pointing towards it.
- Stage 4:** Two dashed boxes are shown. The left dashed box contains a triangle labeled λ pointing up and a box labeled A pointing down. The right dashed box contains a triangle labeled λ pointing up and a box labeled A^\dagger pointing down. The horizontal lines pass through A , ρ , and A^\dagger in sequence.
- Stage 5:** A box labeled K is placed above the left λ triangle, and a box labeled K^\dagger is placed above the right λ triangle. The horizontal lines pass through K , ρ , and K^\dagger in sequence.

(2.128)

The proof that Kraus operators satisfy the completeness relation follows from the trace

preserving property of Λ in Eq. (2.98):

$$\begin{aligned}
 & \text{---} \boxed{K^\dagger} \boxed{K} \text{---} = \text{---} \diamond \lambda \diamond \phi \text{---} \text{---} \diamond \phi \diamond \lambda \text{---} \\
 & = \text{---} \diamond \phi \diamond \lambda^2 \diamond \phi \text{---} \\
 & = \text{---} \boxed{\Lambda} \text{---} \\
 & = \text{---} \text{---} \\
 & = \text{---} \text{---}
 \end{aligned}
 \tag{2.129}$$

Note that since Λ , and the χ -matrix are related by a unitary change of basis, the Kraus representations constructed from their respective spectral decompositions will also be related by the same transformation. Each will give a unitarily equivalent Canonical Kraus representation of \mathcal{E} since the eigen-vectors are orthogonal. Thus we have described the arrow in Fig. 2.1 connecting the Choi-matrix to the Kraus representation. It is represented as a dashed arrow as it involves a non-linear decomposition, and is single directional as this representation transformation is injective, but not surjective. Surjectivity fails as we can only construct the canonical Kraus representations for \mathcal{E} . The reverse transformation is given by the Jamiolkowski isomorphism described in Section 2.4.3.

Starting with a system-environment representation with joint unitary $U \in L(\mathcal{X} \otimes \mathcal{Z})$ and initial environment state $|v_0\rangle \in \mathcal{Z}$, we first choose an orthonormal basis $\{|\alpha\rangle : \alpha =$

$0, \dots, D-1\}$ for \mathcal{Z} . We then construct the Kraus representation by decomposing the partial trace in this basis as follows

$$\mathcal{E}(\rho) = \text{Tr}_E [U (\rho \otimes |v\rangle\langle v|) U^\dagger] \quad (2.130)$$

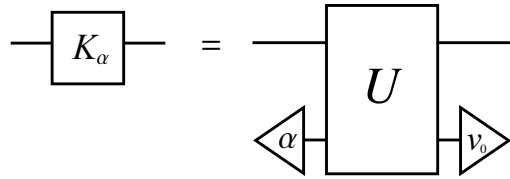
$$= \sum_{\alpha=0}^{D-1} \langle \alpha | U | v_0 \rangle \rho \langle v_0 | U^\dagger | \alpha \rangle \quad (2.131)$$

$$= \sum_{\alpha=0}^{D-1} K_\alpha \rho K_\alpha^\dagger. \quad (2.132)$$

Hence we may define Kraus matrices

$$K_\alpha = \langle \alpha | U | v_0 \rangle \quad (2.133)$$

leading to the tensor network



$$\quad (2.134)$$

The graphical proof of Eq. (2.133) and Eq. (2.134) is as follows

(2.135)

Though the Kraus and system-environment representations are both non-unique, for a fixed environment basis this partial trace decomposition is an injective transformation between the Kraus and Stinespring representations (or equivalently between the Kraus and system-environment representations when the joint unitary is restricted to a fixed initial state of the environment). To see this let $\{K_\alpha\}$ and $\{J_\alpha\}$ be two Kraus representations for a CPTP-map $\mathcal{E} \in C(\mathcal{X}, \mathcal{Y})$, constructed from Stinespring representations A and B

respectively. We have that

$$K_\alpha = J_\alpha \Leftrightarrow (K_\alpha)_{ij} = (J_\alpha)_{ij} \quad (2.136)$$

$$\Leftrightarrow A_{i\alpha,j} = B_{i\alpha,j} \quad (2.137)$$

$$\Leftrightarrow A = B. \quad (2.138)$$

Since the Stinespring operators satisfy $A = U|v_0\rangle$ and $B = V|v_0\rangle$ for some joint unitaries U and V , we must have that $U_0 = V_0$ where U_0 and V_0 are the joint unitaries restricted to the subspace of the environment spanned by $|v_0\rangle$.

This transformation can be thought of as the reverse application of the Stinespring dilation theorem, and hence for a fixed choice of basis (and initial state of the environment) it is invertible. The inverse transformation is the Stinespring dilation, and as we will show in [Section 2.4.5](#), since the inverse transformation is also injective this transformation is a bijection. However, since the partial trace decomposition involves a choice of basis for the environment it is non-linear — hence we use a dashed bi-directional arrow to represent the transformation from the system-environment representation to the Kraus representation in [Fig. 2.1](#).

2.4.5 Transformations to the System-Environment representation

We now describe the final remaining transformation given in [Fig. 2.1](#), the bijective non-linear transformation from the Kraus representation to the system-environment, or Stinespring, representation. The system-environment representation is the most cumbersome to transform to since it involves the unitary completion of a Stinespring dilation of a Kraus representation. Thus from a superoperator one must first reshuffle to the Choi-matrix, then from the Choi-matrix description one must then spectral decompose to the canonical Kraus representation, before finally constructing the system-environment as follows.

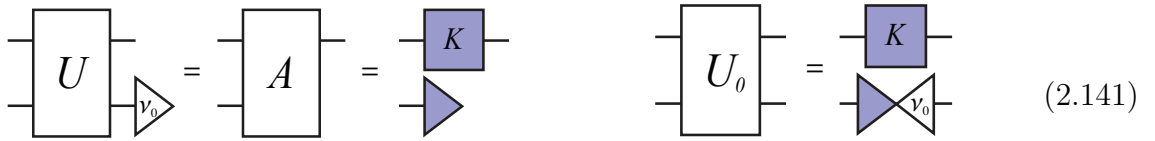
Let $\{K_\alpha : \alpha = 0, \dots, D-1\}$, where $1 \leq D \leq \dim(L(\mathcal{X}, \mathcal{Y}))$, be a Kraus representation for the CP-map $\mathcal{E} \in T(\mathcal{X}, \mathcal{Y})$. Consider an ancilla Hilbert space $\mathcal{Z} \cong \mathbb{C}^D$, this will model the environment. If we choose an orthonormal basis for the environment, $\{|\alpha\rangle : \alpha = 0, \dots, D-1\}$, then by Stinespring's dilation theorem we may construct the Stinespring matrix for the CP map \mathcal{E} by

$$A = \sum_{\alpha=0}^{D-1} K_\alpha \otimes |\alpha\rangle. \quad (2.139)$$

Recall from [Section 2.3.2](#) that the Stinespring representation is essentially the system-environment representation when the joint unitary operator is restricted to the subspace spanned by the initial state of the environment. Hence if we let $|v_0\rangle \in \mathcal{Z}$ be the initial state of the environment system, then this restricted unitary is given by

$$U_0 = \sum_{\alpha} K_{\alpha} \otimes |\alpha\rangle\langle v_0|, . \tag{2.140}$$

The tensor networks for [Eq. \(2.139\)](#) and [Eq. \(2.140\)](#) are:

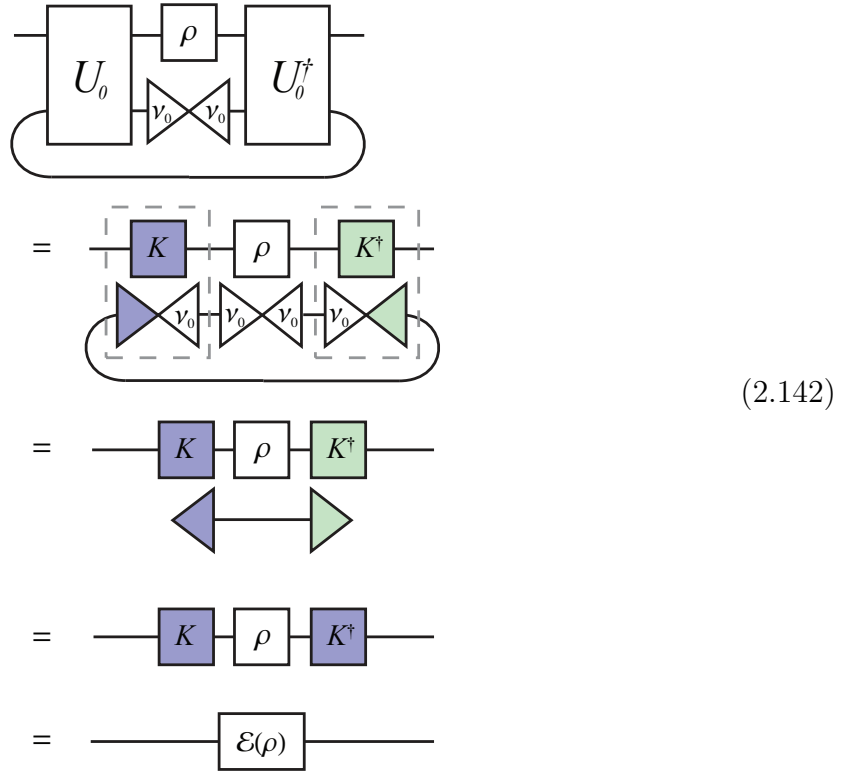


(a) Stinespring operator

(a) Restricted unitary

$$\tag{2.141}$$

The graphical proof that this construction gives the required evolution of a state ρ is as follows



$$\tag{2.142}$$

In principle, one may complete the remaining entries of this matrix to construct the full matrix description for the unitary U , however such a process is cumbersome and is unnecessary to describe the evolution of the CP-map \mathcal{E} [BZ06].

We have now finished characterizing the final transformations depicted in Fig. 2.1 connecting the Kraus representation to the system-environment representation by Stinespring dilation. As previously mentioned in Section 2.4.4, for a fixed choice of basis and initial state for the environment, the transformation between Kraus and Stinespring representations is bijective (and hence so is the transformation between Kraus and system-environment representations when restricted to the subspace spanned by the initial state of the environment). Though both these representations are non-unique, by fixing a basis and initial state for the environment we ensure that this transformation is injective. To see this let U_0 and V_0 be unitaries restricted to the state $|v_0\rangle$ constructed from Kraus representations, $\{K_\alpha\}$ and $\{J_\alpha\}$ respectively, for $\mathcal{E} \in C(\mathcal{X}, \mathcal{Y})$. Then

$$U_0 = V_0 \Leftrightarrow \sum_{\alpha} K_{\alpha} \otimes |\alpha\rangle\langle v_0| = \sum_{\alpha} J_{\alpha} \otimes |\alpha\rangle\langle v_0| \quad (2.143)$$

$$\Leftrightarrow \sum_{\alpha} K_{\alpha} \langle \beta | \alpha \rangle = \sum_{\alpha} J_{\alpha} \langle \beta | \alpha \rangle \quad (2.144)$$

$$\Leftrightarrow K_{\beta} = J_{\beta} \quad (2.145)$$

Bijectivity then follows from the injectivity of the inverse transformation — the previously given construction of a Kraus representation by the partial trace decomposition of a joint unitary operator in Eq. (2.134).

2.5 Composite Channels

We have now introduced all the basic elements of our graphical calculus for open quantum systems and shown how it may be used to graphically depict representations, and transformations between representations, for CP-maps. In this section we move onto more advanced applications of the graphical calculus. We will demonstrate how to apply vectorization to composite quantum systems, and in particular how to compose multiple superoperators together, and construct effective reduced superoperators from tracing out a subsystem. We also demonstrate the superoperator representation of various linear transformations of matrices. These constructions will be necessary for Chapter 3 where we use them for succinct proofs of several common quantities used in quantum information theory.

2.5.1 Vectorization of Composite Systems

We now describe how to deal with vectorization of the general case of composite system of N finite dimensional Hilbert spaces. Let $\mathcal{X}_k \cong \mathbb{C}^{d_k}$ be a d_k -dimensional complex Hilbert space, and let $\{|i_k\rangle : i_k = 0, \dots, d_k - 1\}$ be the standard basis for \mathcal{X}_k . We are interested in the composite system of N such Hilbert spaces,

$$\mathcal{X} = \mathcal{X}_1 \otimes \dots \otimes \mathcal{X}_N = \bigotimes_{k=1}^N \mathcal{X}_k \quad (2.146)$$

which has dimensions $D = \prod_{k=1}^N d_k$. Let $\{|\alpha\rangle : \alpha = 0, \dots, D-1\}$ be the computational basis for \mathcal{X} . We can consider vectors in \mathcal{X} and the dual space \mathcal{X}^\dagger as either 1st-order tensors where their single wire represents an index running over α , or as a N th-order tensor where each of the N wire corresponds to an individual Hilbert space \mathcal{X}_k . The correspondence between these two descriptions is made by the concatenation of the composite indices according to the lexicographical order

$$\alpha = \sum_{k=1}^N c(k) i_k \quad \text{where} \quad c(k) := \frac{D}{\prod_{j=1}^k d_j}. \quad (2.147)$$

Note that one could also consider the object as any order tensor between 1st and N th by the appropriate concatenation of some subset of the the wires.

We define the unnormalized Bell-state on the composite system $\mathcal{X} \otimes \mathcal{X}$ to be the state formed by the column (or row) vectorization of the identity operator $\mathbb{1}_{\mathcal{X}} \in L(\mathcal{X})$

$$\begin{aligned} |\mathbb{1}_{\mathcal{X}}\rangle\rangle &= \sum_{\alpha=0}^D |\alpha\rangle \otimes |\alpha\rangle \\ &= \text{align} \sum_{i_1=0}^{d_1-1} \dots \sum_{i_N=0}^{d_N-1} |i_1, \dots, i_N\rangle \otimes |i_1, \dots, i_N\rangle. \end{aligned} \quad (2.148)$$

where $|i_1, \dots, i_N\rangle := |i_1\rangle \otimes \dots \otimes |i_N\rangle$. The tensor network for this state is

$$\text{Curl} = \begin{matrix} N \\ \left\{ \begin{array}{c} \text{---} \\ \vdots \\ \text{---} \end{array} \right\} \\ N \\ \left\{ \begin{array}{c} \text{---} \\ \vdots \\ \text{---} \end{array} \right\} \end{matrix} \quad (2.149)$$

As with the single system case the column vectorization of a composite linear operator $A \in L(\mathcal{X}, \mathcal{Y})$, where $\mathcal{Y} = \bigotimes_{k=1}^N \mathcal{Y}_k$, is given by bending all the system wires upwards, or equivalently by the identity

$$|A\rangle\rangle \equiv (\mathbb{1} \otimes A)|\mathbb{1}\rangle\rangle. \quad (2.150)$$

Graphically this is given by

$$\quad (2.151)$$

Note that the order of the subsystems for the bent wires is preserved by the vectorization operation.

In some situations it may be preferable to consider vectorization of the composite system in terms of vectorization of the individual component systems. Transferring between this component vectorization and the joint-system vectorization can be achieved by an appropriate index permutation of vectorized operators [GTW09] which has a succinct graphical expression when cast in the tensor network framework.

Suppose the operator $A \in L(\mathcal{X}, \mathcal{Y})$, where $\mathcal{X} = \bigotimes_{k=1}^N \mathcal{X}_k$, $\mathcal{Y} = \bigotimes_{k=1}^N \mathcal{Y}_k$, is composed of subsystem operators such that

$$A = A_1 \otimes \dots \otimes A_N \quad (2.152)$$

where $A_k \in L(\mathcal{X}_k, \mathcal{Y}_k)$ for $k = 1, \dots, N$. As previously stated the vectorized composite operator $|A\rangle\rangle$ is a vector in the Hilbert space $\mathcal{X} \otimes \mathcal{Y}$.

We define an operation \mathcal{V}_N called the *unravelling* operation, the action of which unravels a vectorized matrix $|A\rangle\rangle = |A_1 \otimes \dots \otimes A_N\rangle\rangle$ into the tensor product of vectorized matrices on each individual subsystem $\mathcal{X}_k \otimes \mathcal{Y}_k$ [GTW09]

$$\mathcal{V}_N |A_1 \otimes \dots \otimes A_N\rangle\rangle = |A_1\rangle\rangle \otimes \dots \otimes |A_N\rangle\rangle \quad (2.153)$$

The inverse operation then undoes the unravelling

$$\mathcal{V}_N^{-1}(|A_1\rangle\rangle \otimes \dots \otimes |A_N\rangle\rangle) = |A_1 \otimes \dots \otimes A_N\rangle\rangle. \quad (2.154)$$

More generally the unravelling operation \mathcal{V}_N is given by the map

$$\mathcal{V}_N : |x_{\mathcal{X}}\rangle \otimes |y_{\mathcal{Y}}\rangle \longmapsto \bigotimes_{k=1}^N (|x_k\rangle \otimes |y_k\rangle) \quad (2.155)$$

where $|x_{\mathcal{X}}\rangle \equiv |x_1\rangle \otimes \dots \otimes |x_N\rangle$, $|y_{\mathcal{Y}}\rangle \equiv |y_1\rangle \otimes \dots \otimes |y_N\rangle$. Hence we can write \mathcal{V}_N in matrix form as

$$\mathcal{V}_N = \sum_{i_1, \dots, i_N} \sum_{j_1, \dots, j_N} |i_1, j_1, \dots, i_N, j_N\rangle \langle i_{\mathcal{X}}, j_{\mathcal{Y}}|. \quad (2.156)$$

where $|i_{\mathcal{X}}\rangle \equiv |i_1\rangle \otimes \dots \otimes |i_N\rangle$, $|j_{\mathcal{Y}}\rangle \equiv |j_1\rangle \otimes \dots \otimes |j_N\rangle$, and $|i_k\rangle, |j_l\rangle$ are the standard bases for \mathcal{X}_k and \mathcal{Y}_l respectively.

We can also express \mathcal{V}_N as the composition of SWAP operations between two systems. For the previously considered composite operator $A \in L(\mathcal{X}, \mathcal{Y})$ we have that $|A\rangle\rangle$ has $2N$ subsystems. If we label the SWAP operation between two subsystem Hilbert spaces indexed by k and l by $\text{SWAP}_{k:l}$, where $1 \leq k, l \leq 2N$, then the unravelling operation can be composed as

$$\mathcal{V}_N = W_{N-1} \dots W_1 \quad (2.157)$$

where

$$W_k = \prod_{j=0}^{k-1} \text{SWAP}_{N-k+2j+1:N-k+2j+2}. \quad (2.158)$$

For example

$$\begin{aligned} W_1 &= \text{SWAP}_{N:N+1} \\ W_2 &= \text{SWAP}_{N-1:N} \text{SWAP}_{N+1:N+2} \\ W_{N-1} &= \text{SWAP}_{2:3} \text{SWAP}_{4:5} \dots \text{SWAP}_{2N-2:2N-1}. \end{aligned} \quad (2.159)$$

While this equation looks complicated, it has a more intuitive construction when depicted graphically. The tensor networks for the unravelling operation in the $N = 2, 3$ and 4 cases

are shown below

(a) \mathcal{V}_2 (a) \mathcal{V}_3 (a) \mathcal{V}_4

(2.160)

We also present a graphical proof of this for the $N = 3$ case:

(2.161)

2.5.2 Composing Superoperators and Choi-Matrices

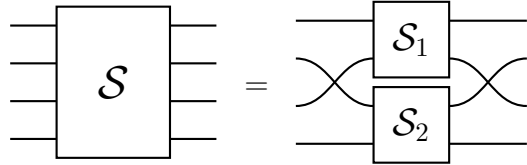
We now discuss how to compose superoperators and Choi-matrices on individual subsystems to form the correct operators on the composite system. As motivation for this we first consider the case of combining two superoperators \mathcal{S}_1 , and \mathcal{S}_2 . If we construct a joint system superoperator via tensor product ($\mathcal{S}_1 \otimes \mathcal{S}_2$) this composite operator acts on the tensor product of vectorized inputs $|\rho_1\rangle\rangle \otimes |\rho_2\rangle\rangle$, rather than the the vectorization of the composite input $|\rho_1 \otimes \rho_2\rangle\rangle$. To construct the correct composite superoperator for input $|\rho_1 \otimes \rho_2\rangle\rangle$ we may use the unravelling operation \mathcal{V}_N from Eq. (2.157) and its inverse.

Consider a set of channels $\{\mathcal{E}_k : k = 1, \dots, N\}$ where $\mathcal{E}_k \in T(\mathcal{X}_k, \mathcal{Y}_k)$ with superoperator and Choi-matrix \mathcal{S}_k and Λ_k respectively. The composite channel $\mathcal{E} \in T(\mathcal{X}, \mathcal{Y})$, where $\mathcal{X} = \bigotimes_{k=1}^N \mathcal{X}_k$, $\mathcal{Y} = \bigotimes_{k=1}^N \mathcal{Y}_k$ may be constructed in terms of the composite superoperator \mathcal{S} and Choi-matrix Λ as

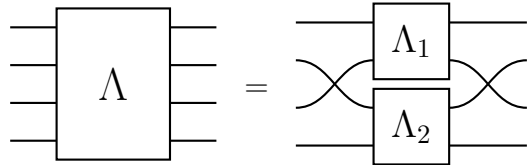
$$\mathcal{S} = \mathcal{V}_N^\dagger (\mathcal{S}_1 \otimes \dots \otimes \mathcal{S}_N) \mathcal{V}_N \quad (2.162)$$

$$\Lambda = \mathcal{V}_N^\dagger (\Lambda_1 \otimes \dots \otimes \Lambda_N) \mathcal{V}_N. \quad (2.163)$$

The reason that the transformation is the same for both superoperators and Choi-matrices, is that the Choi-matrix may be defined in terms of vectorization as was shown in Eq. (2.91). The tensor networks for these transformations in the $N = 2$ case are



$$(2.164)$$



$$(2.165)$$

For the $N = 3$ case they are

(2.166)

(2.167)

Composing channels from individual subsystem channels is useful when performing the same computations for multiple identical systems. For an example we consider vectorization in the Pauli-basis for an N -qubit system. While it is generally computationally more efficient to perform vectorization calculations in the col-vec (or row-vec) basis, as these may be implemented using structural operations on arrays, it is often convenient to express the superoperator in the Pauli basis, or the Choi-matrix in the χ -matrix representation, when we are interested in determining the form of correlated errors. However, transforming from the col-vec to the Pauli-basis for multiple (and possibly arbitrary) number of qubits is inconvenient. Using our unravelling operation we can instead compute the single qubit change of basis superoperator $T_{c \rightarrow \sigma}$ from Eq. (2.55), where $\sigma = \{\mathbb{1}, X, Y, Z\}/\sqrt{2}$ is the Pauli-basis for a single qubit, and use this to generate the transformation operator for multiple qubits. In the case of N -qubits we can construct the basis transformation matrix as

$$T_{c \rightarrow \sigma}^{(N)} = \mathcal{V}_N^\dagger \cdot T_{c \rightarrow \sigma}^{\otimes N} \cdot \mathcal{V}_N. \quad (2.168)$$

The composite system superoperator and χ -matrix in the Pauli-basis are then given by

$$\mathcal{S}^\sigma = T_{c \rightarrow \sigma}^{(N)} \cdot \mathcal{S} \cdot T_{c \rightarrow \sigma}^{(N)\dagger} \quad (2.169)$$

$$\chi = T_{c \rightarrow \sigma}^{(N)} \cdot \Lambda \cdot T_{c \rightarrow \sigma}^{(N)\dagger} \quad (2.170)$$

The same transformation can be used for converting a state $\rho = \rho_1 \otimes \dots \otimes \rho_N$ to the Pauli basis: $|\rho\rangle\rangle_\sigma = T_{c \rightarrow \sigma}^{(N)} |\rho\rangle\rangle_c$. These unravelling techniques are also useful for applying operations to a limited number of subsystems in a tensor network as used in many tensor network algorithms.

2.5.3 Reduced Channels

We now discuss the inverse case of [Section 2.5.2](#) and show how to construct reduced channels from channels on composite systems. We define a reduced channel to be a channel acting on a subsystem or some set of subsystems by *removing* the part of the channel acting on the remainder of the subsystems of a composite system. This removal may be achieved in several ways and we consider the following:

1. *Pre-selection*: Involves taking the action of a channel when a removed subsystem is preselected to be in a certain known state. This channel will be CPTP if the original composite channel is CPTP.
2. *Averaging over inputs*: Involves averaging the action of the channel over all possible input states of the removed subsystems. This can represent a lack of knowledge of the subsystems and can be written as a special case of pre-selection where we pre-select on the maximally mixed state.
3. *Post-selection*: Involves taking the action of a channel when a removed subsystem is measured and post-selected to be in a certain known state. This reduced channel will be CP if the composite channel is CP, however, since this involves a measurement the resulting channel will not generally be TP, and we may not simply renormalize as the probability of the post-selection succeeding will in general depend on the reduced channels input state.
4. *Averaging over outputs*: Involves averaging the action of the channel over all possible output states of the removed subsystems. As with averaging over inputs this can be written as a special case of post-selection where we post-select on the maximally mixed state, however unlike general post-selection, if the original channel is TP the reduced channel will be TP if multiply it be the dimension of \mathcal{Y}_2 . This is equivalent to performing a partial trace over the output.

We will now discuss how to implement each of these cases, and how to combine them to combine them. For simplicity we will restrict ourselves to the bipartite case where we have

a channel $\mathcal{E} \in T(\mathcal{X}, \mathcal{Y})$ across two subsystems with input and output spaces $\mathcal{X} = \mathcal{X}_1 \otimes \mathcal{X}_2$ and $\mathcal{Y} = \mathcal{Y}_1 \otimes \mathcal{Y}_2$ respectively.

We start with the case of pre-selection. Suppose the second subsystem can be assumed to initially be in the state $\rho_2 \in D(\mathcal{X}_2)$. We may construct the superoperator and Choi-matrix for the resulting effective channel $\mathcal{E}_{\rho_2}^{\text{pre}} \in T(\mathcal{X}_1, \mathcal{Y})$ as follows

Diagram (2.171) illustrates the equivalence between a pre-selected channel superoperator and a channel with a pre-selected input state. On the left, a box labeled $\mathcal{S}_{\rho_2}^{\text{pre}}$ has four input lines on the left and four output lines on the right. This is equal to a box labeled \mathcal{S} with four input lines on the left and four output lines on the right. The second output line from the \mathcal{S} box is connected to a triangular state symbol labeled ρ_2 , which then feeds back into the second input line of the \mathcal{S} box.

Diagram (2.172) illustrates the equivalence between a pre-selected channel Choi-matrix and a Choi-matrix with a pre-selected input state. On the left, a box labeled $\Lambda_{\rho_2}^{\text{pre}}$ has four input lines on the left and four output lines on the right. This is equal to a box labeled Λ with four input lines on the left and four output lines on the right. The second output line from the Λ box is connected to a rectangular state symbol labeled ρ_2 , which then feeds back into the second input line of the Λ box.

In terms of equations these are given by

$$\mathcal{S}_{\rho_2}^{\text{pre}} = \mathcal{S} \mathcal{V}_2^\dagger |\rho_2\rangle\rangle' \quad (2.173)$$

$$\Lambda_{\rho_2}^{\text{pre}} = \text{Tr}_{\mathcal{X}_2}[(\mathbb{1}_{\mathcal{X}_1} \otimes \rho_2^T \otimes \mathbb{1}_{\mathcal{Y}}) \Lambda_{\mathcal{E}}] \quad (2.174)$$

where we have assumed the identity on the first subsystem in $|\rho_2\rangle\rangle' \equiv \mathbb{1}_{\mathcal{X}_1} \otimes \mathbb{1}_{\mathcal{X}_1} \otimes |\rho_2\rangle\rangle$. It is easy to verify via the construction in Eq. (2.172) that the channel \mathcal{E} is a CPTP map, then the pre-selected reduced channel \mathcal{E}^{pre} is also CPTP.

In the case of constructing a post-selected channel we suppose that we measure the output state of second subsystem, and post select on finding it to be in the state $\tau_2 \in D(\mathcal{Y}_2)$. We may construct the superoperator and Choi-matrix for the resulting effective channel $\mathcal{E}_{\tau_2}^{\text{post}} \in T(\mathcal{X}, \mathcal{Y}_1)$ as follows

Diagram (2.175) illustrates the equivalence between a post-selected channel superoperator and a channel with a post-selected output state. On the left, a box labeled $\mathcal{S}_{\tau_2}^{\text{post}}$ has four input lines on the left and four output lines on the right. This is equal to a box labeled \mathcal{S} with four input lines on the left and four output lines on the right. The second output line from the \mathcal{S} box is connected to a triangular state symbol labeled τ_2 , which then feeds back into the second input line of the \mathcal{S} box.

$$(2.176)$$

In terms of equations these are given by

$$\mathcal{S}_{\tau_2}^{\text{post}} = \langle\langle \tau_2 |' \mathcal{V}_2 \mathcal{S} \quad (2.177)$$

$$\Lambda_{\tau_2}^{\text{post}} = \text{Tr}_{\mathcal{Y}_2}[(\mathbb{1}_{\mathcal{X}} \otimes \mathbb{1}_{\mathcal{Y}_1} \otimes \tau_2) \Lambda_{\mathcal{E}}] \quad (2.178)$$

where we have assumed the identity on the first subsystem in $\langle\langle \tau_2 |' \equiv \mathbb{1}_{\mathcal{Y}_1} \otimes \mathbb{1}_{\mathcal{Y}_1} \otimes \langle\langle \tau_2 |$. We note that in general this effective channel $\mathcal{E}_{\tau_2}^{\text{post}}$ is not properly normalized as per the conventions in [Section 2.3](#). Hence while $\mathcal{E}_{\tau_2}^{\text{post}}$ will be CP if \mathcal{E} is CP, it will in general be trace decreasing due to the probability of the post-selection succeeding. For it to be able to be renormalized to TP channel we would require $\text{Tr}_{\mathcal{Y}}[(\mathbb{1}_{\mathcal{X}} \otimes \mathbb{1}_{\mathcal{Y}_1} \otimes \tau_2) \Lambda_{\mathcal{E}}] = p \mathbb{1}_{\mathcal{X}}$.

For averaging over the inputs or outputs we use the result that for a state space $\mathcal{Z} \cong \mathbb{C}^d$, since density matrices form a convex set, averaging over states $\rho \in D(\mathcal{Z})$ is equivalent to averaging over pure states $|\psi\rangle\langle\psi| \in D(\mathcal{Z})$ with

$$\int d\rho \rho = \int d\psi |\psi\rangle\langle\psi| = \frac{\mathbb{1}}{d}. \quad (2.179)$$

If we average over input states $\rho_2 \in D(\mathcal{X}_2)$ over the second subsystem we may construct the superoperator and Choi-matrix for the resulting effective channel $\mathcal{E}_{\text{ave}}^{\text{pre}} \in T(\mathcal{X}_1, \mathcal{Y})$ as follows

$$(2.180)$$

$$\Lambda_{\text{ave}}^{\text{pre}} = \Lambda \diamond d_{x_2}^{-1} \quad (2.181)$$

In terms of equations these are given by

$$\mathcal{S}_{\text{ave}}^{\text{pre}} = \frac{1}{d_{x_2}} \mathcal{S} \mathcal{V}_2^\dagger |\mathbb{1}_{\mathcal{X}_2}\rangle\rangle' \quad (2.182)$$

$$\Lambda_{\text{ave}}^{\text{pre}} = \frac{1}{d_{x_2}} \text{Tr}_{\mathcal{X}_2}[\Lambda_{\mathcal{E}}] \quad (2.183)$$

where $\mathcal{X}_2 \cong \mathcal{C}^{d_{x_2}}$, and we have assumed the identity on the first subsystem in $|\mathbb{1}_{\mathcal{X}_2}\rangle\rangle' \equiv \mathbb{1}_{\mathcal{X}_1} \otimes \mathbb{1}_{\mathcal{X}_1} \otimes |\mathbb{1}_{\mathcal{X}_2}\rangle\rangle$.

Finally, if we average over output states $\tau_2 \in D(\mathcal{Y}_2)$ of the second subsystem we may construct the superoperator and Choi-matrix for the resulting effective channel $\mathcal{E}_{\text{ave}}^{\text{post}} \in T(\mathcal{X}, \mathcal{Y}_1)$ as follows

$$\mathcal{S}_{\text{ave}}^{\text{post}} = \mathcal{S} \quad (2.184)$$

$$\Lambda_{\text{ave}}^{\text{post}} = \Lambda \quad (2.185)$$

In terms of equations these are given by

$$\mathcal{S}_{\text{ave}}^{\text{post}} = \langle\langle \mathbb{1}_{\mathcal{Y}_2} |' \mathcal{V}_2 \mathcal{S} \quad (2.186)$$

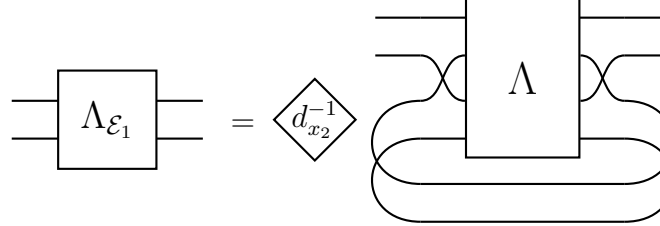
$$\Lambda_{\text{ave}}^{\text{post}} = \text{Tr}_{\mathcal{Y}_2}[\Lambda_{\mathcal{E}}] \quad (2.187)$$

and we have assumed the identity on the first subsystem in $\langle\langle \mathbb{1}_{\mathcal{Y}_2} |' \equiv \mathbb{1}_{\mathcal{Y}_1} \otimes \mathbb{1}_{\mathcal{Y}_1} \otimes \langle\langle \mathbb{1}_{\mathcal{Y}_2} |$.

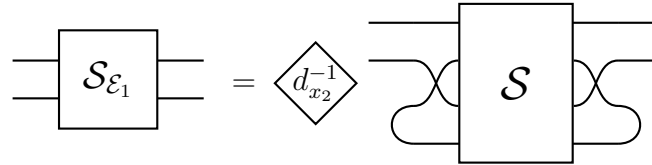
Combining the above constructions we may construct various reduced channels by averaging, pre-selecting or post-selecting channels. For more than two systems we may combine these with the results from [Section 2.5.2](#) to construct reduced channels from an arbitrary number of subsystems. We finish this section with two important cases of reduced channels.

2.5.4 Partial Trace of a Channel

We now formally define what we mean by the partial trace of a channel. This is a reduced channel where we average over both the input and output states of a subsystem. For a channel $\mathcal{E} \in T(\mathcal{X}, \mathcal{Y})$, with $\mathcal{X} = \mathcal{X}_1 \otimes \mathcal{X}_2, \mathcal{Y} = \mathcal{Y}_1 \otimes \mathcal{Y}_2$ the partial trace over subsystem 2 result is the effective channel $\mathcal{E}_1 \in T(\mathcal{X}_1, \mathcal{Y}_1)$



$$\Lambda_{\mathcal{E}_1} = \langle\langle d_{x_2}^{-1} \Lambda \quad (2.188)$$



$$\mathcal{S}_{\mathcal{E}_1} = \langle\langle d_{x_2}^{-1} \mathcal{S} \quad (2.189)$$

In terms of equations these are given by

$$\mathcal{S}_{\mathcal{E}_1} = \frac{1}{d_2} \langle\langle \mathbb{1}_{\mathcal{Y}_2} |' \mathcal{V}_2 \mathcal{S} \mathcal{V}_2^\dagger | \mathbb{1}_{\mathcal{Y}_2} \rangle\rangle' \quad (2.190)$$

$$\Lambda_{\mathcal{E}_1} = \frac{1}{d_2} \text{Tr}_2[\Lambda_{\mathcal{E}}] \equiv \frac{1}{d_2} \text{Tr}_{\mathcal{X}_2, \mathcal{Y}_2}[\Lambda_{\mathcal{E}}] \quad (2.191)$$

where $d_{y_2} = \dim(\mathcal{Y}_2)$ and Tr_2 is short hand for $\text{Tr}_{\mathcal{X}_2, \mathcal{Y}_2}$. This is simply a combination of the reduced channels for averaging over inputs and outputs of the second subsystem.

For the general case of a channel $\mathcal{E} \in T(\mathcal{X}, \mathcal{Y})$, with $\mathcal{X} = \bigotimes_{k=1}^N \mathcal{X}_k$, $\mathcal{Y} = \bigotimes_{k=1}^N \mathcal{Y}_k$ we have that the partial trace over subsystem j is given by

$$\Lambda_{\mathcal{E}_{\bar{j}}} = \frac{1}{d_{y_j}} \text{Tr}_j[\Lambda_{\mathcal{E}}] \equiv \frac{1}{d_{y_j}} \text{Tr}_{\mathcal{X}_j, \mathcal{Y}_j}[\Lambda_{\mathcal{E}}] \quad (2.192)$$

where $d_{y_j} = \dim(\mathcal{Y}_j)$, and we use the notation $\bar{j} = 1, \dots, j-1, j+1, \dots, N$, so that $\mathcal{E}_{\bar{j}}$ is the effective channel with subsystem j removed.

Equivalently we can define an effective channel $\mathcal{E}_j \in T(\mathcal{X}_j, \mathcal{Y}_j)$ acting only subsystem j by tracing over the remaining subsystems

$$\Lambda_{\mathcal{E}_j} = \frac{1}{D_{\bar{j}}} \text{Tr}_{\bar{j}}[\Lambda_{\mathcal{E}}] \quad (2.193)$$

where $D_{\bar{j}} \equiv \prod_{k=1, k \neq j}^N d_{y_k}$.

2.6 Matrix Operations as Channels

We now show how several common matrix manipulations can be written as channels in the superoperator and Choi-matrix representations. The superoperators are obtained by simply vectorizing the transformed operators, and the Choi-matrices can be obtained either by reshuffling or directly by wire bending manipulations. We begin with the trace superoperator S_{Tr} which implements the trace of a matrix $S_{\text{Tr}}|A\rangle\rangle \equiv \text{Tr}[A]$ for a square matrix $A \in L(\mathcal{X})$. This operation is simply given by the adjoint of the Bell-state:

$$S_{\text{Tr}} \equiv \langle\langle \mathbb{1}_{\mathcal{X}} |. \quad (2.194)$$

If \mathcal{X} is itself a composite system, we simply use the definition of the Bell-state for composite systems from Eq. (2.148). This is illustrated in our graphical calculus as

$$S_{\text{Tr}}|A\rangle\rangle = \text{Cup} \triangleleft A \triangleright \text{Cup} \Leftrightarrow S_{\text{Tr}} = \text{Cup} \quad (2.195)$$

In terms of the Choi-matrix, the trace operation is simply an identity $\Lambda_{\text{Tr}} = \mathbb{1}$, where the output space is the trivial vector space \mathbb{C} , $\Lambda_{\text{Tr}} \in L(\mathcal{X} \otimes \mathbb{C})$.

$$\Lambda_{\text{Tr}} = \text{Horizontal line} \quad (2.196)$$

where the dotted wire is added for convenience to represent the identity 1 of trivial output space \mathbb{C} .

For a rectangular matrix $B \in L(\mathcal{X}, \mathcal{Y})$, the transpose superoperator S_T which implements the transpose $S_T|B\rangle\rangle = |B^T\rangle\rangle$ is simply a swap superoperator between \mathcal{X} and \mathcal{Y} .

$$S_T = \text{SWAP} \quad (2.197)$$

$$\text{SWAP} : \mathcal{X} \otimes \mathcal{Y} \mapsto \mathcal{Y} \otimes \mathcal{X} \quad (2.198)$$

The tensor network for the swap superoperator is

$$S_T|B\rangle\rangle = \begin{array}{c} \diagup \quad \diagdown \\ \diagdown \quad \diagup \end{array} \triangleleft B \triangleright \quad \Leftrightarrow \quad S_T = \begin{array}{c} \diagdown \quad \diagup \\ \diagup \quad \diagdown \end{array} \quad (2.199)$$

If \mathcal{X} and \mathcal{Y} are composite vector spaces we may simply split the crossed wires into their respective subsystem wires.

By applying the reshuffling transformation to S_T one can see that the Choi-matrix for transposition is also the SWAP operator.

$$\Lambda_T = \begin{array}{c} \diagdown \quad \diagup \\ \diagup \quad \diagdown \end{array} \quad (2.200)$$

Next we give the superoperator representations of the bipartite matrix operations in Eq. (2.40) acting on vectorized square bipartite matrices $M \in L(\mathcal{X}_1 \otimes \mathcal{X}_2)$. These are the partial trace over \mathcal{X}_1 ($S_{\text{Tr}_{\mathcal{X}_1}}$) (and $S_{\text{Tr}_{\mathcal{X}_2}}$ over \mathcal{X}_2), transposition S_T , and col-reshuffling (S_{R_c}).

$$\mathcal{S}_{\text{Tr}_{\mathcal{X}_1}} : \mathcal{X}_1 \otimes \mathcal{X}_2 \otimes \mathcal{X}_1 \otimes \mathcal{X}_2 \mapsto \mathcal{X}_2 \otimes \mathcal{X}_2 \quad (2.201)$$

$$\mathcal{S}_{\text{Tr}_{\mathcal{X}_2}} : \mathcal{X}_1 \otimes \mathcal{X}_2 \otimes \mathcal{X}_1 \otimes \mathcal{X}_2 \mapsto \mathcal{X}_1 \otimes \mathcal{X}_1 \quad (2.202)$$

$$\mathcal{S}_{R_c} : \mathcal{X}_1 \otimes \mathcal{X}_2 \otimes \mathcal{X}_1 \otimes \mathcal{X}_2 \mapsto \mathcal{X}_1 \otimes \mathcal{X}_1 \otimes \mathcal{X}_2 \otimes \mathcal{X}_2 \quad (2.203)$$

$$\mathcal{S}_S : \mathcal{X}_1 \otimes \mathcal{X}_2 \otimes \mathcal{X}_1 \otimes \mathcal{X}_2 \mapsto \mathcal{X}_2 \otimes \mathcal{X}_1 \otimes \mathcal{X}_2 \otimes \mathcal{X}_1 \quad (2.204)$$

$$\mathcal{S}_T : \mathcal{X}_1 \otimes \mathcal{X}_2 \otimes \mathcal{X}_1 \otimes \mathcal{X}_2 \mapsto \mathcal{X}_1 \otimes \mathcal{X}_2 \otimes \mathcal{X}_1 \otimes \mathcal{X}_2 \quad (2.205)$$

The graphical representation of the superoperators for these operations are

$$\mathcal{S}_{\text{Tr}_{\mathcal{X}_1}} = \begin{array}{c} \text{---} \\ \diagdown \quad \diagup \\ \text{---} \end{array}, \quad \mathcal{S}_{\text{Tr}_{\mathcal{X}_2}} = \begin{array}{c} \text{---} \\ \diagup \quad \diagdown \\ \text{---} \end{array}, \quad \mathcal{S}_{R_c} = \begin{array}{c} \text{---} \\ \diagdown \quad \diagup \\ \text{---} \end{array}, \quad (2.206)$$

$$\mathcal{S}_S = \begin{array}{c} \diagdown \quad \diagup \\ \diagup \quad \diagdown \end{array}, \quad \mathcal{S}_T = \begin{array}{c} \diagdown \quad \diagup \\ \diagup \quad \diagdown \\ \diagdown \quad \diagup \\ \diagup \quad \diagdown \end{array} \equiv \begin{array}{c} \diagdown \quad \diagup \\ \diagup \quad \diagdown \\ \diagdown \quad \diagup \\ \diagup \quad \diagdown \end{array} \quad (2.207)$$

Algebraically they are given by

$$\mathcal{S}_{\text{Tr}_{\mathcal{X}}} = [\langle\langle \mathbb{1}_{\mathcal{X}} | \otimes \mathbb{1}_{\mathcal{Y}} \otimes \mathbb{1}_{\mathcal{Y}} \rangle\rangle \mathcal{V}_2] \quad (2.208)$$

$$\mathcal{S}_{\text{Tr}_{\mathcal{Y}}} = [\mathbb{1}_{\mathcal{X}} \otimes \mathbb{1}_{\mathcal{X}} \otimes \langle\langle \mathbb{1}_{\mathcal{Y}} | \rangle\rangle \mathcal{V}_2] \quad (2.209)$$

$$\mathcal{S}_{R_c} = \mathcal{V}_2 \quad (2.210)$$

$$\mathcal{S}_S = \text{SWAP}_{1:2} \text{SWAP}_{3:4} \quad (2.211)$$

$$\mathcal{S}_T = \text{SWAP}_{1:3} \text{SWAP}_{2:4} \quad (2.212)$$

where \mathcal{V}_2 is the unravelling operation from [Eq. \(2.157\)](#).

In terms of Choi-matrices we have

$$\Lambda_{\text{Tr}_{\mathcal{X}_1}} : L(\mathcal{X}_1 \otimes \mathcal{X}_2) \mapsto L(\mathcal{X}_2) \quad (2.213)$$

$$\Lambda_{\text{Tr}_{\mathcal{X}_2}} : L(\mathcal{X}_1 \otimes \mathcal{X}_2) \mapsto L(\mathcal{X}_1) \quad (2.214)$$

$$\Lambda_{R_c} : L(\mathcal{X}_1 \otimes \mathcal{X}_2) \mapsto L(\mathcal{X}_1 \otimes \mathcal{X}_1, \mathcal{X}_2 \otimes \mathcal{X}_2) \quad (2.215)$$

$$\Lambda_S : L(\mathcal{X}_1 \otimes \mathcal{X}_2) \mapsto L(\mathcal{X}_2 \otimes \mathcal{X}_1) \quad (2.216)$$

$$\Lambda_T : L(\mathcal{X}_1 \otimes \mathcal{X}_2) \mapsto L(\mathcal{X}_1 \otimes \mathcal{X}_2) \quad (2.217)$$

The graphical representation of these Choi-matrices are

$$\Lambda_{\text{Tr}_{\mathcal{X}_1}} = \begin{array}{c} \text{---} \\ \diagdown \quad \diagup \\ \text{---} \end{array}, \quad \Lambda_{\text{Tr}_{\mathcal{X}_2}} = \begin{array}{c} \text{---} \\ \diagup \quad \diagdown \\ \text{---} \end{array}, \quad \Lambda_{R_c} = \begin{array}{c} \text{---} \\ \diagdown \quad \diagup \\ \text{---} \end{array}, \quad \Lambda_S = \begin{array}{c} \diagdown \quad \diagup \\ \diagup \quad \diagdown \end{array}, \quad \Lambda_T = \begin{array}{c} \diagdown \quad \diagup \\ \diagup \quad \diagdown \\ \diagdown \quad \diagup \\ \diagup \quad \diagdown \end{array} \quad (2.218)$$

While these Choi-matrices have succinct graphical representations their algebraic forms are cumbersome and so we refrain from listing them.

In the general multipartite case for a composite matrix $A \in L(\mathcal{X})$ where $\mathcal{X} = \bigotimes_{k=1}^N \mathcal{X}_k$, we can trace out or transpose a subsystem j by using the unravelling operation in Eq. (2.157) to insert the appropriate superoperator for that subsystem with identity superoperators on the remaining subsystems:

$$\mathcal{S}_{O_j} = \mathcal{V}_{N-1}^{-1} \left[\left(\bigotimes_{k=1}^{j-1} \mathcal{S}_{I_k} \right) \otimes \mathcal{S}_O \times \left(\bigotimes_{k=j+1}^N \mathcal{S}_{I_k} \right) \right] \mathcal{V}_N$$

where $\mathcal{S}_O \in T(\mathcal{X}_j)$ is the superoperator acting on system j and $\mathcal{S}_{I_k} \in T(\mathcal{X}_k)$ is the identity superoperator for subsystem $L(\mathcal{X}_k)$. Similarly by inserting the appropriate operators at multiple subsystem locations we can perform the partial trace or partial transpose of any number of subsystems.

2.7 Summary

In this chapter we have reviewed the description of evolution of an open quantum system using the channel formalism, and in doing so we have cast it into the tensor network formalism the graphical calculus. The various representations of CP-maps reviewed, the transformation between them, and the graphical manipulations used to facilitate these transformation will be of use in later chapters. In addition we have also demonstrated the application of the graphical calculus for describing the vectorization of composite quantum systems to allow one to freely transform between a description of the vectorized composite system, and the composite system of individually vectorized systems. These tools are useful for constructing composite system superoperators and effective reduced system superoperators, and for applications where we wish to update or modify a subset of a composite system, and will be of particular use in Chapter 3 where we discuss channel fidelity measures and unitary designs, and in Chapter 4 where we develop a generalization of the quantum channel formalism.

The study of completely-positive trace-preserving maps is an old topic, so it is perhaps surprising that there are still new insights to be gained by investigating their structure using new techniques. Further, while the application of CPTP-maps to describing the evolution of open quantum systems is well understood, it is a surprisingly difficult task to find a concise summary of the properties of, and transformations between, the various

mathematically equivalent representations used in the quantum information processing literature. The graphical calculus for open quantum systems presented in this paper has enabled us to unify, and hence transform freely between, the various common representations of CPTP-maps by performing diagrammatic manipulations of their respective tensor networks. A summary of these transformations between the different representations was given in Fig. 2.1.

We found that many of these transformations between representations of CPTP-maps corresponded to wire bending dualities in our graphical calculus, which have a particularly succinct tensor network description. These transformations are depicted by solid arrows between two boxes labelling representations in Fig. 2.1. Of these duality transformations, only the reshuffling operation connecting the Choi-matrix and Liouville-superoperator is bi-directional — the reshuffling operation is bijective and self-inverse, and hence the same transformation takes the Choi-Matrix to the superoperator as taking the superoperator to the Choi-matrix. The two other wire bending dualities are vectorization, which transforms both the Kraus and system-environment representations to the superoperator representation, and the Choi-Jamiolkowski isomorphism, which transforms the same two representations to the Choi-matrix. These duality transformations are only single directional as they are not injective, and hence depicted by a one-way arrow connecting the appropriate boxes labelling these representations in Fig. 2.1. The reason these transformations are single directional, as apposed to the bi-directional transformation between the Choi-matrix and superoperator, is due to the non-uniqueness of the Kraus and system-environment representations of a CPTP-map. The transformation is a many-to-one (surjective) mapping and not strictly invertible without first specifying some form of decomposition of the superoperator or Choi-matrix.

The transformations we presented for converting from the Choi-matrix to the Kraus representation, and between the Kraus and system-environment representations, were not based solely on wire bending dualities. These transformations are depicted by dashed arrows in Fig. 2.1, where the dash is meant to indicate that they are non-linear transformations. This non-linearity arose from the decompositions and constructions involved, for example the spectral decomposition of a positive-semi definite operator in the Choi-matrix to Kraus representation transformation. In our case, these non-linear transformations were all also one directional due to the non-uniqueness of the representation being transformed to. There is unitary freedom in constructing them — for the Choi-matrix to Kraus representation transformation, one could change the basis of the eigenvectors with respect to a vectorization convention and still arrive at a valid Kraus representation; for Kraus to the system environment representation one may choose any orthonormal basis in the construction of the joint system-environment unitary in Eq. (2.140); and for the system environment

to Kraus representation one may decompose the partial trace over the environment in any orthonormal basis.

Chapter 3

Characterization of Open Quantum Systems

3.1 Introduction

The quantum channel formalism provides an operational description of the evolution of an open quantum system. In the real world however, this description is often not known a priori and must be reconstructed from experiments which characterize the noise and behaviour of a physical system. In this chapter we review several common methods and measures for the experimental characterization of the evolution of an open quantum system. While this aims to be a pedagogical overview of known techniques for characterizing channels, the summary presented here is clearer than what was found in the literature and collects together several different approaches under a unified framework. In particular we make use of the graphical calculus developed in [Chapter 2](#) to simplify several proofs of known expressions.

There are numerous measures for characterizing a quantum channel by comparing a noisy quantum channel to some ideal target. A good measure should have an operational interpretation related to evaluating the performance of an experiment, but in practice this must be balanced with a requirement that the measure be easy to compute and satisfy useful properties for theoretical reasoning as discussed in [\[GLN05\]](#). Three which we describe are the completely bounded (CB) trace norm [\[Kit97\]](#), entanglement fidelity [\[Sch96\]](#), and average gate fidelity [\[Nie02\]](#). The CB trace norm, which we review in [Section 3.2.1](#), is often also referred to as the *diamond norm* and has seen wide spread use in studies of quantum error correction and fault-tolerance [\[Kit97\]](#). This norm has several useful properties

including its stability with respect to the tensor product which makes it useful for reasoning about errors in many quantum information applications that involve reasoning with composite channels; it has an operational interpretation in terms of channel discrimination analogous to how the trace norm on states has an operational interpretation in terms of state discrimination [RW05, Wat08]; it may also be efficiently computed as the optimal value of a semidefinite program (SDP) [Wat09, VB96] which is a class is a convex optimization problem for which there exist efficient methods for solving [VB96, GB08, GB14]. We prove a new upper bound on the CB trace norm for Hermitian preserving operator maps which is defined in terms of the Choi-matrix. We also prove that this upper bound is tight for a class of operator maps where the canonical Kraus decomposition satisfies a condition that $K_\alpha^\dagger K_\alpha$ is diagonal while $K_\alpha^\dagger K_\beta$ for $\alpha \neq \beta$ has no diagonal components with respect to some orthonormal basis for the input Hilbert space.

Entanglement fidelity, which we review in Section 3.2.2, can be thought of as quantifying how well a channel preserves entanglement with an ancilla [Sch96, NC00]. Entanglement fidelity has closed form expressions given in terms of the Kraus representation [NC00] and Choi-matrix [FSW07]. In Section 3.2.2 we present a simple equivalent derivation in terms of the Choi-matrix representation of the channel using graphical techniques from Chapter 2. By applying the channel transformations of Section 2.4 we obtain expressions in terms of the other representations. Average gate fidelity, which we review in Section 3.2.3, is a widely used figure of merit for comparing a noisy channel to a target unitary channel. It has the nice property that it is simple to compute and gives a single parameter by which to evaluate the performance of a channel \mathcal{E} , regardless of the dimension of the input state space. Expressions for average gate fidelity have previously been given in terms of the Kraus representation [HHH99, Nie02], the superoperator [EAŻ05] and the Choi-matrix [JK11]. Using the graphical calculus developed in Chapter 2, we present an equivalent graphical derivation of the average gate fidelity in terms of the Choi-matrix which we believe is simpler than previous derivations.

One of the reasons that average gate fidelity is so widely used is because it can be efficiently estimated using an experimental partial characterization procedure known as *twirling* [EAŻ05, GAE07], and its robust generalization *randomized benchmarking* [MGE11, MGJ⁺12]. We briefly discuss these partial characterization schemes in Section 3.3 though without mention of error analysis and performance which can be readily found in the literature [WF14, GFC15]. In particular we present a simple derivation of twirling and RB as a *unitary 2-design* [RS09] using the graphical calculus developed in Chapter 2. A unitary design is a set of unitary operators which when summed over in a certain way returns the same result as averaging over the entire unitary group with respect to the uniform Haar measure [RS09]. In particular, the Clifford group is a unitary 2-design. This

is useful as the Clifford group has many nice properties for quantum information processing: it stabilizes the Pauli group, elements have a quadratic length representation, and there exist fast decomposition algorithms for these representations. In particular this means that approximate and exact unitary 2-design can be efficiently sampled and implemented [DCEL09, KS14, CLLW15].

In the absence of a partial characterization scheme to directly estimate a property of an unknown channel it is generally required to first obtain a complete description of the channel by reconstructing the Choi-matrix, or other representation of the channel, from the measurement outcomes of suitably chosen set of experiments. The first step of such a procedure is the reconstruction of a description of an unknown density operator by performing a set *quantum state tomography* which we describe in Section 3.4. This involves preparing many copies of a quantum state and subjecting each to one of a chosen set of *tomographically completely* set of measurements which span the state space of density matrices. The resulting count statistics can then be inverted to recover an estimation of the unknown state [NC00]. There are many variants of quantum state tomography in the literature including adaptive schemes [MRD⁺13], methods such as maximum likelihood estimation [Hra97, JKMW01], hedged maximum likelihood estimation [BK10a], compressed sensing [SKM⁺11, FGLE12], Bayesian mean estimation [HH12, BK10b], and other approaches have may have improved performance or error scaling in certain situations. In Section 3.4 we consider the simple case of linear inversion [NC00] and show how this is just a least squares fitting maximum likelihood estimation problem. We also consider the constrained MLE problem where one also adds the constraint that the density matrix is positive, and show how this optimization problem can be solved as a SDP. We believe this presentation of quantum state tomography is clearer than is currently presented in the literature, and in particular it unifies most tomographic protocols as special cases of this formalism.

In Section 3.5 we describe *quantum process tomography* [NC00, DLP01] which is the standard method of reconstructing a quantum channel by subjecting a tomographically complete set of input states to the unknown channel, and then performing state tomography on each of the outputs. We show how this can be interpreted as a special case of state tomography where the “state” being reconstructed is the positive semidefinite Choi-matrix of the known quantum channel. We also discuss two variants of quantum process tomography: ancilla assisted process tomography (AAPT)[Leu03, DLP03, ABJ⁺03], and randomized benchmarking process tomography [KdSR⁺14]. AAPT is related to a physical realization of the Choi-Jamiołkowski isomorphism where an experimenter prepares a joint state on the system and an ancilla, subjecting this state to the channel $\mathcal{I} \otimes \mathcal{E}$, where \mathcal{I} is the identity channel on the ancilla, and then performing state tomography to determine

the resulting output state. In the ideal case, if the input state is the maximally entangled bell state then the output state will be a normalized Choi-matrix for the unknown channel. However, the input need not be maximally entangled. A necessary and sufficient condition for the recovery of the channel is that the joint input state state has maximally Schmidt-number [ABJ⁺03, DLP03]. In Section 3.5.1 we present an equivalent though simpler to check condition on the input state, and a method of reconstructing the channel in terms of the reshuffling transformation between Choi-matrices and superoperators. Finally, randomized benchmarking process tomography is a recently proposed combination of interleaved randomized benchmarking and quantum process tomography which aims to be more robust to state preparation and measurement errors than standard quantum process tomography. In Section 3.5.2 we show how this also nicely fits within our framework as a special case of quantum state tomography.

3.2 Measures for Comparing Quantum Channels

There are numerous measures for characterizing a quantum channel by comparing a noisy quantum channel to some ideal target. The goal of a good measure is one which has a useful operational interpretation and significance for evaluating the performance of an experiment, but in practice this must be balanced with a requirement that the measure be easy to compute and satisfy useful properties for theoretical reasoning as discussed in [GLN05]. In this section we review three common measures, the completely bounded trace norm [Kit97], which has an operational interpretation in terms of channel discrimination; entanglement fidelity [Sch96], which can be thought of as quantifying how well a channel preserves entanglement with an ancilla; and average gate fidelity [Nie02], which is an easily computed figure of merit for comparing a channel to a target unitary channel which can be directly estimated using experimental protocols such as twirling and randomized benchmarking.

3.2.1 Completely Bounded Trace Norm

A widely used measure of distance between quantum channels used in quantum information theory is the *completely bounded* (CB) trace norm, often also referred to as the *diamond norm*, which has seen wide spread use in studies of quantum error correction and fault-tolerance [Kit97]. CB norms are defined in terms of *induced norms*, which are induced operator norms on channels analogous to operator norms induced by vector norms. The

induced p -norm of a channel $\mathcal{E} \in T(\mathcal{X}, \mathcal{Y})$ is given by

$$\|\mathcal{E}\|_p = \max \{ \|\mathcal{E}(X)\|_p : X \in L(\mathcal{X}), \|X\|_p \leq 1 \}. \quad (3.1)$$

The CB p -norm of a quantum channel $\mathcal{E} \in T(\mathcal{X}, \mathcal{Y})$ is then defined as the *stabilized* induced p -norm

$$\|\|\mathcal{E}\|\|_p = \sup_{k \geq 1} \|\mathbb{1}_{\mathbb{C}^k} \otimes \mathcal{E}\|_p \quad (3.2)$$

where the stability property ensures that

$$\|\|\mathcal{E} \otimes \mathcal{F}\|\|_p = \|\|\mathcal{E}\|\|_p \|\|\mathcal{F}\|\|_p. \quad (3.3)$$

which in general does not hold for the induced p -norm. The stability property of CB norms makes them preferable to induced forms for many quantum information applications that involve reasoning with composite channels.

The CB trace norm is the $p = 1$ case of Eq. (3.2) and we use the notation

$$\|\mathcal{E}\|_{\diamond} = \|\|\mathcal{E}\|\|_1 = \|\mathbb{1}_{\mathcal{X}} \otimes \mathcal{E}\|_1. \quad (3.4)$$

where the supremum over \mathbb{C}^k in Eq. (3.2) is obtained for $\mathbb{C}^k \cong \mathcal{X}$ for $p = 1$ [Kit97]. The CB trace norm is closely related to the *completely bounded norm* (CB norm) which has been studied extensively in operator theory [Pau02]. This norm corresponds to the $p = \infty$ case in Eq. (3.2):

$$\|\mathcal{E}\|_{\text{CB}} \equiv \|\|\mathcal{E}\|\|_{\infty} = \|\mathbb{1}_{\mathcal{Y}} \otimes \mathcal{E}\|_{\infty}. \quad (3.5)$$

where this time the supremum is obtained for $\mathbb{C}^k \cong \mathcal{Y}$ [Pau02]. The relationship between the CB trace norm and CB norm is given by [JKP09]

$$\|\mathcal{E}\|_{\diamond} = \|\mathcal{E}^{\dagger}\|_{\text{CB}} \quad (3.6)$$

where $\mathcal{E}^{\dagger} \in T(\mathcal{Y}, \mathcal{X})$ is the adjoint-channel of $\mathcal{E} \in T(\mathcal{X}, \mathcal{Y})$.

The CB trace norm has several useful properties. The first is that it has an operational interpretation in terms of channel discrimination analogous to how the trace norm on states has an operational interpretation in terms of state discrimination [RW05]. If we are given access to a copy of a quantum channel that we know is either \mathcal{E}_1 or \mathcal{E}_2 with probability λ and $1 - \lambda$ respectively, the optimal single shot strategy for distinguishing \mathcal{E}_1 and \mathcal{E}_2 succeeds with probability

$$p = \frac{1}{2} + \frac{1}{2} \|\lambda \mathcal{E}_1 - (1 - \lambda) \mathcal{E}_2\|_{\diamond}. \quad (3.7)$$

If we know nothing about the probability of being given channel \mathcal{E}_1 or \mathcal{E}_2 , then the least biased choice for λ is $1/2$ and we have an operationally relevant measure for the distance of channel \mathcal{E}_1 from \mathcal{E}_2 as $\|\mathcal{E}_1 - \mathcal{E}_2\|_\diamond$, corresponding to a probability of successful discrimination of

$$p = \frac{1}{2} + \frac{1}{4}\|\mathcal{E}_1 - \mathcal{E}_2\|_\diamond. \quad (3.8)$$

A second useful property of the CB trace norm is that while it looks difficult to compute due to the maximization, it can be computed as a semidefinite program (SDP), which is a special class of convex optimization problem for which sophisticated numerical algorithms exist for finding the optimum value [VB96]. SDPs are formally stated in terms of a *primal* and *dual* optimization problem, though sophisticated software packages exist which can convert an intuitive minimization problem with a convex constraint into an SDP and then solve the problem [GB08, GB14]. For certain classes of SDPs which satisfy a property called *strong duality*, the optimum value of the primal and dual problems are equal, and the solution is a global optimum [Wat09].

Computing the CB trace norm

Consider an operator map $\mathcal{E} \in T(\mathcal{X}, \mathcal{Y})$ which has Stinespring representation

$$\mathcal{E}(X) = \text{Tr}_{\mathcal{Z}}[A X B^\dagger]. \quad (3.9)$$

with two Stinespring operators $A, B \in L(\mathcal{X}, \mathcal{Y} \otimes \mathcal{Z})$. The fact that there are two Stinespring operators is because we do not assume that \mathcal{E} is CP, as in general we will be interested in the CB trace norm of an operator map defined in terms of the difference of two CP-maps: $\mathcal{E} = \mathcal{E}_1 - \mathcal{E}_2$. The square of the CB trace norm $\|\mathcal{E}\|_\diamond^2$ is given by the optimal value of the SDP [Wat09]

Primal Problem

$$\begin{aligned} &\text{maximize: } \text{Tr}[BB^\dagger W] \\ &\text{subject to: } \text{Tr}_{\mathcal{Y}}[W] = \text{Tr}_{\mathcal{Y}}[A\rho A^\dagger], \\ &\quad \rho \in D(\mathcal{X}) \\ &\quad W \in L(\mathcal{Y} \otimes \mathcal{Z}), W \geq 0. \end{aligned}$$

Dual Problem

$$\begin{aligned} &\text{minimize: } \|A^\dagger(\mathbb{1}_{\mathcal{Y}} \otimes Z)A\|_\infty \\ &\text{subject to: } \mathbb{1}_{\mathcal{Y}} \otimes Z \geq BB^\dagger, \\ &\quad Z \in L(\mathcal{Z}), Z \geq 0. \end{aligned}$$

Bounding the CB trace norm

The CB trace norm can be upper and lower bounded by the operator trace norm of the Choi-matrix representation of the channel. Consider an operator map $\mathcal{E} \in T(\mathcal{X}, \mathcal{Y})$ then

we have [WF14]

$$\|\mathcal{E}\|_{\diamond} \leq \|\Lambda_{\mathcal{E}}\|_1 \leq d\|\mathcal{E}\|_{\diamond} \quad (3.10)$$

or equivalently

$$\frac{1}{d}\|\Lambda_{\mathcal{E}}\|_1 \leq \|\mathcal{E}\|_{\diamond} \leq \|\Lambda_{\mathcal{E}}\|_1 \quad (3.11)$$

where $d = \dim(\mathcal{X})$.

We now prove a tighter upper bound on the CB trace norm, and also a special case where this bound is equal to the value of the CB trace norm.

Theorem 3.2.1. *Let $\mathcal{E} \in T(\mathcal{X}, \mathcal{Y})$ be a HP operator map. Then the CB trace norm of \mathcal{E} is upper bounded by*

$$\|\mathcal{E}\|_{\diamond} \leq \|\mathrm{Tr}_{\mathcal{Y}} |\Lambda_{\mathcal{E}}|\|_{\infty} \quad (3.12)$$

where $|A| \equiv \sqrt{A^{\dagger}A}$.

Theorem 3.2.2. *Let $\mathcal{E} \in T(\mathcal{X}, \mathcal{Y})$ be HP operator map with Choi-matrix*

$$\Lambda_{\mathcal{E}} = \sum_{\alpha} \lambda_{\alpha} |K_{\alpha}\rangle\rangle\langle\langle K_{\alpha}| \quad (3.13)$$

where $\lambda_{\alpha} \in \mathbb{R}$, and $\{|K_{\alpha}\rangle\rangle : \alpha = 1, \dots, \dim(\mathcal{X} \otimes \mathcal{Y})\}$ is an orthonormal basis for $\mathcal{X} \otimes \mathcal{Y}$ satisfying

$$K_{\alpha}^{\dagger}K_{\alpha} = \sum_i s_i^{\alpha} |x_i\rangle\langle x_i| \quad (3.14)$$

and that for all $|x_i\rangle$, and for all $\alpha \neq \beta$,

$$\langle x_i | K_{\alpha}^{\dagger} K_{\beta} | x_i \rangle = 0 \quad (3.15)$$

where $\{|x_i\rangle\}$ is an orthonormal basis for \mathcal{X} . Then the CB trace norm of \mathcal{E} is given by

$$\|\mathcal{E}\|_{\diamond} = \|\mathrm{Tr}_{\mathcal{Y}} |\Lambda_{\mathcal{E}}|\|_{\infty} = \max \left\{ \sum_{\alpha} s_i^{\alpha} |\lambda_{\alpha}| : i = 1, \dots, d_x \right\}. \quad (3.16)$$

The upper bound in [Theorem 3.2.1](#) is tighter than the bound in [Eq. \(3.11\)](#):

$$\|\mathrm{Tr}_{\mathcal{Y}} |\Lambda_{\mathcal{E}}|\|_{\infty} \leq \|\mathrm{Tr}_{\mathcal{Y}} |\Lambda_{\mathcal{E}}|\|_1 \leq \|\Lambda_{\mathcal{E}}\|_1. \quad (3.17)$$

In practice it is often *much* tighter, and for the cases covered by [Theorem 3.2.2](#) is tight. One of the cases covered by [Theorem 3.2.2](#) is a Choi-matrix, or Chi-matrix, which is diagonal in some separable basis:

$$\Lambda_{\mathcal{E}} = \sum_{i=1}^{d_x} \sum_{j=1}^{d_y} \lambda_{ij} |x_i\rangle\langle x_i| \otimes |y_j\rangle\langle y_j|. \quad (3.18)$$

where $\{|x_i\rangle\}, \{|y_j\rangle\}$ are orthonormal bases for \mathcal{X} and \mathcal{Y} respectively.

Proof of [Theorem 3.2.1](#). The *CB representation theorem* [[JKP09](#)] states that for any operator map $\mathcal{E} \in T(\mathcal{X}, \mathcal{Y})$, there exists a generalized Kraus representation $\mathcal{E}(X) = \sum_{\alpha} A_{\alpha} X B_{\alpha}^{\dagger}$ such that the CB norm is given by

$$\|\mathcal{E}\|_{\text{CB}}^2 = \left\| \sum_{\alpha} A_{\alpha} A_{\alpha}^{\dagger} \right\|_{\infty} \left\| \sum_{\alpha} B_{\alpha} B_{\alpha}^{\dagger} \right\|_{\infty} \quad (3.19)$$

As a corollary of the CB representation theorem we have [[JKP09](#)]

$$\|\mathcal{E}\|_{\text{CB}} = \inf \left\{ \left\| \sum_{\alpha} A_{\alpha} A_{\alpha}^{\dagger} \right\|_{\infty}^{1/2} \left\| \sum_{\alpha} B_{\alpha} B_{\alpha}^{\dagger} \right\|_{\infty}^{1/2} \right\}. \quad (3.20)$$

Hence for any Kraus representation

$$\|\mathcal{E}\|_{\text{CB}} \leq \left\| \sum_{\alpha} A_{\alpha} A_{\alpha}^{\dagger} \right\|_{\infty}^{1/2} \left\| \sum_{\alpha} B_{\alpha} B_{\alpha}^{\dagger} \right\|_{\infty}^{1/2}. \quad (3.21)$$

Now to prove our claim. Let $\mathcal{E} \in T(\mathcal{X}, \mathcal{Y})$ be a HP map, this implies that the Choi-matrix $\Lambda_{\mathcal{E}}$ is Hermitian. Hence by the spectral theorem we have that there exists an orthonormal basis $\{|K_{\alpha}\rangle\rangle\}$ for $\mathcal{X} \otimes \mathcal{Y}$ and real numbers λ_{α} such that

$$\Lambda_{\mathcal{E}} = \sum_{\alpha} \lambda_{\alpha} |K_{\alpha}\rangle\rangle\langle\langle K_{\alpha}|. \quad (3.22)$$

Thus we have $|\Lambda_{\mathcal{E}}| = \sum_{\alpha} |\lambda_{\alpha}| |K_{\alpha}\rangle\rangle\langle\langle K_{\alpha}|$, and

$$\text{Tr}_{\mathcal{Y}} |\Lambda_{\mathcal{E}}| = \sum_{\alpha} |\lambda_{\alpha}| \text{Tr}_{\mathcal{Y}} \left[|K_{\alpha}\rangle\rangle\langle\langle K_{\alpha}| \right] = \sum_{\alpha} |\lambda_{\alpha}| K_{\alpha}^T \bar{K}_{\alpha}. \quad (3.23)$$

We can define a canonical generalized Kraus representation for \mathcal{E} as

$$\mathcal{E}(X) = \sum_{\alpha} \lambda_{\alpha} K_{\alpha} X K_{\alpha}^{\dagger} = \sum_{\alpha} A_{\alpha} X B_{\alpha}^{\dagger} \quad (3.24)$$

where

$$A_\alpha = \frac{\lambda_\alpha}{\sqrt{|\lambda_\alpha|}} K_\alpha, \quad B_\alpha = \sqrt{|\lambda_\alpha|} K_\alpha. \quad (3.25)$$

The Choi-matrix of the adjoint channel \mathcal{E}^\dagger is then given by

$$\Lambda_{\mathcal{E}^\dagger} = \sum_{\alpha} \lambda_\alpha |K_\alpha^\dagger\rangle\rangle\langle\langle K_\alpha^\dagger| \quad (3.26)$$

$$\mathcal{E}^\dagger(\rho) = \sum_{\alpha} \lambda_\alpha K_\alpha^\dagger \rho K_\alpha = \sum_{\alpha} A_\alpha^\dagger \rho B_\alpha \quad (3.27)$$

Now, from [Eq. \(3.21\)](#) we have that

$$\|\mathcal{E}\|_{\diamond} = \|\mathcal{E}^\dagger\|_{\text{CB}} \leq \left\| \sum_{\alpha} A_\alpha^\dagger A_\alpha \right\|_{\infty}^{1/2} \left\| \sum_{\alpha} B_\alpha^\dagger B_\alpha \right\|_{\infty}^{1/2} = \left\| \sum_{\alpha} |\lambda_\alpha| K_\alpha^\dagger K_\alpha \right\|_{\infty} \quad (3.28)$$

Noting that $\|\bar{A}\|_{\infty} = \|A\|_{\infty}$ we have

$$\|\text{Tr}_{\mathcal{Y}} |\Lambda_{\mathcal{E}}|\|_{\infty} = \left\| \sum_{\alpha} |\lambda_\alpha| K_\alpha^T \bar{K}_\alpha \right\|_{\infty} = \left\| \sum_{\alpha} |\lambda_\alpha| K_\alpha^\dagger K_\alpha \right\|_{\infty} \quad (3.29)$$

and hence

$$\|\mathcal{E}\|_{\diamond} \leq \|\text{Tr}_{\mathcal{Y}} |\Lambda_{\mathcal{E}}|\|_{\infty}. \quad (3.30)$$

□

Proof of [Theorem 3.2.2](#). Let $\mathcal{E} \in T(\mathcal{X}, \mathcal{Y})$ be an operator map with Choi-matrix satisfying the conditions of [Theorem 3.2.2](#). Then

$$\text{Tr}_{\mathcal{Y}} |\Lambda_{\mathcal{E}}| = \sum_{\alpha} |\lambda_\alpha| \overline{K_\alpha^\dagger K_\alpha} = \sum_i \left(\sum_{\alpha} s_i^\alpha |\lambda_\alpha| \right) |x_i\rangle\langle x_i| \quad (3.31)$$

and hence

$$\|\text{Tr}_{\mathcal{Y}} |\Lambda_{\mathcal{E}}|\|_{\infty} = \max \left\{ \sum_{\alpha} s_i^\alpha |\lambda_\alpha| : i = 1, \dots, d_x \right\}. \quad (3.32)$$

Since \mathcal{E} is HP, from [Theorem 3.2.1](#) we have $\|\mathcal{E}\|_{\diamond} \leq \|\text{Tr}_{\mathcal{Y}} |\Lambda_{\mathcal{E}}|\|_{\infty}$. We now work to prove that we also have $\|\mathcal{E}\|_{\diamond} \geq \|\text{Tr}_{\mathcal{Y}} |\Lambda_{\mathcal{E}}|\|_{\infty}$. We do this using the primal problem of the SDP for the CB trace norm.

We can define a generalized Stinespring representation for \mathcal{E} , $\mathcal{E}(\rho) = \text{Tr}_{\mathcal{Z}}[A\rho B^\dagger]$ where $\mathcal{Z} \cong \mathcal{X} \otimes \mathcal{Y}$ and $A, B \in L(\mathcal{X}, \mathcal{Y} \otimes \mathcal{Z})$ are given by

$$A = \sum_{\alpha} \frac{\lambda_{\alpha}}{\sqrt{|\lambda_{\alpha}|}} K_{\alpha} \otimes |K_{\alpha}\rangle\rangle \quad (3.33)$$

$$B = \sum_{\alpha} \sqrt{|\lambda_{\alpha}|} K_{\alpha} \otimes |K_{\alpha}\rangle\rangle \quad (3.34)$$

$$(3.35)$$

Now we define operators $W_i \in L(\mathcal{Y} \otimes \mathcal{Z})$ and $\rho_i \in D(\mathcal{X})$

$$W_i = \sum_{\alpha, \beta} \sqrt{|\lambda_{\alpha}| |\lambda_{\beta}|} K_{\alpha} |x_i\rangle\langle x_i| K_{\beta}^{\dagger} \otimes |K_{\alpha}\rangle\rangle\langle\langle K_{\beta}| \quad (3.36)$$

$$\rho_i = |x_i\rangle\langle x_i| \quad (3.37)$$

in particular $W \geq 0$ and ρ is a density operator, and

$$\text{Tr}_{\mathcal{Y}}[W] = \sum_{\alpha, \beta} \sqrt{|\lambda_{\alpha} \lambda_{\beta}|} \text{Tr}_{\mathcal{Y}}[K_{\alpha} |x_i\rangle\langle x_i| K_{\beta}^{\dagger}] |K_{\alpha}\rangle\rangle\langle\langle K_{\beta}| \quad (3.38)$$

$$(3.39)$$

$$= \sum_{\alpha, \beta} \sqrt{|\lambda_{\alpha}| |\lambda_{\beta}|} \langle x_i | K_{\beta}^{\dagger} K_{\alpha} | x_i \rangle |K_{\alpha}\rangle\rangle\langle\langle K_{\beta}| \quad (3.40)$$

$$= \sum_{\alpha} s_i^{\alpha} |\lambda_{\alpha}| |K_{\alpha}\rangle\rangle\langle\langle K_{\alpha}| \quad (3.41)$$

Now we note

$$\text{Tr}_{\mathcal{Y}}[A \rho_i A^{\dagger}] = \sum_{\alpha, \beta} \frac{\lambda_{\alpha} \lambda_{\beta}}{\sqrt{|\lambda_{\alpha}| |\lambda_{\beta}|}} \langle x_i | K_{\beta}^{\dagger} K_{\alpha} | x_i \rangle |K_{\alpha}\rangle\rangle\langle\langle K_{\beta}| \quad (3.42)$$

$$= \sum_{\alpha} s_i^{\alpha} \frac{\lambda_{\alpha}^2}{|\lambda_{\alpha}|} |K_{\alpha}\rangle\rangle\langle\langle K_{\alpha}| \quad (3.43)$$

$$= \text{Tr}_{\mathcal{Y}}[W_i] \quad (3.44)$$

So W_i is primal feasible for the CB trace norm SDP, and hence

$$\|\mathcal{E}\|_{\diamond}^2 \geq \text{Tr}[B B^{\dagger} W_i] \quad (3.45)$$

for all $i = 1, \dots, \dim(\mathcal{X})$. Now to compute the value we have

$$\mathrm{Tr}[BB^\dagger W_i] = \sum_{\alpha, \beta, \gamma, \delta} \sqrt{|\lambda_\alpha| |\lambda_\beta|} \sqrt{|\lambda_\gamma| |\lambda_\delta|} \mathrm{Tr}[K_\alpha K_\beta^\dagger K_\gamma |x_i\rangle\langle x_i| K_\delta^\dagger] \langle\langle K_\beta | K_\gamma \rangle\rangle \langle\langle K_\delta | K_\alpha \rangle\rangle \quad (3.46)$$

$$= \sum_{\alpha, \beta} |\lambda_\alpha| |\lambda_\beta| \langle x_i | K_\alpha^\dagger K_\alpha K_\beta^\dagger K_\beta | x_i \rangle \quad (3.47)$$

$$= \sum_{\alpha, \beta} |\lambda_\alpha| |\lambda_\beta| s_i^\alpha s_i^\beta \quad (3.48)$$

$$= \left(\sum_{\alpha} s_i^\alpha |\lambda_\alpha| \right)^2 \quad (3.49)$$

Now choose i such that

$$i = \arg \max \left\{ \sum_{\alpha} s_i^\alpha |\lambda_\alpha| : i = 1, \dots, \dim(\mathcal{X}) \right\}. \quad (3.50)$$

Then we have

$$\mathrm{Tr}[BB^\dagger W_i] = \|\mathrm{Tr}_{\mathcal{Y}} |\Lambda_{\mathcal{E}}|\|_{\infty}^2 \quad (3.51)$$

and hence $\|\mathcal{E}\|_{\diamond} \geq \|\mathrm{Tr}_{\mathcal{Y}} |\Lambda_{\mathcal{E}}|\|_{\infty}$. Since we have

$$\|\mathrm{Tr}_{\mathcal{Y}} |\Lambda_{\mathcal{E}}|\|_{\infty} \leq \|\mathcal{E}\|_{\diamond} \leq \|\mathrm{Tr}_{\mathcal{Y}} |\Lambda_{\mathcal{E}}|\|_{\infty} \implies \|\mathcal{E}\|_{\diamond} = \|\mathrm{Tr}_{\mathcal{Y}} |\Lambda_{\mathcal{E}}|\|_{\infty}. \quad (3.52)$$

□

3.2.2 Entanglement Fidelity

When characterizing the performance of a noisy quantum channels a useful fidelity quantity is the *entanglement fidelity* which quantifies how well a channel preserves entanglement with an ancilla [Sch96, NC00]. For a CPTP map $\mathcal{E} \in C(\mathcal{X})$ and density matrix $\rho \in L(\mathcal{X})$ the entanglement fidelity is given by

$$F_e(\mathcal{E}, \rho) = \inf \{ F(|\psi\rangle\langle\psi|, (\mathcal{I}_{\mathcal{Z}} \otimes \mathcal{E})(|\psi\rangle\langle\psi|)) : \mathrm{Tr}_{\mathcal{Z}}[|\psi\rangle\langle\psi|] = \rho \} \quad (3.53)$$

where $|\psi\rangle \in \mathcal{X} \otimes \mathcal{Z}$ is a purification of ρ over an ancilla \mathcal{Z} . Entanglement fidelity turns out to be independent of the choice of purification $|\psi\rangle$, and a closed form expression has been given in terms of the Kraus representation [NC00] and Choi-matrix [FSW07]. Here

we present a simple equivalent derivation in terms of the Choi-matrix representation of the channel \mathcal{E} using graphical techniques. Then by applying the channel transformations of [Section 2.4](#) we obtain expressions in terms of the other representations.

The resulting expressions for entanglement fidelity are:

$$F_e(\mathcal{E}, \rho) = \langle\langle \rho | \Lambda | \rho \rangle\rangle \quad (3.54)$$

$$= \text{Tr} [(\rho^T \otimes \rho) \mathcal{S}] \quad (3.55)$$

$$= \sum_j |\text{Tr}[\rho K_j]|^2 \quad (3.56)$$

$$= \sum_{i,j} \chi_{ij} \text{Tr}[\rho \sigma_i] \text{Tr}[\rho \sigma_j^\dagger] \quad (3.57)$$

$$= \text{Tr}_{\mathcal{X}}[\rho A^\dagger] \cdot \text{Tr}_{\mathcal{X}}[A \rho] \quad (3.58)$$

where \mathcal{S} , Λ , $\{K_j\}$, χ , A are the superoperator, Choi-matrix, Kraus, χ -matrix and Strine-spring representations for \mathcal{E} respectively. In the case of the χ -matrix representation, χ is defined with respect to a basis $\{\sigma_j\}$ satisfying $\text{Tr}[\sigma_j] = \sqrt{d} \delta_{j,0}$.

For the graphical proof in terms of the Choi-representation we start with [Section 3.2.2](#) and perform the following tensor manipulations

$$\begin{aligned}
 F_e(\mathcal{E}, \rho) &= \inf_{\psi} \text{Diagram 1} \\
 &= \inf_{\psi} \text{Diagram 2} \\
 &= \inf_{\psi} \text{Diagram 3}
 \end{aligned} \quad (3.59)$$

Now since the infimum is over all $|\psi\rangle \in \mathcal{Z} \otimes \mathcal{X}$ satisfying $\text{Tr}_{\mathcal{Z}}[|\psi\rangle\langle\psi|] = \rho$ the result is

independent of the specific purification ψ and we have:

$$\begin{aligned}
 F_e(\mathcal{E}, \rho) &= \inf_{\psi} \text{Diagram 1} \\
 &= \text{Diagram 2} \\
 &= \text{Diagram 3}
 \end{aligned}
 \tag{3.60}$$

The transformations to the other representations are given by the tensor networks

$$\begin{aligned}
 \text{Diagram 3} &= \text{Diagram 4} \\
 &= \text{Diagram 5}
 \end{aligned}
 \tag{3.61}$$

$$\begin{aligned}
 \text{Diagram 3} &= \text{Diagram 6} \\
 &= \text{Diagram 7}
 \end{aligned}
 \tag{3.62}$$

$$\text{Diagram 3} = \text{Diagram 8}
 \tag{3.63}$$

$$\begin{aligned}
 \text{Diagram 3} &= \text{Diagram 9} \\
 &= \text{Diagram 10}
 \end{aligned}
 \tag{3.64}$$

3.2.3 Average Gate Fidelity

Perhaps the most widely used metric for the performance of a quantum channel is *Gate Fidelity*, which characterizes the closeness of a CPTP map $\mathcal{E} \in C(\mathcal{X})$ to a desired quantum channel $\mathcal{F} \in C(\mathcal{X})$. Gate fidelity is defined to be

$$F_{\mathcal{E},\mathcal{F}}(\rho) = F(\mathcal{F}(\rho)\mathcal{E}(\rho)) \quad (3.65)$$

where F is the fidelity function for quantum states given in Eq. (1.27).

In general we are interested in comparing a \mathcal{E} to a unitary channel $\mathcal{U} \in C(\mathcal{X})$ where $\mathcal{U}(\rho) = U\rho U^\dagger$. In this case we have

$$F_{\mathcal{E},\mathcal{U}}(\rho) = \left[\text{Tr} \sqrt{\sqrt{U\rho U^\dagger} \mathcal{E}(\rho) \sqrt{U\rho U^\dagger}} \right]^2 \quad (3.66a)$$

$$= \left[\text{Tr} \sqrt{\sqrt{\rho} U^\dagger \mathcal{E}(\rho) U \sqrt{\rho}} \right]^2 \quad (3.66b)$$

$$= \left[\text{Tr} \sqrt{\sqrt{\rho} \mathcal{U}^\dagger(\mathcal{E}(\rho)) \sqrt{\rho}} \right]^2 \quad (3.66c)$$

$$= F_{\mathcal{U}^\dagger \mathcal{E} \mathcal{I}}(\rho) \quad (3.66d)$$

where \mathcal{I} is the identity channel and $\mathcal{U}^\dagger(\rho) = U^\dagger \rho U$, is the adjoint channel of the unitary channel \mathcal{U} . Thus without loss of generality we may consider the gate fidelity $F_{\mathcal{E}}(\rho) \equiv F_{\mathcal{E},\mathcal{I}}(\rho)$ comparing \mathcal{E} to the identity channel, where we simply define $\mathcal{E} \equiv \mathcal{U}^\dagger \mathcal{F}$ if we wish to compare \mathcal{F} to a target unitary channel \mathcal{U} .

We note that gate fidelity is equivalent to entanglement fidelity for pure states. This can be shown graphically as follows

$$\begin{aligned}
 F_e(\mathcal{E}, |\psi\rangle\langle\psi|) &= \text{Diagram 1} \\
 &= \text{Diagram 2} \\
 &= F_{\mathcal{E}}(|\psi\rangle\langle\psi|)
 \end{aligned} \quad (3.67)$$

Alternatively we can also define the average gate fidelity in terms of the entanglement fidelity with the identity operator

$$\overline{F}_{\mathcal{E}} = \frac{d + F_e(\mathcal{E}, \mathbb{1})}{d(d+1)}. \quad (3.68)$$

The most often used quantity derived from the gate fidelity is the *average gate fidelity* taken by averaging $F_{\mathcal{E}}(\rho)$ over the the Fubini-Study measure. Due to the concavity of quantum states we need only integrate over pure states, which satisfy

$$F_{\mathcal{E}}(|\psi\rangle\langle\psi|) = \langle\psi|\mathcal{E}(|\psi\rangle\langle\psi|)|\psi\rangle. \quad (3.69)$$

Hence, the average gate fidelity is defined by

$$\overline{F}_{\mathcal{E}} = \int d\psi \langle\psi|\mathcal{E}(|\psi\rangle\langle\psi|)|\psi\rangle. \quad (3.70)$$

Average gate fidelity is a widely used figure of merit in part because it is simple to compute. The expression in Eq. (3.70) reduces to explicit expression for $\overline{F}_{\mathcal{E}}$ in terms of a single parameter of the channel \mathcal{E} itself. This has previously been given in terms of the Kraus representation [HHH99, Nie02], superoperator [EAŽ05] and Choi-matrix in [JK11]. We now present an equivalent graphical derivation of the average gate fidelity in terms of the Choi-matrix which we believe is simpler than previous derivations. We start with the tensor network diagram corresponding to Eq. (3.70) and perform graphical manipulations as follows

$$\begin{aligned} \overline{F}_{\mathcal{E}} &= \int d\psi \left\langle \psi \left| \mathcal{E}(|\psi\rangle\langle\psi|) \right| \psi \right\rangle \\ &= \int d\psi \left\langle \psi \left| \Lambda \right| \psi \right\rangle \\ &= \int d\psi \left\langle \psi \left| \Lambda \right| \psi \right\rangle \end{aligned} \quad (3.71)$$

For the next step of the proof we use the result that the average over ψ of a tensor product of states $|\psi\rangle\langle\psi|^{\otimes n}$ is given by

$$\int d\psi |\psi\rangle\langle\psi|^{\otimes n} = \frac{\Pi_{\text{sym}}(n, d)}{\text{Tr}[\Pi_{\text{sym}}(n, d)]} \quad (3.72)$$

where $\Pi_{\text{sym}}(n, d)$ is the projector onto the symmetric subspace of $\mathcal{X}^{\otimes n}$. This projector may be written as [MBKE11]

$$\Pi_{\text{sym}}(n, d) = \frac{1}{n!} \sum_{\sigma} P_{\sigma} \quad (3.73)$$

where P_{σ} are operators for the permutation σ of n -indices. These permutations may be represented as a swap type operator with n tensor wires. For the case of $n = 2$ we have the tensor diagram:

$$\boxed{\Pi_{\text{sym}}(2, d)} = \diamond_{\frac{1}{2}} \left(\text{---} + \text{---} \right) \quad (3.74)$$

Here we can see that

$$\text{Tr}[\Pi_{\text{sym}}(2, d)] = (d^2 + d)/2, \quad (3.75)$$

and hence we have that

$$\Pi_{\text{sym}}(2, d) = \frac{1}{2} (\mathbb{1} \otimes \mathbb{1} + \text{SWAP}) \quad (3.76)$$

$$\text{Tr}[\Pi_{\text{sym}}(2, d)] = \frac{d^2 + d}{2} \quad (3.77)$$

$$\Rightarrow \int d\psi |\psi\rangle\langle\psi|^2 = \frac{\mathbb{1} \otimes \mathbb{1} + \text{SWAP}}{d(d+1)} \quad (3.78)$$

where $\mathcal{X} \cong \mathbb{C}^d$, $\mathbb{1} \in L(\mathcal{X})$ is the identity operator, and SWAP is the SWAP operation on

$\mathcal{X} \otimes \mathcal{X}$. Substituting Eq. (3.74) into Eq. (3.71) completes the proof:

$$\begin{aligned}
\overline{F}_\mathcal{E} &= \int d\psi \text{ (diagram: a box } \Lambda \text{ with two input wires from the left, each labeled } \psi, \text{ and two output wires to the right, all connected by a single line)} \\
&= \frac{1}{d(d+1)} \left(\text{diagram: } \Lambda \text{ with two loops on the left} + \text{diagram: } \Lambda \text{ with two loops on the right} \right) \\
&= \frac{1}{d(d+1)} \left(\text{diagram: } \Lambda \text{ with two loops on the top} + \text{diagram: } \Lambda \text{ with two loops on the bottom} \right)
\end{aligned} \tag{3.79}$$

Using the fact that the Choi-matrix is normalized of a CPTP map is normalized such that $\text{Tr}[\Lambda] = d$, we have that the average gate fidelity in terms of the Choi-matrix is given by

$$\overline{F}_\mathcal{E} = \frac{d + \langle\langle \mathbb{1} | \Lambda | \mathbb{1} \rangle\rangle}{d(d+1)}. \tag{3.80}$$

From this proof one may derive expressions for the other representations using the channel transformations in Section 2.4, as was done in Section 3.2.2 for entanglement fidelity and then using Eq. (3.68). The resulting expressions are

$$\overline{F}_\mathcal{E} = \frac{d + \text{Tr}[\mathcal{S}]}{d(d+1)} \tag{3.81}$$

$$= \frac{d + \langle\langle \mathbb{1} | \Lambda | \mathbb{1} \rangle\rangle}{d(d+1)} \tag{3.82}$$

$$= \frac{d + \sum_j |\text{Tr}[K_j]|^2}{d(d+1)} \tag{3.83}$$

$$= \frac{d + d\chi_{00}}{d(d+1)} \tag{3.84}$$

$$= \frac{d + \text{Tr}_\mathcal{X}[A^\dagger] \cdot \text{Tr}_\mathcal{X}[A]}{d(d+1)} \tag{3.85}$$

where \mathcal{S} , Λ , $\{K_j\}$, χ , A are the superoperator, Choi-matrix, Kraus, χ -matrix and Strine-spring representations for \mathcal{E} respectively. In the case of the χ -matrix representation, χ is defined with respect to a basis $\{\sigma_j\}$ satisfying $\text{Tr}[\sigma_j] = \sqrt{d}\delta_{j,0}$.

Similar techniques can be applied for tensor networks that may be graphically manipulated into containing a term $\int d\psi |\psi\rangle\langle\psi|^{\otimes n}$ for $n > 2$. This could prove useful for computing higher order moments of fidelity functions and other quantities defined in terms of averages over quantum states $|\psi\rangle$. In this case there are $n!$ permutations of the tensor wires for the permutation operator P_σ in Eq. (3.73), and these can be decomposed as a series of SWAP gates. For example, in the case of $n = 3$ we have

$$\begin{aligned} \Pi_{\text{sym}}(3, d) = \frac{1}{6} & (\mathbb{1}^{\otimes 3} + \text{SWAP}_{1:2} + \text{SWAP}_{1:3} + \text{SWAP}_{2:3} \\ & + \text{SWAP}_{1:2}\text{SWAP}_{2:3} + \text{SWAP}_{2:3}\text{SWAP}_{1:2}) \end{aligned} \quad (3.86)$$

$$\text{Tr}[\Pi_{\text{sym}}(3, d)] = \frac{d^3 + 3d^2 + 2d}{6}. \quad (3.87)$$

3.3 Partial Characterization of Quantum Channels

One of the reasons that average gate fidelity as discussed in Section 3.2.3 is a widely used figure of merit for comparing a noisy quantum channel to an ideal channel is because it can be efficiently estimated using an experimental procedure known as *twirling* [EAŽ05, GAE07], and its robust generalization *randomized benchmarking* [MGE11, MGJ⁺12]. These are partial characterization protocols which, rather than trying to obtain a complete description of an unknown quantum channel, simply return a single parameter which is proportional the average gate fidelity. In this section we briefly describe these protocols without any regard to error analysis and performance which can be readily found in the literature [WF14, GFC15].

3.3.1 Unitary Designs for Average Gate Fidelity Estimation

Twirling and randomized benchmarking can be defined in terms of a unitary 2-design. A general unitary t -design for an operator Hilbert space $L(\mathcal{X})$ is a finite set of unitary operators

$$\mathbf{U}_t = \{U_1, \dots, U_{|\mathbf{U}|} : U_j \in L(\mathcal{X})\} \quad (3.88)$$

which satisfy

$$\frac{1}{|\mathbf{U}_t|} \sum_{U \in \mathbf{U}_t} \bar{U}^{\otimes t} \otimes U^{\otimes t} = \int d\eta(U) \bar{U}^{\otimes t} \otimes U^{\otimes t}. \quad (3.89)$$

where $\eta(U)$ is the uniform Haar measure. We define the following notation for a unitary- t design

$$\mathcal{S}_{\mathbf{U}_t} \equiv \frac{1}{|\mathbf{U}_t|} \sum_{U \in \mathbf{U}_t} \bar{U}^{\otimes t} \otimes U^{\otimes t}. \quad (3.90)$$

noting that it has the same form as a superoperator

$$\mathcal{S}_{\mathbf{U}_t} |A\rangle\rangle = \frac{1}{|\mathbf{U}_t|} \sum_{U \in \mathbf{U}_t} U^{\otimes t} A (U^{\otimes t})^\dagger. \quad (3.91)$$

Intuitively a set \mathbf{U}_t being a t -design means that summing over all elements of \mathbf{U}_t is equivalent to integrating over the uniform Haar measure $\eta(U)$. This is particularly relevant for quantum information experiments, as it means that protocols which involve a uniform average over the Haar measure can in principle be performed by summing over a finite set of unitary operators. In quantum information theory the most used cases are the 1-design, where \mathbf{U}_1 is the *generalized Pauli group*, and the 2-design, where $\mathbf{U}_2 = \mathbf{C}$ is the *Clifford Group*. Both of these groups have many nice properties for quantum information theory with an important one being that both approximate and exact unitary 2-design can be efficiently implemented [DCEL09, CLLW15]. By explicitly computing the Haar integral we can evaluate the form of the $\mathcal{S}_{\mathbf{U}_t}$ for a t -design [CS06]. The expressions for $t = 1$ and $t = 2$ are well known and given by [RS09]

$$\mathcal{S}_{\mathbf{U}_1} = \frac{1}{d} |\mathbb{1}\rangle\rangle\langle\langle\mathbb{1}| = \frac{1}{d} \begin{array}{c} \text{)} \\ \text{)} \\ \text{)} \end{array} \begin{array}{c} \text{(} \\ \text{(} \\ \text{(} \end{array} \quad (3.92)$$

$$\mathcal{S}_{\mathbf{U}_2} = \frac{1}{d^2 - 1} \left(\sum_{j=0}^1 |s_j\rangle\rangle\langle\langle s_j| \right) - \frac{1}{d(d^2 - 1)} \left(\sum_{j=0}^1 |s_j\rangle\rangle\langle\langle s_{j+2}| \right) \quad (3.93)$$

$$= \frac{1}{d^2 - 1} \left(\begin{array}{c} \text{)} \\ \text{)} \\ \text{)} \end{array} \begin{array}{c} \text{(} \\ \text{(} \\ \text{(} \end{array} + \begin{array}{c} \text{)} \\ \text{)} \\ \text{)} \end{array} \begin{array}{c} \text{(} \\ \text{(} \\ \text{(} \end{array} \right) - \frac{1}{d(d^2 - 1)} \left(\begin{array}{c} \text{)} \\ \text{)} \\ \text{)} \end{array} \begin{array}{c} \text{(} \\ \text{(} \\ \text{(} \end{array} + \begin{array}{c} \text{)} \\ \text{)} \\ \text{)} \end{array} \begin{array}{c} \text{(} \\ \text{(} \\ \text{(} \end{array} \right) \quad (3.94)$$

where $s_0 = \mathbb{1} \otimes \mathbb{1} = \begin{array}{c} \text{---} \\ \text{---} \end{array}$, $s_1 = \sum_{i,j=1}^d |i\rangle\langle j| \otimes |j\rangle\langle i| = \begin{array}{c} \text{X} \\ \text{X} \end{array}$.

We now show how average gate fidelity can be estimated by a unitary 2-design, which is a protocol sometimes referred to as *twirling* a channel. It involves averaging a channel

$\mathcal{E} \in T(\mathcal{X})$ over the uniform Haar measure of unitary channels. Starting from [Eq. \(3.70\)](#) we may rewrite the expression for average gate fidelity as

$$\bar{F}_{\mathcal{E}} = \int dU \langle \psi_0 | U \mathcal{E}(U^\dagger | \psi_0 \rangle \langle \psi_0 | U) U^\dagger | \psi_0 \rangle \quad (3.95)$$

for some arbitrary state $|\psi_0\rangle$. In terms of superoperators this is given by

$$\bar{F}_{\mathcal{E}} = \int dU \langle \langle \rho | \mathcal{S}_U \mathcal{S}_{\mathcal{E}} \mathcal{S}_U^\dagger | \rho \rangle \rangle = \langle \langle \rho | \mathcal{S}_{\mathcal{W}(\mathcal{E})} | \rho \rangle \rangle \quad (3.96)$$

where we have generalized [Eq. \(3.95\)](#) to allow for a mixed state input and output, and defined a *twirling channel* \mathcal{W} defined by

$$\mathcal{S}_{\mathcal{W}(\mathcal{E})} \equiv \int dU \mathcal{S}_U \mathcal{S}_{\mathcal{E}} \mathcal{S}_U^\dagger. \quad (3.97)$$

For a TP channel \mathcal{E} the action of this twirling operation is given by

$$\mathcal{W}(\mathcal{E}) = \lambda \mathcal{I} + (1 - \lambda) \mathcal{D} \quad (3.98)$$

where \mathcal{I} is the identity channel, \mathcal{D} is the completely depolarizing channel, and

$$\lambda = \frac{\text{Tr}[\mathcal{S}_{\mathcal{E}}] - 1}{d^2 - 1} = \frac{d\bar{F}_{\mathcal{E}} - 1}{d - 1}. \quad (3.99)$$

This follows as a result of [Proposition 4.5.1](#) in [Section 4.5](#). Hence for an arbitrary state ρ ,

$$\text{Tr} [\rho \mathcal{W}(\mathcal{E})(\rho)] = \lambda + \frac{1 - \lambda}{d} \quad (3.100)$$

$$= \frac{(d - 1)}{d} \lambda + \frac{1}{d} \quad (3.101)$$

$$= \left(\frac{d\bar{F}_{\mathcal{E}} - 1}{d - 1} \right) \left(\frac{d - 1}{d} \right) + \frac{1}{d} \quad (3.102)$$

$$= \bar{F}_{\mathcal{E}}. \quad (3.103)$$

3.3.2 Randomized Benchmarking

Using twirling to estimate average gate fidelity assumes one can accurately prepare and measure a fixed initial state ρ , which makes it sensitive to state preparation and measurement (SPAM) errors. In addition one also needs to consider any errors in implementing

the actual gates of the 2-design. Randomized benchmarking is a generalization of twirling which has built in robustness to SPAM errors, and allows for an estimating of gate errors in implementing the 2-design. In this section we briefly show how, in the exact case one can consider RB as a sequence of twirling operations.

Consider a sequence of unitary operators $U_1 \in L(\mathcal{X})$, and a recovery unitary operator $\mathcal{R} \in L(\mathcal{X})$ given by $R = U_k^\dagger \dots U_1^\dagger$. Randomized benchmarking of k copies of a channel \mathcal{E} is given by

$$RB_k(\mathcal{E}) = \frac{1}{|\mathbf{U}_2|^k} \sum_{U_1, \dots, U_k \in \mathbf{U}_2} \mathcal{R} \mathcal{E} U_k \mathcal{E} U_{k-1} \dots \mathcal{E} U_2 \mathcal{E} U_1 \quad (3.104)$$

$$\mathcal{S}_{RB_k(\mathcal{E})} = \frac{1}{|\mathbf{U}_2|^k} \sum_{U_1, \dots, U_k \in \mathbf{U}_2} \mathcal{S}_{\mathcal{R}} \mathcal{S}_{\mathcal{E}} \mathcal{S}_{U_k} \mathcal{S}_{\mathcal{E}} \mathcal{S}_{U_{k-1}} \dots \mathcal{S}_{\mathcal{E}} \mathcal{S}_{U_2} \mathcal{S}_{\mathcal{E}} \mathcal{S}_{U_1} \quad (3.105)$$

To see the action of the ideal case we define unitaries

$$V_1 = U_1^\dagger \quad (3.106)$$

$$V_2 = U_2^\dagger V_1 = U_2^\dagger U_1^\dagger \quad (3.107)$$

$$V_3 = U_3^\dagger V_2 = U_3^\dagger U_2^\dagger U_1^\dagger \quad (3.108)$$

$$\vdots \quad (3.109)$$

$$V_k = \prod_{j=1}^k U_j^\dagger = U_k^\dagger \dots U_1^\dagger \quad (3.110)$$

Then we may rewrite

$$RB_k(\mathcal{E}) = \frac{1}{|\mathbf{U}_2|^k} \sum_{V_1, \dots, V_k} (\mathcal{V}_k \mathcal{E} \mathcal{V}_k^\dagger) (\mathcal{V}_{k-1} \mathcal{E} \mathcal{V}_{k-1}^\dagger) \dots (\mathcal{V}_1 \mathcal{E} \mathcal{V}_1^\dagger) = \mathcal{W}(\mathcal{E})^k \quad (3.111)$$

Using the expression for the twirled channel in Eq. (3.98) we have

$$\mathcal{W}(\mathcal{E})^k = \lambda^k \mathcal{I} + (1 - \lambda^k) \mathcal{D} \quad (3.112)$$

where \mathcal{I} is the identity channel, \mathcal{D} is the completely depolarizing channel, and

$$\lambda = \frac{\text{Tr}[S_{\mathcal{E}}] - 1}{d^2 - 1} = \frac{d\overline{F}_{\mathcal{E}} - 1}{d - 1}. \quad (3.113)$$

In the ideal case of preparing and measuring a state ρ we have that the probability of success of the measurement is given by

$$p_k = \frac{1}{d} + \left(\frac{d-1}{d}\right) \lambda^k \quad (3.114)$$

so by measuring p_k for many values of k we may fit to an exponential decay to estimate λ , and then in turn estimate $\overline{F}_{\mathcal{E}}$.

As previously mentioned, one benefit of RB over twirling is that it works even in the presence of SPAM errors. To see this suppose instead of preparing and measuring ρ , we actually prepare a state ρ_1 and perform a measurement of whether the state is ρ_2 or not. In this case we have

$$p_k = \frac{1}{d} + A_0 \lambda^k \quad (3.115)$$

for some parameter unknown $A_0 = \text{Tr}[\rho_2 \rho_1] - \frac{1}{d}$. Further, if we don't implement a perfect twirl $\mathcal{W}(\mathcal{E})$ then we will not in general have perfect identity and depolarizing channels. Hence for RB the zeroth order approximation is to fit the data to a model

$$p_k = A_0 \lambda^k + B_0 \quad (3.116)$$

where A_0 and B_0 are unknown parameters.

3.4 Quantum State Tomography

Quantum state tomography is an experimental procedure to reconstruct an unknown density operator for an open quantum system by repeated measurements. Let $\mathcal{X} \cong \mathbb{C}^d$ be the pure state Hilbert space for a d -dimensional quantum system, and let $\mathbf{P} = \{\Pi_j : j = 1, \dots, M\}$ be a set of rank-1 projectors corresponding to pure quantum states $\Pi_j = |\psi_j\rangle\langle\psi_j| \in D\mathcal{X}$. In quantum state tomography, an experimenter performs measurements of each Π_j on many copies of an unknown quantum state ρ , and attempts to reconstruct a description of ρ from the observed measurement counts. We can loosely classify the usefulness of the set \mathcal{P} for tomography in the following way

1. *Tomographically complete:* The set \mathbf{P} has $|\mathbf{P}| = d^2$ elements and spans $D(\mathcal{X})$.
2. *Tomographically over-complete:* The set \mathbf{P} has $|\mathbf{P}| > d^2$ elements and spans $D(\mathcal{X})$.

3. *Tomographically incomplete*: The set \mathbf{P} does not span $D(\mathcal{X})$.

In principle it is possible to completely reconstruct an unknown quantum state via measurements of a tomographically complete or over-complete set of measurements \mathbf{P} .

The first method of reconstruction we present is called *linear inversion*. If we perform a measurement of a projector $\Pi_j \in \mathbf{P}$ on a state ρ the probability of success of the measurement is given by Born's rule

$$p_j = \text{Tr}[\Pi_j \rho]. \quad (3.117)$$

If we wish to infer p_j from experimental data, the most naive approach given N_j identical copies of ρ is to measure Π_j on each of the copies. If we observe n_j positive outcomes then our naive estimate of p_j is

$$\hat{p}_j = \frac{n_j}{N_j}. \quad (3.118)$$

This is the essence of linear inversion.

To reconstruct an estimation of the state ρ from these observed probabilities is slightly more complicated as a set \mathbf{P} , even if tomographically complete, is not an orthonormal basis for $D(\mathcal{X})$. The standard way to reconstruction of the state ρ from the linear inversion of probabilities p_j is done using a *dual basis* [DMP00], and is generally only done for a tomographically complete set. The derivation we present however holds for any set of measurement projectors \mathbf{P} , tomographically complete and overcomplete ones will simply have some desirable properties.

The superoperator for the projector onto the subspace spanned by \mathbf{P} is given by

$$\mathcal{S}_{\mathbf{P}} = \sum_{j=1}^M |\Pi_j\rangle\rangle\langle\langle \Pi_j|. \quad (3.119)$$

We define a dual basis $\mathbf{D} = \{D_j : j = 1, \dots, M\}$ for \mathbf{P} by the construction.

$$|D_j\rangle\rangle = \mathcal{S}_{\mathbf{P}}^{-1} |\Pi_j\rangle\rangle \quad (3.120)$$

where $\mathcal{S}_{\mathbf{P}}^{-1}$ refers to the *Moore-Penrose pseudo inverse* of the square matrix $\mathcal{S}_{\mathbf{P}}$ [Pen56]. If the matrix is invertible (non-singular) this will be equal to the standard matrix inverse. Our estimate of ρ via linear inversion may then be reconstructed as

$$\hat{\rho}_L = \sum_j \frac{n_j}{N_j} D_j. \quad (3.121)$$

To prove this we note that

$$|\hat{\rho}_L\rangle\rangle = \sum_j \frac{n_j}{N_j} |D_j\rangle\rangle \quad (3.122a)$$

$$= \sum_j \hat{\rho}_j \mathcal{S}_{\mathbf{P}}^{-1} |\Pi_j\rangle\rangle \quad (3.122b)$$

$$= \sum_j \mathcal{S}_{\mathbf{P}}^{-1} |\Pi_j\rangle\rangle \langle\langle \Pi_j | \rho_{\mathbf{n}} \rangle\rangle \quad (3.122c)$$

$$= (\mathcal{S}_{\mathbf{P}}^{-1} \mathcal{S}_{\mathbf{P}}) |\rho_{\mathbf{n}}\rangle\rangle, \quad (3.122d)$$

where $\rho_{\mathbf{n}}$ is the operator that generated an observed sequence of counts $n_j = N_j \text{Tr}[\Pi_j \rho_{\mathbf{n}}]$. We see here that for an arbitrary state our linear inversion estimate will be equal to the state $\hat{\rho}$ consistent with the measured counts if and only if

$$\mathcal{S}_{\mathbf{P}}^{-1} \mathcal{S}_{\mathbf{P}} = \mathbb{1}, \quad (3.123)$$

which is only the case if the matrix $\mathcal{S}_{\mathbf{P}}$ is invertible. This is true for any tomographically complete or overcomplete set \mathbf{P} . If the set \mathbf{P} is tomographically incomplete the pseudo-inverse will be the identity matrix for states which may be completely specified by the support of the set

$$\rho = \sum_{j=1}^M \lambda_j \Pi_j. \quad (3.124)$$

Otherwise we will recover the projection onto the support of \mathbf{P} .

For states entirely outside of the support of \mathbf{P} the pseudo inverse will return a maximally mixed state on the subspace spanned by these states. That is to say if $\text{Tr}[\Pi_j \rho] = 0$ for all $\Pi_j \in \mathbf{P}$ then

$$\mathcal{S}_{\mathbf{P}}^{-1} \mathcal{S}_{\mathbf{P}} |\rho\rangle\rangle = \left| \frac{\mathbb{1}}{d} \right\rangle\rangle \quad (3.125)$$

If we have a state with partial support on \mathbf{P} ,

$$\rho = p \rho_{\Pi} + (1 - p) \rho_{\perp} \quad (3.126)$$

then (under perfect counting statistics) the linear inversions estimate would be given by

$$\hat{\rho}_L = p \rho_{\Pi} + (1 - p) \frac{\mathbb{1}}{d}. \quad (3.127)$$

Note that due to finite count statistics the operator $\rho_{\mathbf{n}}$ that generated an observed sequence of counts n_j may not necessarily be positive semidefinite nor have unit trace, and hence

need not be a valid density matrix. One can always renormalize the recovered operator to have the correct trace, but if one adds the additional constraint that $\rho \geq 0$ then we are doing a form of constrained MLE rather than linear inversion which we discuss in [Section 3.4.1](#), but first we consider unconstrained MLE in terms of least squares fitting.

3.4.1 Maximum Likelihood Estimation

We now show how to formulate linear inversion as a *maximum likelihood estimation* (MLE) problem. Let $P(n_j|\rho)$ be the probability of observing n_j counts (successes) from measuring Π_j on N_j copies of a state ρ . The probability of observing a vector $\mathbf{n} = \{n_1, n_2, \dots, n_M\}$ counts from $\mathbf{N} = \{N_1, \dots, N_M\}$ trails for the set \mathbf{P} is then the product of the individual distributions

$$P(\mathbf{n}|\rho) = \prod_{j=1}^M P(n_j|\rho). \quad (3.128)$$

In an experiment the initial state ρ is unknown and it is the measured counts \mathbf{n} that are known. We may rewrite [Eq. \(3.128\)](#) as a *likelihood function*

$$L(\rho|\mathbf{n}) = P(\mathbf{n}|\rho), \quad (3.129)$$

which quantifies the likelihood of the state ρ to have generated the observed set of counts. We may then maximize the likelihood function over ρ to find the state $\hat{\rho}_{\text{mle}}$ with the highest probability of generating the observed counts:

$$\hat{\rho}_{\text{mle}} = \arg \max_{\rho} L(\rho|\mathbf{n}). \quad (3.130)$$

This procedure called maximum likelihood estimation. In practice maximizing $\mathcal{L}(\rho)$ is often numerically unstable due to the product of exponentials for normally distributed variables as in [Eq. \(3.132\)](#). Instead, what is usually done is to minimize the negative log-likelihood function:

$$\hat{\rho}_{\text{mle}} = \arg \min_{\rho} \left[-\log L(\rho|\mathbf{n}) \right]. \quad (3.131)$$

The Moore-Penrose pseudo-inverse of a matrix is the solution to a least squares fitting optimization problem [[Pen56](#)]. We will now show how one may phrase quantum state tomography as a weighted least squares fitting, with linear inversion corresponding to

case of uniform weights for all measurements. In the limit of large N_j the probability distributions in [Eq. \(3.128\)](#) for n_j can be approximated by the normal distribution

$$P(n_j | \rho) = \frac{1}{\sqrt{2\pi\sigma_j^2}} \exp\left(-\frac{(N_j \text{Tr}[\Pi_j \rho] - n_j)^2}{2\sigma_j^2}\right). \quad (3.132)$$

For a normal distribution the negative log-likelihood function for a given set of counts \mathbf{n} is proportional to

$$\log L(\rho | \mathbf{n}) = \sum_{j=1}^M \log P(n_j | \rho) \quad (3.133a)$$

$$= \sum_{j=1}^M \left(-\frac{1}{2} \log [2\pi\sigma_j^2] - \frac{(N_j \text{Tr}[\Pi_j \rho] - n_j)^2}{\sigma_j^2} \right) \quad (3.133b)$$

$$= -C - \sum_{j=1}^M \frac{(N_j \text{Tr}[\Pi_j \rho] - n_j)^2}{\sigma_j^2}, \quad (3.133c)$$

where $C = \sum_{j=1}^M \log [2\pi\sigma_j^2] / 2$. If we make a further approximation that the variance σ_j does not depend on $\text{Tr}[\Pi_j \rho]$ and instead is the variance of the observed counts n_j , then C is constant and our MLE of ρ becomes

$$\hat{\rho}_{\text{mle}} = \arg \min_{\rho} \sum_{j=1}^M \frac{(N_j \text{Tr}[\Pi_j \rho] - n_j)^2}{\sigma_j^2}. \quad (3.134)$$

If we place no constraints on the operator ρ we may solve this minimization exactly. To show this we define the following operators and vectors

$$S = \sum_{j=1}^M |j\rangle\langle\Pi_j|, \quad |f\rangle = \sum_{j=1}^M \frac{n_j}{N_j} |j\rangle, \quad W = \sum_{j=1}^M \frac{N_j}{\sigma_j} |j\rangle\langle j|, \quad (3.135a)$$

where S is similar to a vectorization change of basis operator in [Section 2.2.2](#) (although \mathbf{P} is not an orthonormal basis), $|f\rangle$ is a column vector of the observed probabilities, and W is a diagonal matrix of the weights given by the number of samples divided by the variance

σ_j . Using this we may then rewrite Eq. (3.134) as

$$\hat{\rho}_{\text{mle}} = \arg \min_{\rho} \sum_{j=1}^N \frac{(N_j \text{Tr}[\Pi_j \rho] - n_j)^2}{\sigma_j^2} \quad (3.136a)$$

$$= \arg \min_{\rho} \sum_{j=1}^N \left(\langle\langle \rho | \Pi_j \rangle\rangle - \frac{n_j}{N_j} \right) \frac{N_j^2}{\sigma_j^2} \left(\langle\langle \Pi_j | \rho \rangle\rangle - \frac{n_j}{N_j} \right) \quad (3.136b)$$

$$= \arg \min_{\rho} \sum_{i,j,k,l=1}^N \left(\langle\langle \rho | \Pi_i \rangle\rangle \langle i | - \frac{n_i}{N_i} \langle i | \right) W^\dagger W \left(|l\rangle \langle\langle \Pi_l | \rho \rangle\rangle - \frac{n_l}{N_l} |l\rangle \right) \quad (3.136c)$$

$$= \arg \min_{\rho} \left(\langle\langle \rho | S^\dagger - \langle f | \rangle \right) W^\dagger W \left(| \rho \rangle\rangle - |f\rangle \right) \quad (3.136d)$$

$$= \arg \min_{\rho} \left\| W(|\rho\rangle\rangle - |f\rangle) \right\|_2^2. \quad (3.136e)$$

Since the norm is always positive we can drop the square root and our MLE estimate of ρ is given by the weighted least squares fitting problem

$$\hat{\rho}_{\text{mle}} = \arg \min_{\rho} \left\| W(|\rho\rangle\rangle - |f\rangle) \right\|. \quad (3.137)$$

If we place no constraints on ρ this has an analytic solution

$$|\hat{\rho}_{\text{mle}}\rangle\rangle = (S^\dagger W^\dagger W S)^{-1} S^\dagger W^\dagger W |f\rangle. \quad (3.138)$$

We further note that if we define an operator $\rho_{\mathbf{n}}$ which is the ideal operator to generate an observed set of counts \mathbf{n} via $n_j = N_j \text{Tr}[\Pi_j \rho_{\mathbf{n}}]$, then

$$\langle\langle \Pi_j | \rho_{\mathbf{n}} \rangle\rangle = \frac{n_j}{N_j} \implies |f\rangle = \sum_j |j\rangle \langle\langle \Pi_j | \rho' \rangle\rangle = S | \rho_{\mathbf{n}} \rangle\rangle, \quad (3.139)$$

and hence we may rewrite Eq. (3.138) as

$$|\hat{\rho}\rangle\rangle = (S^\dagger W^\dagger W S)^{-1} S^\dagger W^\dagger W S | \rho_{\mathbf{n}} \rangle\rangle. \quad (3.140)$$

The linear inversion reconstruction in Eq. (3.122d) is a special case of Eq. (3.140) where we set the variance to be $\sigma_j = N_j$ so that $W = \mathbb{1}$, and

$$|\hat{\rho}\rangle\rangle = (S^\dagger S)^{-1} S^\dagger S | \rho_{\mathbf{n}} \rangle\rangle = \mathcal{S}_{\mathbf{P}}^{-1} \mathcal{S}_{\mathbf{P}} | \rho_{\mathbf{n}} \rangle\rangle, \quad (3.141)$$

where $\mathcal{S}_{\mathbf{P}} = S^\dagger S$. Hence we may also do weighted linear inversion reconstruction of ρ by Eq. (3.138), which corresponds to defining

$$\mathcal{S}_{\mathbf{P}} = \sum_j w_j^2 |\Pi_j\rangle\rangle\langle\langle \Pi_j|, \quad (3.142)$$

with weights $w_j = N_j/\sigma_j$.

While least squares fitting is a special case of unconstrained MLE, in practice what is usually referred to as MLE quantum state tomography is the case of *constrained* MLE. Due to finite count statistics, detector dark counts, and other sources of experimental measurement error, for any real experiment data the ideal operator $\rho_{\mathbf{n}}$ which generates a sequence \mathbf{n} of observed counts will not actually be the true state ρ , and indeed need not even be a valid density matrix. If the true state is very close to a pure state, the operator $\hat{\rho}_L$ reconstructed from linear inversion will often have a small negative eigenvalues. Hence the standard usage of MLE-tomography refers to weighted linear inversion where we add an additional constraint that ρ must be positive semidefinite and normalized to trace 1. The cost of doing this is that there is no longer an analytic solution to the optimization problem in Eq. (3.137).

Formally we define MLE state tomography to be the following optimization problem

$$\begin{aligned} & \text{minimize} && \|W(S|\rho\rangle\rangle - |f\rangle\rangle)\|_2 \\ & \text{subject to:} && \rho \geq 0, \text{Tr}[\rho] = 1 \end{aligned}$$

Since the objective function being minimized is convex we may solve it using powerful the techniques of convex optimization. In particular since the constraint $\rho \geq 0$ is a semidefinite constant this optimization problem may be phrased as a *semidefinite program* (SDP), for which sophisticated algorithms exist to find a global optimum [VB96].

There has been a large body of literature on variations of the likelihood function to be minimized for MLE quantum state tomography. We briefly mention two here that serve opposite purposes. Hedged maximum likelihood estimation (HMLE) [BK10a] is the optimization problem

$$\begin{aligned} & \text{minimize} && \frac{1}{2} \|W(S|\rho\rangle\rangle - |f\rangle\rangle)\|_2^2 - \beta \log \det(\rho) \\ & \text{subject to:} && \rho \geq 0, \text{Tr}[\rho] = 1. \end{aligned}$$

Trace constrained maximum likelihood estimation (TCMLE) [FGLE12], also known as

compressed sensing or matrix-lasso, is the optimization problem

$$\begin{aligned} & \text{minimize} && \frac{1}{2} \|W(S|\rho\rangle\rangle - |f\rangle\rangle)\|_2^2 + \mu \|\rho\|_1 \\ & \text{subject to:} && \rho \geq 0, \text{Tr}[\rho] = 1 \end{aligned}$$

HMLE can be thought of as a skeptics fit which penalizes low rank fits to the data, while in contrast, TCMLE can be thought of as an optimists fit and favours low rank fits to the data. Both of these are still convex optimization problems with semidefinite constraints, and so can still be solved numerically as SDPs. The parameters β and μ quantify the degree to which we penalize low rank or high rank fits in hedged and trace constrained MLE respectively. Note that in both of these definitions we minimize $\frac{1}{2} \|W(S|\rho\rangle\rangle - |f\rangle\rangle)\|_2^2$ instead of $\|W(S|\rho\rangle\rangle - |f\rangle\rangle)\|_2$, as once we add an additional cost term the specific value of the expression is relevant when minimizing the objective function, not just the value of ρ at the minimum.

3.4.2 Tomography with Two Outcome Measurements

In many experimental situations our measurements will be two two outcomes, given by a POVM $\{\Pi_j, \mathbb{1} - \Pi_j\}$, where we either measure the unknown quantum state to be in Π_j or we don't. Let $\mathbb{P}(n_j|\rho)$ be the probability of obtaining exactly n_j successful measurement outcomes from from POVM $\{\Pi_j, \mathbb{1} - \Pi_j\}$ on N_j copies of ρ . For a two outcome POVM this is given by a binomial distribution

$$\mathbb{P}(n_j|\rho) = \binom{N_j}{n_j} p_j^{n_j} (1 - p_j)^{N_j - n_j} \quad (3.143)$$

where $p_j = \text{Tr}[\Pi_j\rho]$ is the probability of success. To make this more computationally tractable it is general practice to approximate the binomial distribution by a normal distribution with mean $\mu_j = N_j p_j$ and standard deviation $\sigma_j^2 = N_j p_j (1 - p_j)$. With this approximation the MLE objective function in Eq. (3.134) function has a particular simple form

$$\hat{\rho}_{\text{mle}} = \arg \min_{\rho} (-\ln \mathcal{L}(\rho|\{m_j\})) \quad (3.144)$$

$$= \arg \min_{\rho} \sum_j N_j \left(\frac{(f_j - p_j)^2}{p_j(1 - p_j)} \right) \quad (3.145)$$

where $f_j = n_j/N_j$ are the probabilities obtained by the observed number of counts n_j . To further simplify the problem one may assume that the observed f_j are close to the true p_j , which is valid for large N_j . Under this approximation we may take a second order Taylor series expansion in p_j about f_j to obtain

$$\hat{\rho} \approx \arg \min_{\rho} \sum_j \frac{N_j(\text{Tr}[\Pi_j^\dagger \rho] - f_j)^2}{f_j(1 - f_j)} \quad (3.146)$$

In effect this is assuming that the standard deviation of each measurement $\sigma_j^2 = p_j(1 - p_j)$ is approximately equal to the observed standard deviation $f_j(1 - f_j)$.

We can see that the above reconstruction is ill defined if the observed probabilities f_j are either 0 or 1. If we trust our estimate completely either of these two values would completely determine the state $\rho = \Pi_j$. If we allow that this may not be the true value due to finite statistics we can circumvent this issue by adding a small amount of noise to our data. One way to do this is to define our observed probabilities to be [\[FBK12\]](#)

$$f_j = \frac{n_j + \beta}{N_j + 2\beta}. \quad (3.147)$$

rather than n_j/N_j , where β is a real number that incorporates our uncertainty. In practice $\beta \approx 1/2$ is useful for many situations, though if we are confident that the true state is pure and our number of samples N_j is large it may be better to set it to a smaller value.

Another point to note is that in many experimental situations the total number of copies of a state being measured is not known in advance, only the resulting measurement count statistics are known. In this case, if for each two outcome measurement of an operator Π_j the failure corresponds to a measurement of another rank-1 projector, $Q_j = \mathbb{1} - \Pi_j$, one can estimate the total number of counts of counts by summing the observed counts for Π_j and Q_j . For example if for a qubit one measures the Pauli operators X, Y, Z , each operator corresponds to a two outcome measurement $\{\Pi_{\alpha+}, \Pi_{\alpha-}\}$ where $\alpha = x, y, z$, where $\Pi_{\alpha\pm}$ are the \pm eigenstates of the operator. Let the number of outcomes for measurement of the $+1$ and -1 eigenstates of Z be n_{z+} and n_{z-} respectively. The total number of counts for measurements of the two Z projectors are $N_{z+} = N_{z-} = n_{z+} + n_{z-}$, and similarly for X and Y . Hence we may phrase this as state tomography with the tomographically overcomplete set with six elements

$$\mathcal{P}_\sigma = \{\Pi_{x+}, \Pi_{x-}, \Pi_{y+}, \Pi_{y-}, \Pi_{z+}, \Pi_{z-}\} \quad (3.148)$$

with observed counts $\mathbf{n} = \{n_{x+}, n_{x-}, n_{y+}, n_{y-}, n_{z+}, n_{z-}\}$, and total numbers of copies per measurement of $N_{\alpha\pm} = n_{\alpha-} + n_{\alpha+}$.

3.5 Quantum Process Tomography

Quantum process tomography is a procedure for reconstructing the description of an arbitrary quantum channel $\mathcal{E} \in T(\mathcal{X}, \mathcal{Y})$ through measurement statistics. The full characterization of all parameters of an unknown channel \mathcal{E} requires a tomographically complete input states $\mathbf{Q} = \{\rho_j\}$ for $D(\mathcal{X})$, and a tomographically complete set of measurements $\mathbf{P} = \{\Pi_j\}$ for $D(\mathcal{Y})$. Many copies of each input state $\rho_j \in \mathbf{Q}$ are then sent through the unknown channel \mathcal{E} , and state tomography using \mathbf{P} is performed on the output $\mathcal{E}(\rho_j)$. This procedure is often referred to as *standard quantum process tomography* (SQPT) [DLP01] as other variants exist.

Quantum process tomography can be succinctly phrased as a special case of quantum state tomography, where the unknown state is the Choi-matrix of the channel. This is because, as was described in Section 2.3.4, a Choi-matrix $\Lambda_{\mathcal{E}}$ for a CP-map \mathcal{E} is isomorphic to a bipartite quantum state $\rho_{\mathcal{E}} \in L(\mathcal{X} \otimes \mathcal{Y})$ where $\rho_{\mathcal{E}} = \Lambda/d_x$. We will now detail how one may implement quantum process tomography in this way.

Let $\mathcal{E} \in C(\mathcal{X}, \mathcal{Y})$ be a channel with Choi-matrix $\Lambda_{\mathcal{E}}$. Let $\mathbf{Q} = \{\rho_1, \dots, \rho_K\}$ be a set of states on $D(\mathcal{X})$, and let $\mathbf{P} = \{\Pi_1, \dots, \Pi_M\}$ be a set of rank one projectors acting on $D(\mathcal{Y})$. Recall from Eq. (2.92) that the evolution of a state ρ is given by

$$\mathcal{E}(\rho_i) = \text{Tr}_{\mathcal{X}}[(\rho_i^T \otimes \mathbb{1})\Lambda] \quad (3.149)$$

Hence if we measure a projector $\Pi_j \in \mathbf{P}$, the probability of success is given by

$$p_{ij} = \text{Tr}[\Pi_j^\dagger \mathcal{E}(\rho_i)] \quad (3.150a)$$

$$= \text{Tr}[\Pi_j^\dagger \text{Tr}_1[(\rho_i^T \otimes \mathbb{1})\Lambda]] \quad (3.150b)$$

$$= \text{Tr}[(\rho_i^T \otimes \Pi_j^\dagger)\Lambda] \quad (3.150c)$$

$$= \langle\langle \bar{\rho}_i \otimes \Pi_j | \Lambda \rangle\rangle \quad (3.150d)$$

This can be interpreted as a measurement of $\Lambda_{\mathcal{E}}$ with measurement operator

$$\Pi_{ij} = \bar{\rho}_i \otimes \Pi_j. \quad (3.151)$$

Hence we may define a set of measurement operators on $L(\mathcal{X} \otimes \mathcal{Y})$ by

$$\mathbf{P}_{\mathcal{E}} = \{\Pi_{ij} : i = 1, \dots, K, j = 1, \dots, M\}, \quad \Pi_{ij} \equiv \bar{\rho}_i \otimes \Pi_j \quad (3.152)$$

Let N_{ij} the number of copies of Λ measured with Π_{ij} , or equivalently the number of copies of $\mathcal{E}(\rho_i)$ measured with Π_j , and n_{ij} the number of successes for these measurements. We

can then implement process tomography by defining the operators

$$f_{ij} = \frac{n_{ij} + \beta}{N_{ij} + 2\beta} \quad |f\rangle = \sum_{ij} f_{i,j} |i, j\rangle \quad (3.153a)$$

$$S = \sum_{ij} |i, j\rangle \langle\langle \Pi_{ij} | \quad W = \sum_{ij} w_{ij} |i, j\rangle \langle i, j|, \quad (3.153b)$$

where for a binomial distributed measurement outcome we have

$$w_{ij} = \sqrt{\frac{N_{ij}}{f_{ij}(1 - f_{ij})}}, \quad (3.154)$$

and we have incorporated the hedging of the frequencies f_{ij} as described in [Section 3.4.2](#). We then reconstruct the Choi-matrix by either linear inversion:

$$|\hat{\Lambda}\rangle\rangle = (S^\dagger W^\dagger W S)^{-1} S^\dagger W^\dagger W |f\rangle \quad (3.155)$$

Or the constrained MLE:

$$\begin{aligned} & \text{minimize } \|W(S|\Lambda\rangle\rangle - |f\rangle)\|_2 \\ & \text{subject to: } \Lambda \geq 0, \text{ Tr}[\Lambda] = d. \end{aligned} \quad (3.156)$$

Similarly one may also implement the other variants of MLE mentioned in [Section 3.4.1](#).

3.5.1 Ancilla Assisted Process Tomography

An alternative procedure for perform quantum process tomography is for an experimenter to perform state tomography on the Choi-matrix directly by using a method known as *ancilla assisted process tomography* (AAPT) [[Leu03](#), [DLP03](#), [ABJ⁺03](#)]. The simplest case of AAPT is *entanglement assisted process tomography* (EAPT) which is an experimental realization of the Choi-Jamiołkowski isomorphism from [Eq. \(2.87\)](#).

Let $\mathcal{E} \in C(\mathcal{X}, \mathcal{Y})$ be an unknown channel to characterized, and suppose an experimenter has access to an ancilla system $\mathcal{Z} \cong \mathcal{X}$. To impliment EAPT an experimenter prepares a a maximally entangled input state $\rho_\Phi \in D(\mathcal{Z} \otimes \mathcal{X})$

$$\rho_\Phi = \frac{1}{d} |\mathbb{1}\rangle\rangle \langle\langle \mathbb{1}| \quad (3.157)$$

across the system of interest \mathcal{X} and the ancilla \mathcal{Z} , and then subjects the system part of the state ρ_Φ to the unknown channel \mathcal{E} , and the ancilla part of ρ_Φ to an identity channel \mathcal{I} . The output of this joint system-ancilla channel is the rescaled Choi-matrix:

$$\rho'_\phi = (\mathcal{I} \otimes \mathcal{E})(\rho_\Phi) = \frac{\Lambda}{d}. \quad (3.158)$$

where $d = \dim(\mathcal{X})$. Hence the experimenter may perform EAPT by performing state tomography on the output state $\rho'_\phi \in D(\mathcal{Z} \otimes \mathcal{Y})$.

General AAPT may also be implemented with an input state $\rho_A \in D(\mathcal{Z} \otimes \mathcal{X})$ that is not the maximally entangled state. Indeed, AAPT has been demonstrated experimentally with a state which does not have any entanglement at all, though this comes with the cost of an increase in the estimation error of the reconstructed channel [ABJ⁺03]. A necessary and sufficient condition for a general state ρ_A to allow complete recovery of the Choi-matrix of an unknown channel via AAPT is that it have a *Schmidt number* equal to d^2 where d is the dimension of the state space \mathcal{X} [ABJ⁺03]. This conditions has previously been called *faithfulness* of the input state, and one can recover the original Choi-matrix for the unknown channel \mathcal{E} by applying an appropriate inverse map to the output state in post-processing [DLP03]. We now provide an arguably simpler derivation of this condition, and the explicit construction of the inverse recovery operator, which makes use of the graphical techniques introduced in Chapter 2.

Proposition 3.5.1 (AAPT). *Let $\mathcal{E} \in C(\mathcal{X}, \mathcal{Y})$ be a channel, and consider a state $\rho_A \in D(\mathcal{Z} \otimes \mathcal{X})$ where $\mathcal{Z} \cong \mathcal{X}$ is an ancilla. Let $\rho'_A = (\mathcal{I} \otimes \mathcal{E})(\rho_A)$ where $\mathcal{I} \in C(\mathcal{Z})$ is the identity channel. If we define a superoperator $\mathcal{S}_A = \rho_A^R$ by reshuffling the input state, then*

1. *We may recover the Choi-matrix Λ for an arbitrary channel \mathcal{E} if and only if reshuffled density matrix \mathcal{S}_A is invertible.*
2. *If \mathcal{S}_A is invertible, we may recover the Choi-matrix via $\Lambda = (\mathcal{R} \otimes \mathcal{I})(\rho'_A)$ where \mathcal{R} is the recovery channel with superoperator $\mathcal{S}_\mathcal{R} = (\mathcal{S}_A^T)^{-1}$.*

Proof. The essence of this proof is that we can treat the bipartite state $\rho_A \in D(\mathcal{Z} \otimes \mathcal{X})$ as a Choi-matrix for an effective channel via the Choi-Jamiołkowski isomorphism (but with trace normalization of 1 instead of d), hence we can apply the reshuffling superoperator

from Eq. (2.207) to transform this state to a superoperator

(3.159a)

where $\mathcal{S}_A = \rho_A^R$. Next we rearrange the tensor diagram by sliding this super operator around the wires two swap the roles of state \mathcal{S}_A and the channel super operator \mathcal{S} :

(3.159b)

(3.159c)

Finally, we may reapply the reshuffling operator to convert \mathcal{S} to its Choi-matrix Λ

(3.159d)

Hence we see that we may completely recover Λ if and only \mathcal{S}_A^T is invertible. □

In particular, we note that if the initial state ρ_A is maximally entangled, then it can be expressed as $\rho_A = |V\rangle\langle V|$ for some unitary $V \in L(\mathcal{X})$. In this case the reshuffled superoperator of the state corresponds to a unitary channel $\mathcal{S}_A = \bar{V} \otimes V$, and hence is invertible with $\mathcal{S}_A^{-1} = \mathcal{S}_A^\dagger$. If the input state is not maximally entangled, then the closer it is to a singular matrix, the larger the condition number and hence the larger the amplification in error when inverting the matrix.

3.5.2 Robust Quantum Process Tomography

The twirling protocol described in [Section 3.3.1](#) can also be used to perform quantum process tomography, albeit with a tomographically under complete set. While in the ideal case there isn't any real advantage to implementing tomography this way, we illustrate it for pedagogical purposes as it provides a clear and intuitive justification for implementing a randomized benchmarking based tomographic protocol such as was recently described in [\[KdSR+14\]](#). Unlike the twirled case, the randomized benchmarking has distinction of being much more robust to state preparation and measurement errors, which may make standard tomographic protocols unreliable.

Suppose we may accurately prepare a single state $\rho \in D(\mathcal{X})$ with $\mathcal{X} \cong \mathbb{C}^d$, and further that we can perform a measurement of this state corresponding using POVM $M = \{\rho, \mathbb{1} - \rho\}$. Consider an unknown channel $\mathcal{E} \in C(\mathcal{X})$, a let \mathbf{C} be the set of Clifford channels.

The twirling procedure involves implementing the channel

$$\bar{F}_j = \frac{1}{|\mathbf{C}|} \sum_{\mathcal{C}_i \in \mathbf{C}} \text{Tr}[\rho \mathcal{C}_i \mathcal{C}_j^\dagger \mathcal{E} \mathcal{C}_i^\dagger(\rho)] \quad (3.160)$$

$$= \frac{d + \text{Tr}[\mathcal{S}_{\mathcal{C}_j}^\dagger \mathcal{S}_{\mathcal{E}}]}{d(d+1)}. \quad (3.161)$$

Now,

$$\text{Tr}[\mathcal{S}_{\mathcal{C}_j}^\dagger \mathcal{S}_{\mathcal{E}}] = \langle\langle \mathcal{S}_{\mathcal{C}_j} | \mathcal{S}_{\mathcal{E}} \rangle\rangle = \langle\langle \Lambda_{\mathcal{C}_j} | \Lambda_{\mathcal{E}} \rangle\rangle. \quad (3.162)$$

Hence the probability of measuring ρ for the twirled channel $\mathcal{C}_j^\dagger \mathcal{E}$ is given by

$$p_j = \bar{F}_j = \frac{1}{d+1} + \frac{\langle\langle \Lambda_{\mathcal{C}_j} | \Lambda_{\mathcal{E}} \rangle\rangle}{d(d+1)}. \quad (3.163)$$

Next, we define a measurement set of Clifford channels

$$\mathbf{P}_{\mathcal{C}} = \left\{ \Pi_j = \frac{\Lambda_{\mathcal{C}}}{d} : \mathcal{C} \in \mathbf{C} \right\}. \quad (3.164)$$

We define a number p'_j to be the ‘‘probability’’ of successful measuring Π_j for a state given by the Choi-matrix $\Lambda_{\mathcal{E}}$:

$$p'_j = \langle\langle \Pi_j | \Lambda_{\mathcal{E}} \rangle\rangle = (d+1)p_j - 1. \quad (3.165)$$

We may then reconstruct $\Lambda_{\mathcal{E}}$ via linear inversion or MLE fitting with the set of projectors $\mathbf{P}_{\mathcal{C}}$ and observed probabilities $\{p'_j\}$.

Generalizing twirled tomography to randomized benchmarking based tomography as described in [KdSR⁺14] follows the exact same procedure as used above, but with randomized benchmarking used to estimate the average gate fidelity. Once this has been estimated one can then compute the estimate of p'_j and proceed in the same manner as with the twirling reconstruction.

We now mention some caveats of this approach. For a d -dimensional system we need d^4 linearly independent rank-1 states Π_j to have a tomographically complete set. However, the only channels which have rank-1 Choi matrices are projective measurements and unitary operations. If we restrict ourselves to unitary operations then we cannot span the full Hilbert space and hence will have a tomographically undercomplete set. To show this, we note that the superoperator of a unitary is $S_U = \bar{U} \otimes U$. The subspace of $\mathcal{L}(\mathcal{X} \otimes \mathcal{X})$ spanned by these tensor-product superoperators has dimension $d^4 - 2d^2 + 2$ [RS09]. The span of Clifford channels $\{S_{\mathcal{C}} : \mathcal{C} \in \mathcal{C}(d)\}$ is however equal to the span of all unitary channels, and the support of this set is all unital CPTP maps. Hence it in principle allows one to reconstruct the unital part of any unknown CPTP map \mathcal{E} . In an actual experimental implementation there is a significant issue that could arise due to finite sampling statistics or other count errors. The mean \mathbf{E} and variance σ of the derived probabilities p'_j used from the reconstruction are given by

$$\mathbf{E}[p'_j] = (d + 1)\mathbf{E}[p_j] - 1 \tag{3.166}$$

$$\sigma[p'_j] = (d + 1)\sigma[p_j] \tag{3.167}$$

and so the minimum average number of counts observed in an ideal twirling experiment is $p_j = N_j/(1 + d)$. If our experimentally observed number of counts is below this threshold then the estimated mean of our reconstruction of p'_j will be negative. Hence we see that our reconstruction of p'_j does not correspond to a probability.

3.6 Summary

In this chapter we reviewed several common methods and measures for characterizing quantum channels and presented several new results including a new proof and condition for a state being suitable for AAPT, and tighter upper bounds on the CB trace norm, that are exact for certain classes of channels. In particular this bound is tight for the CB trace norm of the difference of two channels with Choi-matrices that are diagonal in the

same basis. The CB trace norm and these new bounds will be of use when we use it to characterize generalizations of quantum channels in [Chapters 4 and 5](#). In the average gate fidelity example we were able to give a shorter proof of the average gate fidelity of a quantum channel by using the graphical representation of the average over states $\int d\psi |\psi\rangle\langle\psi|^{\otimes n}$ in terms of the projector onto the symmetric subspace of $\mathcal{X}^{\otimes n}$. This projector can be expressed as the sum of $n!$ permutation operators which have a natural representation as tensor network diagrams consisting of a series of SWAP operations corresponding to all left-to-right permutations of n wires. The power of tensor network framework was to manipulate the string diagram for a given expression to form the tensor product $|\psi\rangle\langle\psi|^{\otimes n}$ irregardless of where the n copies of $|\psi\rangle\langle\psi|$ appear in the original expression. After substituting in the projector onto the symmetric subspace we can contract the $n!$ resulting diagrams to arrive at the final value. Similar techniques could prove useful for calculating other quantities such as higher order moments of fidelity functions and other quantities defined in terms of averages over quantum states.

Our presentation of quantum tomography showed that many forms of quantum process tomography can be interpreted as special cases of quantum state tomography where the quantum state being reconstructed is the unknown channels Choi-matrix. This formalism will be especially useful in [Chapter 4](#) where we introduce a new formalism for generalized quantum channels and show how tomography of these more general objects can again be stated as a form of quantum state tomography. We also demonstrated a use of graphical calculus for quantum channels which allowed us to derive a succinct expression for a necessary sufficient condition for a bipartite state to be usable for ancilla assisted process tomography. By vectorizing bipartite matrices and their bipartite matrix transformations we found that we could consider a bipartite density matrix over a system and ancilla as a type of Choi-matrix via the Choi-Jamilkowski isomorphism. By applying the reshuffling transformation we can convert this into an effective superoperator channel which must be invertible to recover the Choi-matrix. While this is equivalent to the perviously known result we believe it is a simpler proof, and the resulting recovery channel is simpler to construct using the presented method.

Our presentation of partial characterization via unitary 2-designs and randomized benchmarking will be useful in [Chapter 4](#) where we consider the twirling procedure as type of generalized channel and show how to generalize this to describe randomized benchmarking and twirling of multiple subsystems which are related to other proposed characterization techniques [[ESM⁺07](#), [GCM⁺12](#)]. While quantum process tomography and randomized benchmarking are two of the most widely used characterization techniques, there exist many proposals for alternative characterization techniques aiming to address various issues which arise with these protocols [[ESM⁺07](#), [MGS⁺13](#), [BK13](#)]. The main prob-

lem with quantum process tomography is that it requires a number of experiments scaling exponentially with the number of qubits which makes it impossible to implement for more than a couple of qubits, and is also highly susceptible to SPAM errors. The latter issue has been addressed to some extent in recent generalizations such as gate-set tomography and randomized benchmarking tomography [MGS⁺13, BK13, KdSR⁺14]. RB on the other hand is a scalable protocol that is robust to certain SPAM errors, however average gate fidelity is usually too coarse grained a parameter for honestly estimating fault-tolerance thresholds, or if one is interested in more general properties of the noise on a system. For example, sometimes a channel description is not sufficient and one needs to characterize the parameters of a Hamiltonian or dissipator which generates the resulting dynamics rather than the resulting dynamics themselves, for which there exist characterization schemes based on Bayesian parameter estimation methods [GFWC12, WGFC14].

Chapter 4

Quantum Superchannels

4.1 Introduction

In this chapter we introduce a generalization of quantum channels which we call *quantum superchannels* — effective quantum channels which take regular quantum channels as their inputs, rather than a density matrix describing the quantum state of a specific system. In practice superchannels will arise as effective channels constructed from a given channel sequence by a tensor rearrangement that changes the effective input and outputs of the map. This rearrangement is done by using graphical tensor manipulations introduced in [Chapter 2](#). Our choice of naming is analogous to how superoperators are operator maps, so we use the name superchannel for channel maps. The abstract description of superchannels has previously been introduced under the name *quantum supermaps* [[CDP08](#)]. The approach presented here generalizes this formalism further and introduces several new applications.

The benefit of this approach is that it provides natural description of many realistic experimental situations that lead to dynamics strictly broader than those that can be described by CP maps acting on density matrices for a quantum system. In particular the standard CP-map formalism of quantum channels breaks down in the presence of initial system-environment correlations or non-markovian noise process between and in this case the evolution of the system may appear to be non-completely-positive [[Pec94](#), [JSS04](#), [WHE+04](#), [Zim06](#), [KMRRS07](#), [CTZ08](#), [Woo08](#), [MS10](#), [BGTW11](#), [Mod12](#)]. There have been numerous proposals com how to treat “physical” non-CP which can arise in the presence of initial correlations [[ŠB01](#), [SSGF04](#), [JSS04](#), [SS05](#), [SL09](#), [RRMAG10](#), [Mod12](#), [DL15](#)], while

for non-markovian environments one generally has to resort to a continuous time master-equation description of the system and environment [WECC08, RHP10, RHP14].

An operational approach for describing experiments performed on a quantum system with initial system-environment correlations by an effective channel was proposed by Modi in the *M-map* construction which is motivated by explicitly considering the role of state preparation in device characterization [Mod12]. The M-map is an effective map which takes the channel corresponding to the state preparation procedure as the input, and allows for an operational description of the dynamics of the initially correlated system. The superchannel formalism we develop here is a direct generalization of the M-map approach, and we return to the example of describing an initially correlated system in Chapter 5.

Since superchannels are themselves quantum channels, they have all the same properties, representations, and transformations as discussed in Chapter 2. Consider two Hilbert spaces of operator maps $T(\mathcal{X}_1, \mathcal{Y}_1)$ and $T(\mathcal{X}_2, \mathcal{Y}_2)$ for Hilbert spaces $\mathcal{X}_1, \mathcal{X}_2, \mathcal{Y}_1, \mathcal{Y}_2$. We define the space of super channels to be maps of the form $\mathcal{M} : T(\mathcal{X}_1, \mathcal{Y}_1) \rightarrow T(\mathcal{X}_2, \mathcal{Y}_2)$. Let $\mathcal{E}_1 \in T(\mathcal{X}_1, \mathcal{Y}_1)$,

$$\mathcal{M} : L(\mathcal{X}_1 \otimes \mathcal{Y}_1) \rightarrow L(\mathcal{X}_2 \otimes \mathcal{Y}_2) \quad (4.1)$$

$$\mathcal{M} : \mathcal{E}_1 \mapsto \mathcal{E}_2 = \mathcal{M}(\mathcal{E}_1) \quad (4.2)$$

$$\Lambda_{\mathcal{E}_2} = \mathcal{M}(\Lambda_{\mathcal{E}_1}) \quad (4.3)$$

Hence we can see that the set of superchannels can be considered as a set of regular quantum channels, but where the input and output Hilbert spaces are bipartite tensor product spaces: $\mathcal{M} \in T(\mathcal{X}_1 \otimes \mathcal{Y}_1, \mathcal{X}_2 \otimes \mathcal{Y}_2)$. If we want the output space to be density matrices, this is achieved by setting $\mathcal{Y}_2 \cong \mathbb{C}$ as the trivial output space. In effect density matrices are a class of channels which map density matrices to complex numbers via $\text{Tr}[\rho^T \rho_{\mathcal{E}}]$.

Superchannels can also be used to describe other common protocols in quantum information processing such as clifford twirling by a unitary 2-design [TDL01, Cha05, GAE07, RS09, DCEL09], symmetrized characterization [ESM+07], randomized benchmarking [MBKE11, MGE12, MGJ+12] and its recently proposed generalization of simultaneous randomized benchmarking [GCM+12]. We present a description of twirling as a superchannel, and show how ideal interleaved randomized benchmarking can also be considered by taking the k^{th} power of this superchannel. We then generalize this to the n -twirling superchannel which is the simultaneous twirling of n subsystems of a composite channel. This is closely related to the characterization schemes of symmetrized characterization and simultaneous randomized benchmarking. Symmetrized characterization is an example of twirling all individual subsystems of a bipartite quantum system, and then performing a permutation of subsystems to estimate the average errors strength of all weight- k Pauli

errors over the subsystems. Simultaneous randomized benchmarking, which is equivalent to applying k rounds of twirling to each subsystem of a bipartite subsystem, in principle allows for an estimation of correlated errors between subsystems and was originally presented as a partial characterization protocol to assess addressability and estimate cross-talk errors for the control of superconducting qubits. In [Section 4.5.1](#) we present a superchannel for subsystem twirling which describes how in principle this may be applied to any number of subsystems of a multipartite channel.

The outline of this chapter is as follows. In [Section 4.2](#) we formally define a quantum superchannel. In [Section 4.3](#) we give several examples of how to construct superchannels from an underlying sequence of channels, and we discuss the properties of these channels in terms of the properties of the underlying channels. In particular we prove several propositions for when the constructed superchannels will be trace-normalized, hermitian preserving, completely-positive and trace-preserving. In [Section 4.4](#) we develop a generalization of quantum process tomography to describe how in principle an experimenter may completely characterize an unknown quantum superchannel. Finally in [Section 4.5](#) we introduce the twirling and n -twirling superchannels.

4.2 Definition of Quantum Superchannels

A quantum superchannel is itself a quantum channel — a completely positive operator map. Superchannels, however, are in general *effective* channel that we construct out of a sequence of quantum channels which describe a physical process in order to change what we regard to be the input and output of this description. Formally we define a quantum superchannel as follows:

Definition 4.2.1 (Quantum Superchannel). A quantum superchannel is a completely positive operator map $\mathcal{M} \in T(\mathcal{X}_1 \otimes \mathcal{Y}_2, \mathcal{X}_2 \otimes \mathcal{Y}_2)$ which maps bipartite positive operators $\Lambda_1 \in T(\mathcal{X}_1 \otimes \mathcal{Y}_1)$ to bipartite positive operators $\Lambda_2 \in T(\mathcal{X}_2 \otimes \mathcal{Y}_2)$. Let $\mathcal{E} \in C(\mathcal{X}_1, \mathcal{Y}_1)$ be a quantum channel. We define the action of $\mathcal{M}(\mathcal{E})$ acting on this channel by $\Lambda_{\mathcal{M}(\mathcal{E})} \equiv \mathcal{M}(\Lambda_{\mathcal{E}})$.

Since superchannels are simply channels they have all the same representations and transformation for channels as described in [Chapter 2](#). We will only make use of the Choi-matrix and superoperator representations, and we will only consider these operators defined in terms of the col-stacking vectorization convention.

Consider a superchannel $\mathcal{M} \in T(\mathcal{X}_1 \otimes \mathcal{Y}_1, \mathcal{X}_2 \otimes \mathcal{Y}_2)$. The Choi-matrix $\Lambda_{\mathcal{M}} \in L(\mathcal{X}_1 \otimes$

$\mathcal{Y}_1 \otimes \mathcal{X}_2 \otimes \mathcal{Y}_2$) is represented graphically as an 8-index tensor

$$\Lambda_{\mathcal{M}} = \begin{array}{c} \mathcal{X}_1 \\ \mathcal{Y}_1 \\ \mathcal{X}_2 \\ \mathcal{Y}_2 \end{array} \begin{array}{|c|} \hline \Lambda_{\mathcal{M}} \\ \hline \end{array} \begin{array}{c} \mathcal{X}_1 \\ \mathcal{Y}_1 \\ \mathcal{X}_2 \\ \mathcal{Y}_2 \end{array} \quad (4.4)$$

where the top 4 indices correspond to the input Hilbert space $L(\mathcal{X}_1 \otimes \mathcal{Y}_1)$, the bottom 4 indices the output Hilbert space $L(\mathcal{X}_2 \otimes \mathcal{Y}_2)$, and we have labelled the corresponding Hilbert spaces of the wires. Evolution is given by

$$\mathcal{M}(\Lambda_{\mathcal{E}}) = \text{Tr}_{12}[\Lambda_{\mathcal{E}}^T \otimes (\mathbb{1}_{\mathcal{X}_2 \otimes \mathcal{Y}_2}) \Lambda_{\mathcal{M}}] \quad (4.5)$$

where $\text{Tr}_{12} \equiv \text{Tr}_{\mathcal{X}_1, \mathcal{Y}_1}$ refers to tracing over the first 2 indices. This is illustrated as

$$\begin{array}{c} \Lambda_{\mathcal{M}(\rho_{\mathcal{E}})} \end{array} = \begin{array}{c} \rho_{\mathcal{E}} \\ \Lambda_{\mathcal{M}} \end{array} \quad (4.6)$$

The Choi-matrix has the usual properties as discussed in [Section 2.3.4](#). In particular it is defined in terms of the Choi-Jamiołkowski isomorphism

$$\Lambda_{\mathcal{M}} = (\mathbb{1} \otimes \mathcal{M})(|\mathbb{1}_{\mathcal{X}_1 \otimes \mathcal{Y}_1}\rangle\rangle\langle\langle \mathbb{1}_{\mathcal{X}_1 \otimes \mathcal{Y}_1}|). \quad (4.7)$$

A superchannel \mathcal{M} satisfies the following the properties:

$$\text{Tr}[\Lambda_{\mathcal{M}}] = d_{x_1} d_{y_1} \quad (4.8)$$

$$\mathcal{M} \text{ is HP} \Leftrightarrow \Lambda_{\mathcal{M}}^\dagger = \Lambda_{\mathcal{M}} \quad (4.9)$$

$$\mathcal{M} \text{ is CP} \Leftrightarrow \Lambda_{\mathcal{M}} \geq 0 \quad (4.10)$$

$$\mathcal{M} \text{ is TP} \Leftrightarrow \text{Tr}_{\mathcal{X}_1, \mathcal{Y}_2}[\Lambda_{\mathcal{M}}] = \mathbb{1}_{\mathcal{X}_1 \otimes \mathcal{Y}_2} \quad (4.11)$$

where d_{x_1} and d_{y_1} are the dimensions of \mathcal{X}_1 and \mathcal{Y}_1 respectively.

The superoperator representation $\mathcal{S}_{\mathcal{M}} \in L(\mathcal{X}_1 \otimes \mathcal{Y}_1 \otimes \mathcal{X}_1 \otimes \mathcal{Y}_1 \otimes \mathcal{X}_2 \otimes \mathcal{Y}_2 \otimes \mathcal{X}_2 \otimes \mathcal{Y}_2)$ of the superchannel \mathcal{M} is an 8-index tensor represented graphically as

$$\mathcal{S}_{\mathcal{M}} = \begin{array}{c} \mathcal{X}_2 \\ \mathcal{Y}_2 \\ \mathcal{X}_2 \\ \mathcal{Y}_2 \end{array} \begin{array}{|c} \hline \mathcal{S}_{\mathcal{M}} \\ \hline \end{array} \begin{array}{c} \mathcal{X}_1 \\ \mathcal{Y}_1 \\ \mathcal{X}_1 \\ \mathcal{Y}_1 \end{array} \quad (4.12)$$

where the right 4 indices correspond to the vectorized input space, and the left 4-indices to the vectorized output space. As shown in [Section 2.4](#) it is defined in terms of the Choi-matrix by the reshuffling operation $\mathcal{S}_{\mathcal{M}} = \Lambda_{\mathcal{M}}^R$. The output states are vectorized channels

$$|\Lambda_{\mathcal{M}(\mathcal{E})}\rangle\rangle = \mathcal{S}_{\mathcal{M}}|\Lambda_{\mathcal{E}}\rangle\rangle \quad (4.13)$$

which is depicted graphically as

$$\begin{array}{c} \text{---} \\ \text{---} \\ \text{---} \end{array} \begin{array}{|c} \hline \Lambda_{\mathcal{M}(\Lambda_{\mathcal{E}})} \\ \hline \end{array} = \begin{array}{c} \text{---} \\ \text{---} \\ \text{---} \end{array} \begin{array}{|c} \hline \mathcal{S}_{\mathcal{M}} \\ \hline \end{array} \begin{array}{|c} \hline \Lambda_{\mathcal{E}} \\ \hline \end{array} \quad (4.14)$$

4.3 Constructing Quantum Superchannels

We now describe how superchannels may be constructed from a sequence of quantum channels by rearranging the corresponding tensor network to change the object that is the input. The basic element is a superchannel which takes a single channel \mathcal{E} as input, where \mathcal{E} one channel from a given sequence describing the open system evolution of a quantum system:

Definition 4.3.1 (Single Input Superchannel). Consider the composition of three operator maps $\mathcal{F}_2 \mathcal{E} \mathcal{F}_1 \in T(\mathcal{X}, \mathcal{Y})$ where $\mathcal{F}_1 \in T(\mathcal{X}, \mathcal{X}_1)$, $\mathcal{E} \in T(\mathcal{X}_1, \mathcal{Y}_1)$, $\mathcal{F}_2 \in T(\mathcal{Y}_1, \mathcal{Y})$. We define the single input superchannel $\mathcal{M} \in T(\mathcal{X}_1 \otimes \mathcal{Y}_1, \mathcal{X} \otimes \mathcal{Y})$ by the map

$$\mathcal{M}(\mathcal{E})(\rho) \equiv \mathcal{F}_2 \mathcal{E} \mathcal{F}_1(\rho) \quad (4.15)$$

for all $\mathcal{E} \in T(\mathcal{X}_1, \mathcal{Y}_1)$, $\rho \in L(\mathcal{X})$.

We can derive the properties of the single input superchannel in terms of the properties of the underlying channels it is constructed from. These are given by the following proposition:

Proposition 4.3.1. *Let $\mathcal{M} \in T(\mathcal{X}_1 \otimes \mathcal{Y}_1, \mathcal{X} \otimes \mathcal{Y})$ be a single input superchannel as per Definition 4.3.1. Then \mathcal{M} satisfies the following properties:*

1. $\mathcal{M} = \mathcal{F}_1^T \otimes \mathcal{F}_2$.
2. $\text{Tr}[\Lambda_{\mathcal{M}}] = \text{Tr}[\Lambda_{\mathcal{F}_1}] \text{Tr}[\Lambda_{\mathcal{F}_2}]$.
3. \mathcal{M} is HP if and only if \mathcal{F}_1 and \mathcal{F}_2 are HP.
4. \mathcal{M} is CP if and only if \mathcal{F}_1 and \mathcal{F}_2 are CP.
5. \mathcal{M} is TP if and only if \mathcal{F}_1 is a unital and \mathcal{F}_2 is TP.

Note that while the output space of the single input superchannel can be thought of as the space of channels $T(\mathcal{X}, \mathcal{Y})$, in terms of the underlying physical system it is actually the tensor product of the input and output Hilbert spaces $L(\mathcal{X}) \otimes L(\mathcal{Y}) \cong L(\mathcal{X} \otimes \mathcal{Y})$. Hence, for example, we can't use this superchannel to produce an entangled output in $L(\mathcal{X} \otimes \mathcal{Y})$ unless we entangle both the input and output physical spaces with an ancilla system.

We can also compose superchannels to construct bipartite superchannels which act on a bipartite input state space using the same techniques discussed for bipartite channels in Section 2.5. A more interesting case is where the output of ones superchannel feeds into the other which would arise, for example, if we construct a superchannel to describe the sequence of channels

$$\mathcal{F}_3 \mathcal{E}_2 \mathcal{F}_2 \mathcal{E}_1 \mathcal{F}_1 \in T(\mathcal{X}, \mathcal{Y}) \quad (4.16)$$

where we want to have $\mathcal{E}_1 \otimes \mathcal{E}_2$ be the input to an effective superchannel description. We make the following definition for composition superchannels:

Definition 4.3.2 (Composition Superchannel). Consider two superchannels

$$\mathcal{M}_1 \in T(\mathcal{X}_1 \otimes \mathcal{Y}_1, \mathcal{X} \otimes \mathcal{Z}), \quad \mathcal{M}_2 \in T(\mathcal{X}_2 \otimes \mathcal{Y}_2, \mathcal{Z} \otimes \mathcal{Y}).$$

We define the *composition superchannel* $\mathcal{M}_{12} \in T(\mathcal{X}_1 \otimes \mathcal{X}_2 \otimes \mathcal{Y}_1 \otimes \mathcal{Y}_2, \mathcal{X} \otimes \mathcal{Y})$ by the map

$$\mathcal{M}_{12}(\mathcal{E}_1 \otimes \mathcal{E}_2) \equiv \mathcal{M}_2(\mathcal{E}_2) \circ \mathcal{M}_1(\mathcal{E}_1) \quad (4.17)$$

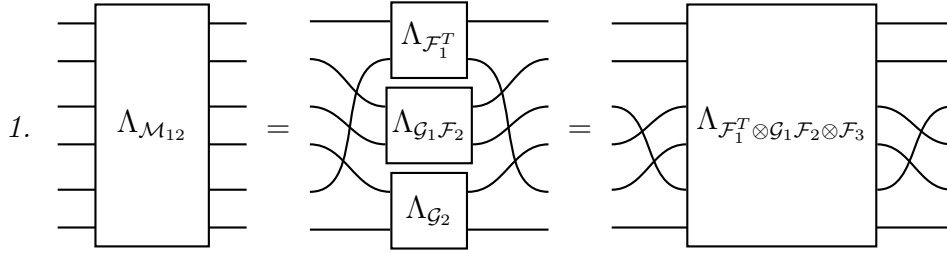
Where

$$\mathcal{M}_2(\mathcal{E}_2) \circ \mathcal{M}_1(\mathcal{E}_1)(\rho) = \mathcal{M}_2(\mathcal{E}_2)(\mathcal{M}_1(\mathcal{E}_1)(\rho)) \quad (4.18)$$

for all $\mathcal{E}_1 \in T(\mathcal{X}_1, \mathcal{Y}_1)$, $\mathcal{E}_2 \in T(\mathcal{X}_2, \mathcal{Y}_2)$, $\rho \in L(\mathcal{X})$.

The properties of the composition superchannel are similar to those for single-input superchannel. However if we explicitly assume that the two channels being composed are themselves single-input superchannels we may deduce the following properties for \mathcal{M}_{12} :

Proposition 4.3.2. *Consider two single input superchannels $\mathcal{M}_1 \in T(\mathcal{X}_1 \otimes \mathcal{Y}_1, \mathcal{X} \otimes \mathcal{Z})$, $\mathcal{M}_2 \in T(\mathcal{X}_2 \otimes \mathcal{Y}_2, \mathcal{Z} \otimes \mathcal{Y})$ where $\mathcal{M} = \mathcal{F}_1^T \otimes \mathcal{F}_2$ and $\mathcal{M}_2 = \mathcal{G}_1^T \otimes \mathcal{G}_2$. The composition superchannel \mathcal{M}_{12} satisfies the following properties:*



2. $\text{Tr}[\Lambda_{\mathcal{M}_{12}}] = \text{Tr}[\Lambda_{\mathcal{F}_1}] \text{Tr}[\Lambda_{\mathcal{G}_1 \mathcal{F}_2}] \text{Tr}[\Lambda_{\mathcal{G}_2}]$.

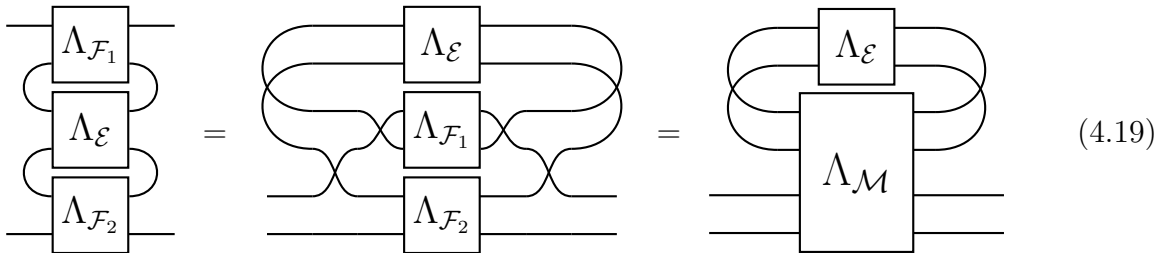
3. \mathcal{M}_{12} is HP if and only if $\mathcal{F}_1, \mathcal{G}_1 \mathcal{F}_2$ and \mathcal{G}_2 are HP.

4. \mathcal{M}_{12} is CP if and only if $\mathcal{F}_1, \mathcal{G}_1 \mathcal{F}_2$ and \mathcal{G}_2 are CP.

5. \mathcal{M} is TP if and only if \mathcal{F}_1 is unital, \mathcal{G}_2 is TP, and $\Lambda_{\mathcal{G}_1 \mathcal{F}_2} = \mathbb{1}_{\mathcal{X}_2 \otimes \mathcal{Y}_1}$.

We see here that we can construct a superchannel for the sequence of channels $\mathcal{F}_3 \mathcal{E}_2 \mathcal{F}_2 \mathcal{E}_1 \mathcal{F}_1$ with inputs $\mathcal{E}_1, \mathcal{E}_2$ by setting $\mathcal{G}_1 = \mathcal{I}$, and $\mathcal{G}_2 = \mathcal{F}_3$ in [Proposition 4.3.2](#). We further note the channel constructed in [Proposition 4.3.2](#) can never be TP as the condition $\Lambda_{\mathcal{G}_1 \mathcal{F}_2} = \mathbb{1}_{\mathcal{X}_2 \otimes \mathcal{Y}_1}$ would require that $\mathcal{G}_1 \mathcal{F}_2$ be trace increasing. We now sketch proofs of these propositions.

Proof of [Proposition 4.3.1](#). We begin by proving $\mathcal{M} = \mathcal{F}_1^T \otimes \mathcal{F}_2$. We may do this in terms of the Choi-matrix representation by the following tensor network manipulations.



where we have defined a superchannel $\mathcal{M} \in T(\mathcal{X}_1 \otimes \mathcal{Y}_1, \mathcal{X} \otimes \mathcal{Y})$ by

$$\Lambda_{\mathcal{M}} = \begin{array}{c} \Lambda_{\mathcal{F}_1} \\ \Lambda_{\mathcal{F}_2} \end{array} \quad (4.20)$$

$$\Lambda_{\mathcal{M}} = \begin{array}{c} \Lambda_{\mathcal{F}_1^T} \\ \Lambda_{\mathcal{F}_2} \end{array} = \Lambda_{\mathcal{F}_1^T \otimes \mathcal{F}_2} \quad (4.21)$$

and hence $\mathcal{M} = \mathcal{F}_1^T \otimes \mathcal{F}_2$. Alternatively we can also see this using vectorization identities:

$$|\Lambda_{\mathcal{F}_1 \varepsilon \mathcal{F}_1}\rangle\rangle = \mathcal{V}_2 |\mathcal{S}_{\mathcal{F}_2} \mathcal{S}_{\varepsilon} \mathcal{S}_{\mathcal{F}_1}\rangle\rangle \quad (4.22a)$$

$$= \mathcal{V}_2 (\mathcal{S}_{\mathcal{F}_1}^T \otimes \mathcal{S}_{\mathcal{F}_2}) |\mathcal{S}_{\varepsilon}\rangle\rangle \quad (4.22b)$$

$$= \mathcal{V}_2 (\mathcal{S}_{\mathcal{F}_1}^T \otimes \mathcal{S}_{\mathcal{F}_2}) \mathcal{V}_2^\dagger |\Lambda_{\varepsilon}\rangle\rangle \quad (4.22c)$$

$$= \mathcal{S}_{\mathcal{F}_1^T \otimes \mathcal{F}_2} |\Lambda_{\varepsilon}\rangle\rangle \quad (4.22d)$$

$$= \mathcal{S}_{\mathcal{M}} |\Lambda_{\varepsilon}\rangle\rangle \quad (4.22e)$$

where the superoperator representation of \mathcal{M} is given by

$$\mathcal{S}_{\mathcal{M}} = \begin{array}{c} \mathcal{S}_{\mathcal{F}_1^T} \\ \mathcal{S}_{\mathcal{F}_2} \end{array} = \mathcal{S}_{\mathcal{F}_1^T \otimes \mathcal{F}_2} \quad (4.23)$$

Properties 2, 3 and 4 follow directly from the fact that $\mathcal{M} = \mathcal{F}_1^T \otimes \mathcal{F}_2$, and that \mathcal{F}_1^T is CP if and only if \mathcal{F}_1 is CP, and similarly for HP.

For property 5 we have

$$\Lambda_{\mathcal{M}} \text{ is TP} \Leftrightarrow \begin{array}{c} \text{---} \\ \text{---} \\ \Lambda_{\mathcal{M}} \\ \text{---} \\ \text{---} \\ \text{---} \end{array} = \text{---} \quad (4.24)$$

$$\Leftrightarrow \begin{array}{c} \text{---} \\ \text{---} \\ \Lambda_{\mathcal{F}_1} \\ \text{---} \\ \text{---} \\ \Lambda_{\mathcal{F}_2} \\ \text{---} \\ \text{---} \end{array} = \text{---} \quad (4.25)$$

$$\Leftrightarrow \begin{array}{c} \text{---} \\ \text{---} \\ \Lambda_{\mathcal{F}_1} \\ \text{---} \\ \text{---} \\ \Lambda_{\mathcal{F}_2} \\ \text{---} \\ \text{---} \end{array} = \text{---} \quad (4.26)$$

Now $\text{Tr}_1[\Lambda_{\mathcal{F}_1}] = \mathbb{1}_{\mathcal{X}}$ if and only if \mathcal{F}_1 is unital, and $\text{Tr}_2[\Lambda_{\mathcal{F}_2}] = \mathbb{1}_{\mathcal{Y}}$ if and only if \mathcal{F}_2 is HP. \square

Proof of Proposition 4.3.2. To prove property 1 we have

$$(4.27)$$

$$(4.28)$$

Properties 2,3,4 follow directly from the construction of \mathcal{M}_{12} in property 1. For property 5 we have that partial tracing over the last two indices of $\Lambda_{\mathcal{M}_{12}}$ is

$$\text{Tr}_{56}[\Lambda_{\mathcal{M}_{12}}] = \text{Tr}_1[\Lambda_{\mathcal{F}_1}] \otimes \Lambda_{\mathcal{G}_1\mathcal{F}_2} \otimes \text{Tr}_2[\Lambda_{\mathcal{G}_2}] \quad (4.29)$$

Hence we have $\text{Tr}_{56}[\Lambda_{\mathcal{M}_{12}}] = \mathbb{1}_{\mathcal{X}_1 \otimes \mathcal{X}_2 \otimes \mathcal{Y}_1 \otimes \mathcal{Y}_2}$ if and only if

$$\begin{aligned} \text{Tr}_{56}[\Lambda_{\mathcal{M}_{12}}] = \mathbb{1}_{\mathcal{X}_1 \otimes \mathcal{X}_2 \otimes \mathcal{Y}_1 \otimes \mathcal{Y}_2} &\Leftrightarrow \text{Tr}_1[\Lambda_{\mathcal{F}_1}] \otimes \Lambda_{\mathcal{G}_1\mathcal{F}_2} \otimes \text{Tr}_2[\Lambda_{\mathcal{G}_2}] = \mathbb{1}_{\mathcal{X}_1 \otimes \mathcal{X}_2 \otimes \mathcal{Y}_1 \otimes \mathcal{Y}_2} \\ &\Leftrightarrow \text{Tr}_1[\Lambda_{\mathcal{F}_1}] = \mathbb{1}_{\mathcal{X}_1} \\ &\quad \Lambda_{\mathcal{G}_1\mathcal{F}_2} = \mathbb{1}_{\mathcal{X}_2 \otimes \mathcal{Y}_1} \\ &\quad \text{Tr}_2[\Lambda_{\mathcal{G}_2}] = \mathbb{1}_{\mathcal{Y}_2} \end{aligned}$$

and so we have \mathcal{M}_{12} is TP if and only if \mathcal{F}_1 is unital, \mathcal{G}_2 is TP, and $\Lambda_{\mathcal{G}_1\mathcal{F}_2} = \mathbb{1}_{\mathcal{X}_2 \otimes \mathcal{Y}_1}$. \square

4.3.1 Post-Selected Superchannels

As previously mentioned the output space of a superchannel \mathcal{M} is actually the tensor product of the input and output Hilbert spaces for the underlying physical system for which we constructed the superchannel from a composition sequence of channels. If we

wish to include a fixed input or measurement so that we are post-selecting (technically pre-selecting for the fixed input) on the system to be in a fixed state ρ we can do this by contracting with the appropriate index of \mathcal{M} . We define a pre and post selected reduced superchannel as follow:

Definition 4.3.3 (Input Post-Selected Superchannel). Consider a superchannel $\mathcal{M} \in T(\mathcal{X}_1 \otimes \mathcal{Y}_1, \mathcal{X}_\otimes \mathcal{Y})$ and an operator $A \in L(\mathcal{X})$. We define the *input post-selected superchannel* $\mathcal{M}_A^{\text{in}} \in T(\mathcal{X}_1 \otimes \mathcal{X}_2, \mathcal{Y})$ to be the channel satisfying $\mathcal{M}_A^{\text{pre}}(\mathcal{E}) = \mathcal{M}(\mathcal{E})(A)$ for all $\mathcal{E} \in T(\mathcal{X}_1, \mathcal{Y}_1)$.

Definition 4.3.4 (Output Post-Selected Superchannel). Consider a superchannel $\mathcal{M} \in T(\mathcal{X}_1 \otimes \mathcal{Y}_1, \mathcal{X}_\otimes \mathcal{Y})$ and an operator $B \in L(\mathcal{Y})$. We define the *output post-selected superchannel* $\mathcal{M}_B^{\text{out}} \in T(\mathcal{X}_1 \otimes \mathcal{X}_2, \mathcal{X})$ to be the channel satisfying $\mathcal{M}_B^{\text{post}}(\mathcal{E})(A) = \text{Tr}[B^\dagger \mathcal{M}(\mathcal{E})(A)]$ for all $A \in L(\mathcal{X})$ and $\mathcal{E} \in T(\mathcal{X}_1, \mathcal{Y}_1)$.

The Choi-matrix for $\mathcal{M}_A^{\text{in}}$ and $\mathcal{M}_B^{\text{out}}$ are given by

$$\Lambda_{\mathcal{M}_A^{\text{in}}} = \Lambda_{\mathcal{M}} \text{ with } A \text{ feedback} \quad (4.30)$$

$$\Lambda_{\mathcal{M}_B^{\text{post}}} = \Lambda_{\mathcal{M}} \text{ with } \bar{B} \text{ feedback} \quad (4.31)$$

If \mathcal{M} is a single input superchannel, then the properties of these post-selected channels follow from [Proposition 4.3.1](#). We detail them with the following corollaries:

Corollary 4.3.1. *Let $\mathcal{M} \in T(\mathcal{X}_1 \otimes \mathcal{Y}_1, \mathcal{X} \otimes \mathcal{Y})$ be a single input superchannel given by $\mathcal{M} = \mathcal{F}_1^T \otimes \mathcal{F}_2$, and let $A \in L(\mathcal{X})$. Then $\mathcal{M}_A^{\text{in}}$ has the following properties:*

1. $\Lambda_{\mathcal{M}_A^{\text{in}}} = \mathcal{F}_1(\rho) \otimes \Lambda_{\mathcal{F}_2}$.

2. $\text{Tr}[\Lambda_{\mathcal{M}_A^{\text{in}}}] = \text{Tr}[\mathcal{F}_1(\rho)] \text{Tr}[\Lambda_{\mathcal{F}_2}]$.
3. $\mathcal{M}_A^{\text{in}}$ is HP $\iff \mathcal{F}_1, \mathcal{F}_2$ are HP and A is Hermitian.
4. $\mathcal{M}_A^{\text{in}}$ is CP $\iff \mathcal{F}_1, \mathcal{F}_2$ are CP and $A \geq 0$.
5. $\mathcal{M}_A^{\text{in}}$ is TP $\iff \mathcal{F}_2$ is TP and $\mathcal{F}_1(A) = \mathbb{1}_{\mathcal{X}_1}$.

Corollary 4.3.2. *Let $\mathcal{M} \in T(\mathcal{X}_1 \otimes \mathcal{Y}_1, \mathcal{X} \otimes \mathcal{Y})$ be a single input superchannel given by $\mathcal{M} = \mathcal{F}_1^T \otimes \mathcal{F}_2$, and let $B \in L(\mathcal{Y})$. Then $\mathcal{M}_B^{\text{out}}$ has the following properties:*

1. $\Lambda_{\mathcal{M}_B^{\text{out}}} = W_{2:3}(\Lambda_{\mathcal{F}_1} \otimes \mathcal{F}_2^\dagger(B))W_{2:3}^\dagger$.
2. $\text{Tr}[\Lambda_{\mathcal{M}_B^{\text{out}}}] = \text{Tr}[\Lambda_{\mathcal{F}_1}] \text{Tr}[\mathcal{F}_2^\dagger(B)]$.
3. $\mathcal{M}_B^{\text{out}}$ is HP $\iff \mathcal{F}_1, \mathcal{F}_2$ are HP and B is Hermitian.
4. $\mathcal{M}_B^{\text{out}}$ is CP $\iff \mathcal{F}_1, \mathcal{F}_2$ are CP and $B \geq 0$.
5. $\mathcal{M}_B^{\text{out}}$ is TP $\iff \mathcal{F}_1$ is unital and $\mathcal{F}_2^\dagger(B) = \mathbb{1}_{\mathcal{Y}_2}$.

Proof. The proof of [Corollaries 4.3.1](#) and [4.3.2](#) follow straight forwardly from the graphical representations of the post-selected channels, and \mathcal{M} . \square

We comment that our definition of our definition for post-selected superchannels is similar to reduced channels in [Section 2.5.3](#), however the difference being the index that is contracted for the pre- and post-selected superchannel is the first and last index of the bipartite output Hilbert space respectively. We can also consider tracing over the output of the superchannel by setting $B = \mathbb{1}$ in [Definition 4.3.4](#). We make a point that these pre- and post-selected superchannels need not be normalized in the usual convention as per [Eq. \(4.8\)](#). To see this suppose $\Lambda_{\mathcal{M}}$ has trace $\text{Tr}[\Lambda_{\mathcal{M}}] = d_{x_1}d_{y_1}$. Now suppose we construct the pre-selected channel by contracting with the maximally mixed state $\rho = \mathbb{1}/d_x$, where $d_x = \dim(\mathcal{X})$. In this case we have

$$\text{Tr}[\Lambda_{\mathcal{M}_\rho^{\text{pre}}}] = \frac{d_{x_1}d_{y_1}}{d}. \quad (4.32)$$

Similarly we have the same for the post-selected channel. Of particular note is that the input post-selected channel can never to TP if \mathcal{F}_1 is trace non-increasing and the operator A is a density matrix, as the output will have trace d_{x_2} . This is because pre and post selecting can be thought of as condition on a measurement outcome which has some probability of

success $0 \leq p \leq 1$. Hence if p is less than one this will be a trace decreasing channel. If we wish to preserve this property we can always renormalize by this probability to ensure the desired trace.

Our final type of superchannel we discuss is when the input and output Hilbert space are themselves composite vector spaces. In this case the single input channel in [Definition 4.3.1](#) takes as input channels $\mathcal{E} = \mathcal{E}_1 \otimes \mathcal{E}_2$ and the physical input and output spaces will themselves be bipartite spaces. Similarly all the other classes of superchannels straightforwardly generalize to this situation. We note that post-selected superchannels can be useful for describing interesting dynamics when the underlying physical system is a composite system, as in this case we can pre or post select on a specific input channel or physical input or output state on subset of subsystems only. In the bipartite case, these kinds of superchannels can describe situations where one of the subsystems is a physical system, and the other is an environment system which is not directly accessible to an experimenter. In [Chapter 5](#) we give an explicit construction of a superchannel of this type that generalizes the quantum channel formalism to also describe a system that is initially correlated with its environment. In this case it can be useful to define a superchannel where we preselect on a channel input for one of the subsystems. In this case the subsystem preselected on will simply contract to a fixed channel between its underlying physical input and output space.

4.4 Characterizing Superchannels

In principle many measures and techniques for characterizing channels can also be used to characterize superchannels. In practice however, some measures such as average gate fidelity don't really make sense for a single input superchannel as defined in [Proposition 4.3.1](#) as the input and output state spaces are not really equivalent, even when they are of equal dimension. To see this consider that if the input is a unitary channel it will have a Choi-matrix isomorphic to a maximally entangled state $\Lambda_{\mathcal{U}} = |U\rangle\rangle\langle\langle U|$. While the output Hilbert space corresponds to tensor products of input and output density matrices $\rho_{\text{in}} \otimes \rho_{\text{out}}$. In general any measures we develop will be specific to the particular application the superchannel was constructed for such as the case of a superchannel for characterizing initial system environment correlations which we introduce in [Chapter 5](#). Regardless of the specific choice of characterization measure it is essential that it is, at least in principle, possible to experimentally reconstruct a complete description of an superchannel via a generalization of quantum process tomography, which is the main consideration of this section.

4.4.1 Quantum Superchannel Tomography

Let $\mathcal{X}_1, \mathcal{Y}_1, \mathcal{X}_2$ and \mathcal{Y}_2 be complex Hilbert spaces with dimension x_1, y_1, x_2 and y_2 respectively. Consider a quantum superchannel $\mathcal{M} \in T(\mathcal{X}_2 \otimes \mathcal{Y}_2, \mathcal{X}_1 \otimes \mathcal{Y}_1)$ constructed as per [Proposition 4.3.1](#), so that \mathcal{M} is a CPTP map. The Choi-matrix for \mathcal{M} is an 8-index tensor and hence in general has $x_1^2 \times y_1^2 \times x_2^2 \times y_2^2$ unknown parameters which must be determined. If all Hilbert spaces are of dimension d this corresponds to d^8 .

In order to completely specify these parameters we require that a tomographically complete input set of input channels \mathbf{E} which has minimum size $|\mathbf{E}| = x_2^2 y_2^2$, and a tomographically complete set of output channels \mathbf{M} which has minimum size $|\mathbf{M}| = x_1^2 y_1^2$. In practice the output set can be taken to be the tensor product of a tomographically complete set of input states \mathbf{Q}_1 and measurements \mathbf{P}_1 in the underlying physical systems:

$$\mathbf{M} = \{\mathcal{P}_{ij} : \Lambda_{\mathcal{P}_{ij}} = \bar{\rho}_i \otimes \rho_j, \rho_i \in \mathbf{Q}_1, \rho_j \in \mathbf{P}_1\}. \quad (4.33)$$

where the measurement is a 2 outcome measurement of $\{\rho_j, \mathbb{1} - \rho_j\}$.

Suppose for now that the dimensions of the input channel are equal $x_2 = y_2 = d$. If we restrict ourselves to unitary input channels $\mathcal{U} \in T(\mathcal{X}_2, \mathcal{Y}_2)$ such that $\Lambda_{\mathcal{U}} \equiv |U\rangle\rangle\langle\langle U|$, than we cannot span the full input Hilbert space of the superchannel. This is because the dimension of the subspace of superoperators spanned by the set of unitary channels is of dimension $d^4 - 2d^2 + 2$ [[RS09](#), [KdSR⁺14](#)].

Thus to completely characterize a superchannel we must have access to input channels which are not unitary. This can be done, for example, by using input channels corresponding to a *preparation procedure* or projective measurement [[MBC⁺12](#)]. A projective preparation procedure $\mathcal{P}_{ij} \in T(\mathcal{X}_1, \mathcal{X}_2)$ consists of an initial projection (or postselection) onto the state ρ_i followed by a rotation to the state ρ_j , so that

$$\mathcal{P}_{ij}(\rho) = \rho_j \text{Tr}[\rho_i^\dagger \rho]. \quad (4.34)$$

The corresponding preparation map on the system is described by the Choi matrix

$$\Lambda_{\mathcal{P}_{ij}} = \bar{\rho}_i \otimes \rho_j. \quad (4.35)$$

Hence we can define a tomographically complete set of projective preparation channels

$$\mathbf{E} = \{\Lambda_{\mathcal{P}_{ij}} : \bar{\rho}_i \otimes \rho_j, \rho_i \in \mathbf{Q}_2, \rho_j \in \mathbf{P}_2\}. \quad (4.36)$$

where $\mathbf{Q}_2, \mathbf{P}_2$ are tomographically complete sets of input states and measurements on $D(\mathcal{X}_2)$ and $D(\mathcal{Y}_2)$ respectively. We see here that this is a trace decreasing channel as $\text{Tr}[\Lambda_{ij}] = 1$ rather than d , and in particular the trace of the output is given by

$$\text{Tr}[\mathcal{P}_{ij}(\rho)] = \text{Tr}[\rho_i^\dagger \rho] \in [0, 1]. \quad (4.37)$$

The intuition for this is we can think of the preparation procedure being a two outcome measurement $\{\rho_i, \mathbb{1} - \rho_i\}$ where we are throwing away our results corresponding to the “failure” outcome $\mathbb{1} - \rho_i$. We note that in practice one need only a single projective measurement onto a fixed state ρ , and then can obtain the full set \mathbf{E} by combining this measurement with appropriate unitary rotations both before and after the measurement.

For the case of projective preparation procedures the probability of observing a count for a single input superchannel \mathcal{M} for input $\mathcal{P}_{ij} \in \mathbf{E}$ and output measurement $\mathcal{P}_{kl} \in \mathbf{M}$ is given by

$$p_{ijkl} = \text{Tr} [(\Lambda_{\mathcal{P}_{ij}}^T \otimes \Lambda_{\mathcal{P}_{kl}}^\dagger) \Lambda_{\mathcal{M}}] \quad (4.38)$$

$$= \text{Tr} [(\rho_i^\dagger \otimes \rho_j^T \otimes \rho_k^T \otimes \rho_l^\dagger) \Lambda_{\mathcal{M}}] \quad (4.39)$$

$$= \text{Tr} [(\rho_i \otimes \bar{\rho}_j \otimes \bar{\rho}_k \otimes \rho_l)^\dagger \Lambda_{\mathcal{M}}] \quad (4.40)$$

$$= \langle\langle \Pi_{ijkl} | \Lambda_{\mathcal{M}} \rangle\rangle \quad (4.41)$$

where $\Pi_{ijkl} \equiv \bar{\Lambda}_{\mathcal{E}_\alpha} \otimes \bar{\rho}_i \otimes M_j$. From here one may then reconstruct the Choi-matrix using any of the methods described in [Section 3.4](#). In the case of MLE tomographic reconstruction we can use the SDP from [Section 3.4.1](#) for the reconstruction of \mathcal{M} . This is equivalent to

$$\text{minimize:} \quad \sum_{i,j,k,l} \frac{N_{ijkl} (\text{Tr}[\Pi_{ijkl}^\dagger \Lambda_{\mathcal{M}}] - p_{ijkl})^2}{p_{ijkl} (1 - p_{ijkl})} \quad (4.42)$$

$$\text{subject to:} \quad \Lambda_{\mathcal{M}} \geq 0, \quad \text{Tr}[\Lambda_{\mathcal{M}}] = x_2 y_2$$

where we have assumed a normal approximation for the distribution of the observed probabilities p_{ijkl} defined by

$$p_{ijkl} = \frac{n_{ijkl} + \beta}{N_{ijkl} + K\beta} \quad (4.43)$$

where n_{ijkl} is the observed number of counts out of N_{ijkl} trials for a given projector Π_{ijkl} , K is the number of possible outcomes of a given measurement configuration, and β is a small hedging parameter used to overcome the problem that the normal approximation is not well-defined for $p_{ijkl} = 0, 1$ [[BK10a](#)]. We note that for a single input superchannel with a tomographic input set of projective preparations and two outcome measurements the value of K is 4 — two outcomes for the measurement involved in preparation, and two outcomes for the final measurement of the output physical state.

To perform quantum process tomography on composition superchannels, or superchannels on composite systems we can simply take tensor products of the corresponding tomographically complete measurement sets \mathbf{E} and \mathbf{M} . For the case of post-selected superchannels tomographic we simply drop the index in [Eq. \(4.41\)](#), [Eq. \(4.42\)](#) and [Eq. \(4.43\)](#) for the

fixed input our output post selected state and reconstruct the corresponding reduced rank tensor.

4.5 Twirling Superchannel

We now give an explicit example of a quantum superchannel that is frequently used in quantum information theory, though not usually thought of in these terms. The Clifford twirling operation discussed in [Section 3.3.1](#) is actually a quantum superchannel which can be thought of as the map which sends a channel \mathcal{E} to the twirled channel $\mathcal{W}(\mathcal{E})$.

Let $\mathcal{E} \in T(\mathcal{X})$ be a channel, the twirling superchannel $\mathcal{W} \in T(\mathcal{X} \otimes \mathcal{X})$ defined by [Eq. \(3.97\)](#)

$$\mathcal{S}_{\mathcal{W}(\mathcal{E})} \equiv \int dU \mathcal{S}_U \mathcal{S}_{\mathcal{E}} \mathcal{S}_U^\dagger. \quad (4.44)$$

is a single-input superchannel as defined in [Definition 4.3.1](#). Hence by [Proposition 4.3.1](#) we have

$$\mathcal{W} = \int dU (\mathcal{U}^\dagger)^T \otimes \mathcal{U} = \int dU \bar{\mathcal{U}} \otimes \mathcal{U} \quad (4.45)$$

We can generalize this further for comparing \mathcal{E} to a target unitary channel \mathcal{U}_0 by setting $\mathcal{E} \mapsto \mathcal{U}_0^\dagger \mathcal{E}$ and defining this twirling superchannel by

$$\mathcal{W}_{\mathcal{U}_0} = \int dU \bar{\mathcal{U}} \otimes (\mathcal{U} \mathcal{U}_0^\dagger) \quad (4.46)$$

To evaluate an explicit representation for the superchannel \mathcal{W} and its action on a channel \mathcal{E} we prove the following results:

Proposition 4.5.1 (Twirling Superchannel). *The twirling superchannel $\mathcal{W}_U \in T(\mathcal{X} \otimes \mathcal{X})$, where $U \in C(\mathcal{X})$ is a unitary channel, has superoperator and Choi-matrix representations*

$$\mathcal{S}_{\mathcal{W}_U} = |\Lambda_{\mathcal{I}}\rangle\rangle\langle\langle \Lambda_{\mathcal{R}_U} | + |\Lambda_{\mathcal{D}}\rangle\rangle\langle\langle \Lambda_{\mathcal{D}} | - |\Lambda_{\mathcal{D}}\rangle\rangle\langle\langle \Lambda_{\mathcal{R}_U} | \quad (4.47)$$

$$\Lambda_{\mathcal{W}_U} = \Lambda_{\mathcal{R}_U} \otimes \Lambda_{\mathcal{I}} + (\Lambda_{\mathcal{D}} - \Lambda_{\mathcal{R}_U}) \otimes \Lambda_{\mathcal{D}} \quad (4.48)$$

where \mathcal{I} is the identity channel \mathcal{D} is the completely depolarizing channel, and \mathcal{R}_U is defined as

$$\mathcal{R}_U = \frac{U - \mathcal{D}}{d^2 - 1} \quad (4.49)$$

Proposition 4.5.2 (Twirled Channel). *Let $\mathcal{E} \in T(\mathcal{X})$ be a channel, let $\mathcal{W} \in T(\mathcal{X} \otimes \mathcal{X})$ be the twirling superchannel, and let k be a positive integer. Then we have*

$$\mathcal{W}_{\mathcal{U}}(\mathcal{E})^k = \lambda^k \mathcal{I} + (p^k - \lambda^k) \mathcal{D} \quad (4.50)$$

$$p = \text{Tr}[\Lambda_{\mathcal{D}}\Lambda_{\mathcal{E}}] = \frac{1}{d} \text{Tr}[\Lambda_{\mathcal{E}}] \quad (4.51)$$

$$\lambda = \text{Tr}[\Lambda_{\mathcal{R}_{\mathcal{U}}}\Lambda_{\mathcal{E}}] = \frac{\text{Tr}[S_{\mathcal{U}}^{\dagger}S_{\mathcal{E}}] - p}{d^2 - 1} = \frac{\langle\langle U | \Lambda_{\mathcal{E}} | U \rangle\rangle - p}{d^2 - 1} \quad (4.52)$$

where \mathcal{I} is the identity channel, \mathcal{D} is the completely depolarizing channel, and $\Lambda_{\mathcal{U}} = |U\rangle\rangle\langle\langle U|$.

We can recover the standard twirling channel \mathcal{W} by setting $\mathcal{U} = \mathcal{I}$ to be the identity operation. Note that [Proposition 4.5.2](#) is related to the generalization of the twirling channel used in defining the RB channel in [Section 3.3.2](#). The basic twirling case is recovered by setting $k = 1$. We also note that in the case where \mathcal{E} is a TP channel we have $p = 1$.

We now prove these results.

Proof of [Proposition 4.5.1](#). First we note that $\mathcal{W}_{\mathcal{U}_0} = \mathcal{W}(\mathcal{I} \otimes \mathcal{U}_0)$, and then from [Eq. \(4.45\)](#) we have

$$\mathcal{S}_{\mathcal{W}} = \int dU \mathcal{V}_2^{\dagger}(\mathcal{S}_{\bar{U}} \otimes \mathcal{S}_U)\mathcal{V}_2 = \int dU (U \otimes \bar{U} \otimes \bar{U} \otimes U) \quad (4.53)$$

If we just consider the where \mathcal{V}_2 is the unravelling operator. To explicitly evaluate this

operator we use our definition of the unitary 2-design $\mathcal{S}_{\mathcal{U}_2}$ in Eq. (3.94):

$$\mathcal{S}_{\mathcal{W}} = \int dU \begin{array}{c} \boxed{U} \\ \boxed{\bar{U}} \\ \boxed{\bar{U}} \\ \boxed{U} \end{array} = \int dU \begin{array}{c} \boxed{\bar{U}} \\ \boxed{\bar{U}} \\ \boxed{U} \\ \boxed{U} \end{array} = \mathcal{S}_{\mathcal{U}_2} \quad (4.54)$$

$$\mathcal{S}_{\mathcal{W}} = \int dU \begin{array}{c} \boxed{U} \\ \boxed{\bar{U}} \\ \boxed{\bar{U}} \\ \boxed{U} \end{array} = \int dU \begin{array}{c} \boxed{\bar{U}} \\ \boxed{\bar{U}} \\ \boxed{U} \\ \boxed{U} \end{array} = \mathcal{S}_{\mathcal{U}_2} \quad (4.55)$$

$$= \frac{1}{d^2 - 1} \left(\begin{array}{c} \curvearrowright \curvearrowleft \\ \curvearrowright \curvearrowleft \end{array} + \begin{array}{c} \curvearrowright \curvearrowright \\ \curvearrowleft \curvearrowleft \end{array} \right) - \frac{1}{d(d^2 - 1)} \left(\begin{array}{c} \curvearrowright \curvearrowleft \\ \curvearrowleft \curvearrowright \end{array} + \begin{array}{c} \curvearrowright \curvearrowright \\ \curvearrowleft \curvearrowleft \end{array} \right) \quad (4.56)$$

$$= \left(\frac{1}{d^2 - 1} \right) (|\Lambda_{\mathcal{I}}\rangle\langle\Lambda_{\mathcal{I}}| + d^2|\Lambda_{\mathcal{D}}\rangle\langle\Lambda_{\mathcal{D}}| - |\Lambda_{\mathcal{I}}\rangle\langle\Lambda_{\mathcal{D}}| - |\Lambda_{\mathcal{D}}\rangle\langle\Lambda_{\mathcal{I}}|) \quad (4.57)$$

$$= |\Lambda_{\mathcal{I}}\rangle\left(\frac{\langle\langle\Lambda_{\mathcal{I}}| - \langle\langle\Lambda_{\mathcal{D}}|}{d^2 - 1} \right) + |\Lambda_{\mathcal{D}}\rangle\left(\frac{d^2\langle\langle\Lambda_{\mathcal{D}}| - \langle\langle\Lambda_{\mathcal{I}}|}{d^2 - 1} \right) \quad (4.58)$$

$$= |\Lambda_{\mathcal{I}}\rangle\langle\Lambda_{\mathcal{R}}| + |\Lambda_{\mathcal{D}}\rangle\langle\Lambda_{\mathcal{D}}| - \langle\langle\Lambda_{\mathcal{D}}| - \langle\langle\Lambda_{\mathcal{R}}|. \quad (4.59)$$

where $\langle\langle\Lambda_{\mathcal{R}}| = (\langle\langle\Lambda_{\mathcal{I}}| - \langle\langle\Lambda_{\mathcal{D}}|)/(d^2 - 1)$, and the Choi-matrices for the identity and completely depolarizing channels are given by

$$\begin{array}{c} \boxed{\Lambda_{\mathcal{I}}} \\ \text{---} \end{array} = \begin{array}{c} \curvearrowright \curvearrowleft \\ \curvearrowright \curvearrowleft \end{array} \Rightarrow |\Lambda_{\mathcal{I}}\rangle = \begin{array}{c} \Lambda_{\mathcal{I}} \\ \text{---} \end{array} = \begin{array}{c} \curvearrowright \\ \curvearrowleft \end{array} \quad (4.60)$$

$$\begin{array}{c} \boxed{\Lambda_{\mathcal{D}}} \\ \text{---} \end{array} = \begin{array}{c} \diamond \frac{1}{d} \\ \text{---} \end{array} = \begin{array}{c} \text{---} \\ \text{---} \end{array} \Rightarrow |\Lambda_{\mathcal{D}}\rangle = \begin{array}{c} \Lambda_{\mathcal{D}} \\ \text{---} \end{array} = \begin{array}{c} \curvearrowright \\ \curvearrowleft \end{array} \diamond \frac{1}{d} \quad (4.61)$$

Finally we note that

$$\begin{array}{c} \triangleleft \Lambda_{\mathcal{I}} \\ \text{---} \text{---} \text{---} \\ \square \mathcal{S}_{\mathcal{I} \otimes \mathcal{U}^\dagger} \\ \text{---} \text{---} \end{array} = \begin{array}{c} \text{---} \text{---} \\ \text{---} \text{---} \\ \square \mathcal{S}_{\mathcal{U}^\dagger} \\ \text{---} \text{---} \end{array} = \begin{array}{c} \triangleleft \mathcal{S}_{\mathcal{U}} \\ \text{---} \text{---} \end{array} = \begin{array}{c} \triangleleft \Lambda_{\mathcal{U}} \\ \text{---} \text{---} \end{array} \quad (4.62)$$

and

$$\begin{array}{c} \triangleleft \Lambda_{\mathcal{D}} \\ \text{---} \text{---} \end{array} \square \mathcal{S}_{\mathcal{I} \otimes \mathcal{U}^\dagger} = \diamond \frac{1}{d} \begin{array}{c} \text{---} \text{---} \\ \text{---} \text{---} \\ \square \mathcal{S}_{\mathcal{U}^\dagger} \\ \text{---} \text{---} \end{array} = \diamond \frac{1}{d} \begin{array}{c} \text{---} \\ \text{---} \\ \text{---} \end{array} = \begin{array}{c} \triangleleft \Lambda_{\mathcal{D}} \\ \text{---} \text{---} \end{array} \quad (4.63)$$

as \mathcal{U}^\dagger is TP implies that $\langle\langle \mathbb{1} | \mathcal{S}_{\mathcal{U}^\dagger} = \langle\langle \mathbb{1} |$.

Hence we have

$$\mathcal{S}_{\mathcal{W}_U} = |\Lambda_{\mathcal{I}}\rangle\rangle\langle\langle \Lambda_{\mathcal{R}_U}| + |\Lambda_{\mathcal{D}}\rangle\rangle\langle\langle \Lambda_{\mathcal{D}}| - |\Lambda_{\mathcal{D}}\rangle\rangle\langle\langle \Lambda_{\mathcal{R}_U}| \quad (4.64)$$

$$\Lambda_{\mathcal{W}_U} = \Lambda_{\mathcal{R}_U} \otimes \Lambda_{\mathcal{I}} + (\Lambda_{\mathcal{D}} - \Lambda_{\mathcal{R}_U}) \otimes \Lambda_{\mathcal{D}}. \quad (4.65)$$

□

Proof of Proposition 4.5.2. We start with the case for $k = 1$. Let $\mathcal{E} \in T(\mathcal{X})$ be a channel and let $\mathcal{W} \in T(\mathcal{X} \otimes \mathcal{X})$ be the twirling superchannel. Then we have

$$|\Lambda_{\mathcal{W}_U(\mathcal{E})}\rangle\rangle = \mathcal{S}_{\mathcal{W}_U} |\Lambda_{\mathcal{E}}\rangle\rangle \quad (4.66)$$

$$= |\Lambda_{\mathcal{I}}\rangle\rangle\langle\langle \Lambda_{\mathcal{R}_U}| \Lambda_{\mathcal{E}}\rangle\rangle + |\Lambda_{\mathcal{D}}\rangle\rangle\langle\langle (\langle\langle \Lambda_{\mathcal{D}} | \Lambda_{\mathcal{E}}\rangle\rangle - \langle\langle \Lambda_{\mathcal{R}_U} | \Lambda_{\mathcal{E}}\rangle\rangle) \quad (4.67)$$

Now using

$$\langle\langle \Lambda_{\mathcal{D}} | \Lambda_{\mathcal{E}}\rangle\rangle = \frac{1}{d} \text{Tr}[\Lambda_{\mathcal{E}}] = p \quad (4.68)$$

$$\langle\langle \Lambda_{\mathcal{U}} | \Lambda_{\mathcal{E}}\rangle\rangle = \langle\langle \mathcal{S}_{\mathcal{U}} | \mathcal{S}_{\mathcal{E}}\rangle\rangle = \text{Tr}[\mathcal{S}_{\mathcal{U}}^\dagger \mathcal{S}_{\mathcal{E}}] \quad (4.69)$$

we have

$$\langle\langle \Lambda_{\mathcal{R}_U} | \Lambda_{\mathcal{E}}\rangle\rangle = \frac{\text{Tr}[\mathcal{S}_{\mathcal{U}}^\dagger \mathcal{S}_{\mathcal{E}}] - p}{d^2 - 1} = \lambda \quad (4.70)$$

and hence $\mathcal{W}_U(\mathcal{E}) = \lambda \mathcal{I} + (p - \lambda) \mathcal{D}$

Now to prove Eq. (4.50) we proceed by induction. Assume that $\mathcal{W}_U(\mathcal{E})^n = \lambda^n \mathcal{I} + (p^n - \lambda^n) \mathcal{D}$. Noting that for channels \mathcal{I} and \mathcal{D} we have

$$\mathcal{S}_I \mathcal{S}_I = \mathcal{S}_I \tag{4.71}$$

$$\mathcal{S}_D \mathcal{S}_D = \mathcal{S}_D \mathcal{S}_I = \mathcal{S}_I \mathcal{S}_D = \mathcal{S}_D \tag{4.72}$$

and hence

$$\mathcal{W}(\mathcal{E})^{n+1} = \mathcal{W}(\mathcal{E})\mathcal{W}(\mathcal{E})^n \tag{4.73}$$

$$= (\lambda \mathcal{I} + (p - \lambda) \mathcal{D})(\lambda^n \mathcal{I} + (p^n - \lambda^n) \mathcal{D}) \tag{4.74}$$

$$= \lambda^{n+1} \mathcal{I} + [(p - \lambda)(p^n - \lambda^n) + \lambda(p^n - \lambda^n) + (p - \lambda)\lambda^n] \mathcal{D} \tag{4.75}$$

$$= \lambda^{n+1} \mathcal{I} + [p(p^n - \lambda^n) + p\lambda^n - \lambda^{n+1}] \mathcal{D} \tag{4.76}$$

$$= \lambda^{n+1} \mathcal{I} + (p^{n+1} - \lambda^{n+1}) \mathcal{D}. \tag{4.77}$$

and hence is true for $k = n + 1$. Thus by induction this is true for all positive integers k . \square

4.5.1 Subsystem Twirling Superchannel

Treating the twirling operation as a superchannel naturally allows us to consider the case where we only twirl a subsystem, or several subsystems, of a channel acting on a composite system. This is closely related to the proposal for *symmetrized characterization* [ESM⁺07], and can be readily generalized to subsystem randomized benchmarking, which has been discussed in the case of a bipartite system [GCM⁺12].

Definition 4.5.1 (Subsystem Twirling). Consider a bipartite Hilbert space $\mathcal{L}(\mathcal{X}_1 \otimes \mathcal{X}_2)$ and channels $\mathcal{E}_1 \in T(\mathcal{X}_1), \mathcal{E}_2 \in T(\mathcal{X}_2)$. We define the subsystem twirling superchannel on the first subsystem to be the map $\mathcal{W}_{1,U}$ satisfying

$$\mathcal{W}_U^{(1)}(\mathcal{E}_1 \otimes \mathcal{E}_2) = \mathcal{W}_U(\mathcal{E}_1) \otimes \mathcal{E}_2 \tag{4.78}$$

for all $\mathcal{E}_1 \in T(\mathcal{X}_1), \mathcal{E}_2 \in T(\mathcal{X}_2)$.

From this definition it is straightforward to generalize to subtwirling any number of subsystems of a multipartite channel. To compute the action of $\mathcal{W}_U^{(1)}$ on a channel $\mathcal{E} \in T(\mathcal{X}_1 \otimes \mathcal{X}_2)$ that is not necessarily a tensor product we prove the following result.

Proposition 4.5.3 (Subsystem Twirling). *Consider a bipartite Hilbert space $\mathcal{L}(\mathcal{X}_{12})$, and a channel $\mathcal{E} \in T(\mathcal{X}_{12})$ where $\mathcal{X}_{12} = \mathcal{X}_1 \otimes \mathcal{X}_2$. The action of twirling the subsystem channel $T(\mathcal{X}_1)$ with $\mathcal{W}_{\mathcal{U}}^{(1)}$ is given by*

$$\mathcal{W}_{\mathcal{U}}^{(1)}(\mathcal{E}) = \mathcal{I} \otimes \mathcal{R}_{\mathcal{U}}^{(1)}(\mathcal{E}) + \mathcal{D} \otimes (\mathcal{Q}_{\mathcal{D}}^{(1)}(\mathcal{E}) - \mathcal{R}_{\mathcal{U}}^{(1)}(\mathcal{E})) \quad (4.79)$$

and for a sequence of k copies of \mathcal{E} we have

$$\mathcal{W}_{\mathcal{U}}^{(1)}(\mathcal{E})^k = \mathcal{I} \otimes \mathcal{R}_{\mathcal{U}}^{(1)}(\mathcal{E})^k + \mathcal{D} \otimes (\mathcal{Q}_{\mathcal{D}}^{(1)}(\mathcal{E})^k - \mathcal{R}_{\mathcal{U}}^{(1)}(\mathcal{E})^k) \quad (4.80)$$

where we have defined superchannels

$$\mathcal{Q}_{\mathcal{F}}^{(1)}, \mathcal{R}_{\mathcal{F}}^{(1)} \in T(\mathcal{X}_{12} \otimes \mathcal{X}_{12}, \mathcal{X}_2 \otimes \mathcal{X}_2) \quad (4.81)$$

with

$$\mathcal{R}_{\mathcal{F}}^{(1)} = \frac{1}{d^2 - 1} (\mathcal{Q}_{\mathcal{F}}^{(1)} - \mathcal{Q}_{\mathcal{D}}^{(1)}) \quad (4.82)$$

and $\mathcal{Q}_{\mathcal{F}}^{(1)}$ is defined by the superoperator

$$\mathcal{S}_{\mathcal{Q}_{\mathcal{F}}^{(1)}} = (\langle\langle \Lambda_{\mathcal{F}} | \otimes \mathcal{S}_I \rangle\rangle \mathcal{V}_2) = \begin{array}{c} \text{Diagram: A triangle labeled } \Lambda_{\mathcal{F}} \text{ with three input lines on the right and two output lines on the left. The two output lines cross each other. Below the triangle are three horizontal lines representing the input to the superoperator } \mathcal{V}_2. \end{array} \quad (4.83)$$

where $\mathcal{F} \in T(\mathcal{X}_1)$.

The operators $\mathcal{Q}_{\mathcal{F}}^{(1)}(\mathcal{E})$ and $\mathcal{R}_{\mathcal{F}}^{(1)}(\mathcal{E})$ are reduced channels in $T(\mathcal{X}_2)$. We see here we can recover the twirling superchannel case by setting $\mathcal{X}_2 \cong \mathcal{C}$ to be a trivial subsystem. In this case the reduced channels are numbers: $\mathcal{Q}_{\mathcal{F}}^{(1)}(\mathcal{E}) = \langle\langle \Lambda_{\mathcal{F}} | \Lambda_{\mathcal{E}} \rangle\rangle$, and in particular $\mathcal{Q}_{\mathcal{D}}^{(1)}(\mathcal{E}) = p$ and $\mathcal{R}_{\mathcal{U}}^{(1)}(\mathcal{E}) = \lambda$ in [Proposition 4.5.1](#).

We can twirl multiple subsystems by composing the subsystem twirling superchannel. For example, consider a tripartite subsystem where we wish to twirl subsystems 1 and 2 separately. Then we have

$$\mathcal{W}_{\mathcal{U}_2}^{(2)} \mathcal{W}_{\mathcal{U}_1}^{(1)}(\mathcal{E}_1 \otimes \mathcal{E}_2 \otimes \mathcal{E}_3) = \mathcal{W}_{\mathcal{U}_1}(\mathcal{E}_1) \otimes \mathcal{W}_{\mathcal{U}_2}(\mathcal{E}_2) \otimes \mathcal{E}_3. \quad (4.84)$$

We will now prove the general expression for twirling n -subsystems, where we allow the possibility of an $(n + 1)^{th}$ subsystem which is not twirled.

Proposition 4.5.4 (*n*-Twirling). Consider a bipartite Hilbert space $\mathcal{L}(\mathcal{X})$ where $\mathcal{X} = \bigotimes_{j=1}^{n+1} \mathcal{X}_j$ and a channel $\mathcal{E} \in T(\mathcal{X})$. The action of individually twirling the first n subsystems channel $T(\mathcal{X}_j)$ with $\mathcal{W}_U^{\{n\}} = \mathcal{W}_{U_1}^{(1)} \dots \mathcal{W}_{U_n}^{(n)}$ is given by

$$\mathcal{W}_U^{\{n\}}(\mathcal{E}) = \sum_b \mathcal{O}_b \otimes \left(\sum_{b' \leq b} (-1)^{w(b)+w(b')} \mathcal{G}_{b'}(\mathcal{E}) \right) \quad (4.85)$$

where $b = b_1, \dots, b_n$ and $b' = b'_1, \dots, b'_n$ are bit-strings of length n ($b_i, b'_i \in \{0, 1\}$), $w(b) = \sum_{j=1}^n b_j$ is the Hamming weight of b , and $b' \leq b \Leftrightarrow b'_j \leq b_j \forall j$, and we have defined channels

$$\mathcal{O}_b = \bigotimes_{j=1}^N (\delta_{b_j,0} \mathcal{I}^{(j)} + \delta_{b_j,1} \mathcal{D}^{(j)}) \quad (4.86)$$

$$\mathcal{G}_b = \bigotimes_{j=1}^N (\delta_{b_j,0} \mathcal{R}^{(j)} + \delta_{b_j,1} \mathcal{Q}_D^{(j)}). \quad (4.87)$$

For *n*-Twirling each of a sequence of k copies of \mathcal{E} we have

$$\mathcal{W}_U^{\{n\}}(\mathcal{E})^k = \sum_b \mathcal{O}_b \otimes \left(\sum_{b' \leq b} (-1)^{w(b)+w(b')} \mathcal{G}_{b'}(\mathcal{E})^k \right) \quad (4.88)$$

Proof of Proposition 4.5.3. The case for $k = 1$ in Eq. (4.79) follows from the definition of the twirling superchannel \mathcal{S}_W in Proposition 4.5.1 and using the composite system superoperator $\mathcal{S}_{W \otimes \mathcal{I}}$. To prove the expression for arbitrary k in Eq. (4.80), we use proof by induction as for Proposition 4.5.1, with the only difference been that now the quantities are channels rather than numbers, and hence do not commute. We have

$$\begin{aligned} \mathcal{W}_U^{(1)}(\mathcal{E})^{k+1} &= \mathcal{W}_U^{(1)}(\mathcal{E})^k \mathcal{W}_U^{(1)}(\mathcal{E}) \\ &= \left[\mathcal{I} \otimes \mathcal{R}_U^{(1)}(\mathcal{E})^k + \mathcal{D} \otimes (\mathcal{Q}_D^{(1)}(\mathcal{E})^k - \mathcal{R}_U^{(1)}(\mathcal{E})^k) \right] \\ &\quad \times \left[\mathcal{I} \otimes \mathcal{R}_U^{(1)}(\mathcal{E}) + \mathcal{D} \otimes (\mathcal{Q}_D^{(1)}(\mathcal{E}) - \mathcal{R}_U^{(1)}(\mathcal{E})) \right] \\ &= \mathcal{I} \otimes \mathcal{R}_U^{(1)}(\mathcal{E})^{k+1} + \mathcal{D} \otimes \left(\mathcal{R}_U^{(1)}(\mathcal{E})^k \mathcal{Q}_D^{(1)}(\mathcal{E}) - \mathcal{R}_U^{(1)}(\mathcal{E})^{k+1} \right) \\ &\quad + \mathcal{D} \otimes \left(\mathcal{Q}_D^{(1)}(\mathcal{E})^k \mathcal{R}_U^{(1)}(\mathcal{E}) - \mathcal{R}_U^{(1)}(\mathcal{E})^{k+1} \right) \\ &\quad + \mathcal{D} \otimes \left(\mathcal{Q}_D^{(1)}(\mathcal{E})^{k+1} - \mathcal{R}_U^{(1)}(\mathcal{E})^k \mathcal{Q}_D^{(1)}(\mathcal{E}) - \mathcal{Q}_D^{(1)}(\mathcal{E})^k \mathcal{R}_U^{(1)}(\mathcal{E}) + \mathcal{R}_U^{(1)}(\mathcal{E})^{k+1} \right) \\ &= \mathcal{I} \otimes \mathcal{R}_U^{(1)}(\mathcal{E})^{k+1} + \mathcal{D} \otimes \left(\mathcal{Q}_D^{(1)}(\mathcal{E})^{k+1} - \mathcal{R}_U^{(1)}(\mathcal{E})^{k+1} \right) \end{aligned} \quad (4.89)$$

So Eq. (4.80) is true for $k + 1$ if true for k . Hence by induction it is true for all positive integer k . \square

Proof of Proposition 4.5.4. The proof of this expression can be seen by construction. Each twirled subsystem will have a tensor product of an identity \mathcal{I} or a depolarizing channel \mathcal{D} with the corresponding contraction superoperator, \mathcal{R}_U or $(\mathcal{Q}_D - \mathcal{R}_U)$ respectively, which acts on the input channel.

We then have all permutations of \mathcal{I} and \mathcal{D} on the twirled subsystems which are labeled by the bitstring b . For a given bit string we can expand out the right hand side of the tensor product of the corresponding operators. For two subsystems we have

$$b = 00 \Rightarrow \mathcal{G}_{00} = \mathcal{R}_U^{(1)} \mathcal{R}_U^{(2)} \quad (4.90)$$

$$b = 01 \Rightarrow \mathcal{G}_{01} = \mathcal{R}_U^{(1)} (\mathcal{Q}_D^{(2)} - \mathcal{R}_U^{(2)}) = \mathcal{R}_U^{(1)} \mathcal{Q}_D^{(2)} - \mathcal{R}_U^{(1)} \mathcal{R}_U^{(2)} \quad (4.91)$$

$$b = 10 \Rightarrow \mathcal{G}_{01} = (\mathcal{Q}_D^{(1)} - \mathcal{R}_U^{(1)}) \mathcal{R}_U^{(2)} = \mathcal{Q}_D^{(1)} \mathcal{R}_U^{(2)} - \mathcal{R}_U^{(1)} \mathcal{R}_U^{(2)} \quad (4.92)$$

$$\begin{aligned} b = 11 \Rightarrow \mathcal{G}_{11} &= (\mathcal{Q}_D^{(1)} - \mathcal{R}_U^{(1)}) (\mathcal{Q}_D^{(2)} - \mathcal{R}_U^{(2)}) \\ &= \mathcal{Q}_D^{(1)} \mathcal{Q}_D^{(2)} - \mathcal{R}_U^{(1)} \mathcal{Q}_D^{(2)} - \mathcal{Q}_D^{(1)} \mathcal{R}_U^{(2)} + \mathcal{R}_U^{(1)} \mathcal{R}_U^{(2)} \end{aligned} \quad (4.93)$$

It is easy to see that this gives the expression in Eq. (4.85). The expression in Eq. (4.88) then follows by induction as with the proof for Proposition 4.5.3, noting that the terms in bit strings 01, 10 have alternating signs to the corresponding terms in 11. \square

4.6 Summary

In this chapter we have introduced the superchannel formalism and given explicit definitions and methods for constructing and composing general superchannels that can then be applied to specific problems of interest. To demonstrate this we gave the example of the twirling and n -twirling superchannels. We have also proved conditions for the complete positivity, trace preserving and Hermiticity preserving properties of these constructed superchannels in terms of the underlying physical channels and states used in the construction. The superchannel formalism generalizes quantum channels to describe a strictly greater set of dynamics by treating channels themselves as the input. As we will describe in Chapter 5 a superchannel allows for an operational description of the dynamics of a system initially correlated with its environment, without having to resort to trying to justify the physicality of a non-CP description of the dynamics. In principle generalizations of this approach could be used for an operational description of certain types of non-Markovian

interactions, in particular where the environment has a long correlation time. The outline of this approach would be that we can consider a superchannel with two inputs, channels \mathcal{E}_1 and \mathcal{E}_2 , where there is some time τ between these gates in the underlying circuit. By characterizing the superchannel one could look to quantify the memory of the environment between these operations. In particular if the inputs are preparation procedures as required for superchannel tomography, for an uncorrelated environment we would expect the outcome after the second preparation to be independent of the first preparation. If this is not the case then we have an indication of a non-Markovian environmental correlations.

In regards to characterization of superchannels, an open question is what sort of dynamics can be characterized by unitary input channels. In regards to doing standard quantum process tomography with interleaved randomized benchmarking, the set of unitary channels only spans the unital subspace of the channel to be reconstructed. In terms of a superchannel where the input and output states of the underlying physical system are set to be ρ_i and ρ_j , this corresponds to characterizing the unital part of the Choi-matrix $\Lambda_{i,j} = \mathcal{F}_1(\rho_i) \otimes \mathcal{F}_2^\dagger(\rho_j)$. If we assume that \mathcal{F}_1 is TP, then the action of this map on an identity matrix outputs the state $\mathcal{F}_2^\dagger(\rho_j)$ which is in general not the identity. Thus we see that superchannels are generally highly non-unital the non-unital subspace is likely contains the most important parameters to be characterized. A final comment is that in general quantum process tomography is unscalable for large systems as the number of experimental configurations required to completely sample all parameters of an arbitrary channel is d^4 for a d dimensional system. Superchannel tomography only makes matters worse due to the increase in dimension of the object being reconstructed, nevertheless it is important to consider when in principle it is possible and is an important first step to considering partial characterization schemes.

Chapter 5

Characterization of Initially Correlated Open Quantum Systems

5.1 Introduction

The quantum channel formalism reviewed in [Chapter 2](#) provides a complete operational description of reduced dynamics of an open quantum system in the case where the initial system-environment state is uncorrelated [[PB02](#)]. This was made explicit in the formulation of the system-environment representation of CP-maps in [Section 2.3.2](#). However in many situations this central assumption is at best an approximation [[RRMK⁺08](#), [MS10](#)]. In this chapter we demonstrate how one may describe the evolution of an open quantum system initially correlated with its environment using the quantum superchannel formalism introduced in [Chapter 4](#).

A typical quantum experiment can be split into three steps: state preparation, evolution, and measurement. State preparation takes a system from a generally unknown initial state to a desired input state. This state is then subjected to some dynamical process for a fixed time—the state evolution—the output of which may be finally measured. If the initial system state is correlated with the environment, the system preparation procedure in general will also affect the state of the environment. We illustrate this with a simple example: Consider the extreme case of a maximally entangled initial SE state:

$$|\psi_{SE}\rangle = \frac{1}{\sqrt{2}}(|0\rangle \otimes |0\rangle + |1\rangle \otimes |1\rangle). \quad (5.1)$$

where the first and second subsystems correspond to the system and environment respectively. If the experimenter attempts to prepare an input state by an ideal projec-

tive measured which leaves the system in the eigenstate corresponding to the observed output, then a projective preparation of the system into $|0\rangle$ or $|1\rangle$ for example leaves the environment in orthogonal states. Hence, if the subsequent system evolution is not perfectly isolated it is coupled to different environment states which may lead to drastically different reduced dynamics of the system conditional on the used preparation procedure [Woo08, BGTW11, Mod11]. If the experimenter attempts to characterize this channel using standard tomographic characterization techniques, the reconstructed description of the reduced system dynamics may appear unphysical and be given by a non-CP map [KMRRS07, Zim06, CTZ08, Woo08, MS10, BGTW11, Mod12]. This highlights the importance of accounting for initial SE correlations to reliably characterize the system dynamics.

While the environment is typically inaccessible to the experimenter, recent results suggest that at least partial information about the initial joint SE state can be extracted from measurements of the system alone. Initial correlations can be witnessed through the distinguishability [WLB13, LPB10, RRMAG12, GB11] and purity [RWC+11] of quantum states, which has also been explored experimentally [SBC+11, LTLG11, GRP+13]. A more operationally complete characterization can be obtained by explicitly treating the system's preparation procedure, rather than the prepared state, as the input to the reduced description as was proposed in [Mod12]. This approach fits naturally in the quantum superchannel formalism presented in Chapter 4, where the preparation procedure is the input of an effective superchannel that takes into account the initial system environment state, and subsequent evolution. Hence it captures not just the system evolution, but also the dynamical influence of the environment, even in the presence of initial SE correlations.

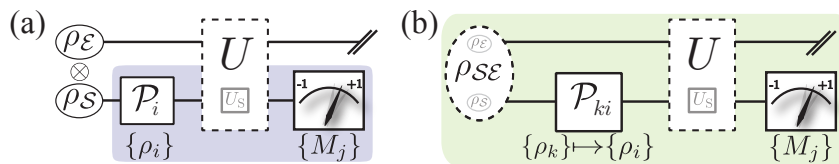


Figure 5.1: System dynamics in the presence of an environment. (a) Special case: with no initial SE correlations the reduced dynamics of the system, which interacts unitarily (U) with an environment, can be completely reconstructed from tomographically complete sets of input states $\{\rho_i\}$ (resulting from preparation procedures $\{\mathcal{P}_i\}$), and measurements $\{M_j\}$. (b) General case: the joint SE state may be initially correlated before the state preparation procedure. The superchannel approach encompasses this situation by treating the preparation procedure \mathcal{P}_{ki} as the input state to a more general description of the reduced system dynamics.

The main differences between standard QPT and the superchannel description of the tomography experiment is illustrated in Fig. 5.1. Recall from Section 3.5 that the Choi matrix of an unknown quantum process can be reconstructed through QPT from the outcomes of a finite set of measurements $\{M_j\}_{j=1}^{d^2}$, performed on a finite set of system input states $\{\rho_i\}_{i=1}^{d^2}$ as shown in Fig. 5.1(a). Crucially, this assumes that the channel \mathcal{E} being characterized is independent of the system's preparation. In the presence of initial SE correlations, this assumption is in general not satisfied. The joint SE state is then $\rho_{SE} \in L(\mathcal{X} \otimes \mathcal{Y})$, as illustrated in Fig. 5.1(b), where \mathcal{X} and \mathcal{Y} are the state spaces of the system and environment, respectively. In the first step of the experiment the system is prepared in the state ρ_s by applying a preparation map \mathcal{P}_i to S alone. Such a preparation will typically leave the environment in a state conditional on \mathcal{P}

$$\rho_{E|\mathcal{P}} = \text{Tr}_{\mathcal{X}}[\mathcal{P}(\rho_{SE})], \quad (5.2)$$

which in turn leads to a conditional evolution $\mathcal{E}_{\mathcal{P}}$ and QPT would return a map which is a combination of the partial reconstructions of the possible $\mathcal{E}_{\mathcal{P}}$. In the following we consider the case of a fully de-correlating preparation procedure:

$$(\mathcal{P} \otimes \mathcal{I})(\rho_{SE}) = \rho_s \otimes \rho_{E|\mathcal{P}} \quad (5.3)$$

where \mathcal{I} is the identity map on E . Denoting by U the channel that describes the subsequent joint evolution, the final output state is given by

$$\rho'_s = \text{Tr}_E [\mathcal{U}((\mathcal{P} \otimes \mathcal{I})(\rho_{SE}))]. \quad (5.4)$$

To characterize the system in the presence of possible initial correlations we describe the dynamics by means of a superchannel

$$\mathcal{M}: \mathcal{P} \rightarrow \rho'_s. \quad (5.5)$$

The output is given by

$$\rho'_s = \mathcal{M}(\Lambda_{\mathcal{P}}) \quad (5.6)$$

where $\Lambda_{\mathcal{P}}$ is the Choi matrix for the preparation map \mathcal{P} [Mod12].

In Section 5.2 we explicitly show how to construct the IC superchannel and discuss its properties, thus incorporating [Mod12] within the general framework of superchannels introduced in Chapter 4. In Section 5.3 we introduce operational measures to quantify the presence and strength of initial correlations. These measures may be computed from the tomographic reconstruction of an unknown superchannel. We explore these measures

using several theoretical examples in [Section 5.4](#). In particular we consider both classically correlated, and entangled initial SE states, and interactions consisting of controlled-unitary dynamics, and a swap gate, which can be used to model several common error modalities for channels. Finally in [Section 5.5](#) we present the results of an experimental demonstration of these characterization techniques performed using photonic qubits. This experiment characterizes the dynamics of a single-photonic qubit that is initially correlated with a simulated single-photon environment.

5.2 Initial Correlation Quantum Superchannel

Consider a system with Hilbert space \mathcal{X}_1 , and an environment with Hilbert space \mathcal{Y}_1 . Let the system and environment initially be in a $\rho_{SE} \in D(\mathcal{X}_1 \otimes \mathcal{Y}_1)$. Consider the case where we first apply to this state a preparation procedure

$$\mathcal{P} = \mathcal{P}_s \otimes \mathcal{I}_E \in T(\mathcal{X}_1 \otimes \mathcal{Y}_1, \mathcal{X}_2, \otimes \mathcal{Y}_2) \quad (5.7)$$

where $\mathcal{Y}_2 = \mathcal{Y}_1$ and $\mathcal{P}_s \in T(\mathcal{X}_1, \mathcal{X}_2)$ acts only on the system to prepare it in a desired input state. This is followed by coupled evolution of the joint system-environment state, described by a CPTP map $\mathcal{U} \in T(\mathcal{X}_2 \otimes \mathcal{Y}_2, \mathcal{X}_3 \otimes \mathcal{Y}_3)$ as depicted in [Fig. 5.1](#). In the typical system-environment formalism of CP maps the channel \mathcal{U} will be a unitary channel on the joint system-environment, however this property isn't necessary in our derivation, and for now we only assume that \mathcal{U} is a general CPTP-map. The reduced output of the joint evolution is given by

$$\begin{aligned} \rho'_s &= \text{Tr}_{\mathcal{Y}_3} [\mathcal{U}(\mathcal{P}_s(\rho_{SE}))] \\ &= \text{Tr}_{\mathcal{X}_2, \mathcal{Y}_2, \mathcal{Y}_3} \left[(\mathcal{P}(\rho_{SE})^T \otimes \mathbb{1}_{se}) \Lambda_{\mathcal{U}} \right] \\ &= \text{Tr}_{\mathcal{X}_2, \mathcal{Y}_2, \mathcal{Y}_3} \left[\left(\text{Tr}_{\mathcal{X}_1, \mathcal{Y}_1} [(\rho_{SE}^T \otimes \mathbb{1}_{SE}) \Lambda_{\mathcal{P}}]^T \otimes \mathbb{1}_{SE} \right) \Lambda_{\mathcal{U}} \right] \end{aligned} \quad (5.8)$$

Note that $\Lambda_{\mathcal{U}}$ and $\Lambda_{\mathcal{P}}$ have each 4 subsystem indices, which correspond to S -input, E -input, S -output and E -output, respectively.

The action of the preparation channel may be represented graphically as:

$$\mathcal{P}(\rho_{SE}) = \begin{array}{c} \text{---} \text{---} \text{---} \text{---} \\ \text{---} \text{---} \text{---} \text{---} \\ \text{---} \text{---} \text{---} \text{---} \\ \text{---} \text{---} \text{---} \text{---} \end{array} \begin{array}{c} \boxed{\rho_{SE}} \\ \boxed{\Lambda_{\mathcal{P}}} \end{array} \begin{array}{c} \text{---} \text{---} \text{---} \text{---} \\ \text{---} \text{---} \text{---} \text{---} \\ \text{---} \text{---} \text{---} \text{---} \\ \text{---} \text{---} \text{---} \text{---} \end{array} \quad (5.9)$$

Hence the graphical representation of Eq. (5.8) is given by

$$\rho'_s = \mathcal{P}(\rho_{\text{se}}) \Lambda_{\mathcal{U}} = \rho_{\text{se}} \Lambda_{\mathcal{P}_s} \Lambda_{\mathcal{U}} = \Lambda_{\mathcal{P}_s} \rho_{\text{se}} \Lambda_{\mathcal{U}} \quad (5.10)$$

Hence we may define the IC quantum superchannel $\mathcal{M}_{\text{IC}} \in T(\mathcal{X}_1 \otimes \mathcal{X}_2, \mathcal{X}_3)$ in terms of the initial system-environment state ρ_{SE} and the interaction Choi matrix $\Lambda_{\mathcal{U}}$ as the tensor network:

$$\Lambda_{\mathcal{M}_{\text{IC}}} = \rho_{\text{se}} \Lambda_{\mathcal{U}} \quad (5.11)$$

so that

$$\rho'_s = \Lambda_{\mathcal{P}_s} \Lambda_{\mathcal{M}_{\text{IC}}} \quad (5.12)$$

Thus the IC superchannel $\mathcal{M}_{\text{IC}} \in T(\mathcal{X}_1 \otimes \mathcal{X}_2, \mathcal{X}_3)$ takes the system-preparation procedure $\Lambda_{\mathcal{P}_s} \in L(\mathcal{X}_1 \otimes \mathcal{X}_2)$ as an input and produces an output quantum state $\rho' \in L(\mathcal{X}_3)$ given by

$$\begin{aligned} \rho' &= \mathcal{M}_{\text{IC}}(\Lambda_{\mathcal{P}_s}) \\ &= \text{Tr}_{\mathcal{X}_1, \mathcal{X}_2}[(\Lambda_{\mathcal{P}_s} \otimes \mathbb{1}_{\mathcal{X}_3})\Lambda_{\mathcal{U}}]. \end{aligned}$$

We note that the IC superchannel is an example of input post-selected superchannel form [Definition 4.3.3](#).

In terms of index contractions we have

$$\begin{aligned}
(\rho_{\text{SE}})_{i_1 i_2; j_1 j_2} &= \langle i_1 i_2 | \rho_{\text{SE}} | j_1 j_2 \rangle \\
(\Lambda_{\mathcal{U}})_{i_1 i_2 i_3 i_4; j_1 j_2 j_3 j_4} &= \langle i_1 i_2 i_3 i_4 | \Lambda_{\mathcal{U}} | j_1 j_2 j_3 j_4 \rangle \\
(\Lambda_{\mathcal{P}_s})_{i_1 i_2; j_1 j_2} &= \langle i_1 i_2 | \Lambda_{\mathcal{P}_s} | j_1 j_2 \rangle \\
(\Lambda_{\mathcal{I}_e})_{i_1 i_2; j_1 j_2} &= \langle i_1 i_2 | \Lambda_{\mathcal{I}_e} | j_1 j_2 \rangle \\
&= \delta_{i_1 i_2} \delta_{j_1 j_2}.
\end{aligned}$$

Hence we can write the action of the preparation map in Eq. (5.9) as

$$\mathcal{P}(\rho_{\text{SE}})_{i_1 i_2; j_1 j_2} = \sum_{n, m} (\rho_{\text{SE}})_{ni_2; mj_2} (\Lambda_{\mathcal{P}_s})_{ni_1; mj_1}, \quad (5.13)$$

and $\Lambda_{\mathcal{M}_{\text{IC}}}$ in Eq. (5.11) is given by

$$(\Lambda_{\mathcal{M}_{\text{IC}}})_{i_1 i_2 i_3; j_1 j_2 j_3} = \sum_{n, m, l} (\rho_{\text{SE}})_{i_1 n; j_1 m} (\Lambda_{\mathcal{U}})_{i_1 n i_3 l; j_1 m j_3 l}. \quad (5.14)$$

In the absence of initial correlations ($\rho_{\text{se}} = \rho_s \otimes \rho_e$) the Choi-matrix for \mathcal{M}_{IC} takes the form

$$\Lambda_{\mathcal{M}_{\text{IC}}} = \rho_s \otimes \Lambda_{\mathcal{E}} \quad (5.15)$$

where $\Lambda_{\mathcal{E}}$ is the Choi-matrix corresponding to the sys-env representation of the joint unitary \mathcal{U} with the environment initially in state ρ_e . In this case the action on a preparation procedure \mathcal{P}_s is given by

$$\mathcal{M}_{\text{IC}}(\mathcal{P}_s) = \mathcal{E}(\mathcal{P}_s(\rho_s)). \quad (5.16)$$

5.2.1 Properties of the IC Superchannel

We may deduce the properties of an IC superchannel \mathcal{M}_{IC} from the properties of superchannels derived in Section 4.3. The first point we make is that \mathcal{M}_{IC} is a CP-map whenever the SE interaction \mathcal{U} is a CP map, regardless of initial SE correlation. This is an important consideration for process tomography experiments as it means that even in the presence of initial correlations we may use constrained MLE methods in good faith the closest fit should be a CP map. The second point we note is that since \mathcal{M}_{IC} is an input-post selected superchannel it is not normalized with respect to the standard convention for a quantum channel. In particular

$$\text{Tr}[\Lambda_{\mathcal{M}_{\text{IC}}}] = d \quad (5.17)$$

where $d = \dim(\mathcal{X}_1)$ is the dimension of the system input Hilbert space. Since the input channel Hilbert space for \mathcal{M}_{IC} is d^2 dimensional the superchannel is sub-normalized. The intuition for this is because the IC superchannel by construction takes Choi-matrices as inputs, which have trace d rather than states, this extra factor of d is accounted for elsewhere. This is worth keeping in mind when using definitions of operator and channel norms, in particular the CB trace norm, on \mathcal{M}_{IC} .

Further, regardless of normalization, \mathcal{M}_{IC} is not trace preserving. With the normalization convention used in Eq. (5.17), for \mathcal{M}_{IC} to be TP we require

$$\text{Tr}_{\mathcal{X}_3}[\Lambda_{\mathcal{M}_{\text{IC}}}] = \mathbb{1}_{\mathcal{X}_1} \otimes \mathbb{1}_{\mathcal{X}_2}. \quad (5.18)$$

However, assuming that the SE interaction is TP, we have

$$\text{Tr}_{\mathcal{X}_3}[\Lambda_{\mathcal{M}_{\text{IC}}}] = \text{Tr}_{\mathcal{Y}_1}[\rho_{\text{SE}}] \otimes \mathbb{1}_{\mathcal{X}_2}. \quad (5.19)$$

Hence \mathcal{M}_{IC} can never be TP as $\text{Tr}_{\mathcal{Y}_1}[\rho_{\text{SE}}]$ will have trace equal to 1 not d . The superchannel can be *pseudo-TP* in the sense that it decreases all states by a constant factor $1/d$. This condition holds if and only if

$$\text{Tr}_{\mathcal{Y}_1}[\rho_{\text{SE}}] = \frac{\mathbb{1}_{\mathcal{X}_1}}{d}. \quad (5.20)$$

Two examples of initial state which do satisfy this quantity are if the system and environment are initially in a maximally entangled initial state, or a completely mixed state. In both these cases the reduced state of the system appears maximally mixed, and hence any measurement used as a preparation procedure will succeed with constant probability. In all other cases different preparation procedures would in general lead to different overall count rates.

Finally we comment that by only considering preparation procedures of the form in Eq. (5.7) we assumed that the preparation channel only acts on the system, and hence acts trivially on the environment. We may relax this condition however for the case where the preparation procedure has some constant action on the environment of the form

$$\mathcal{P} = \mathcal{P}_s \otimes \mathcal{E} \quad \text{or} \quad \mathcal{P} = (\mathcal{E}_{se})(\mathcal{P}_s \otimes \mathcal{I}) \quad (5.21)$$

for example, where \mathcal{E} is fixed for all considered preparation procedures \mathcal{P} . In this case we may simply incorporate the action of the the preparation procedure into our joint SE interaction channel \mathcal{U} . However, in the case where the term \mathcal{E} varies with different preparation procedures we have returned to the original problem where the environment where the object we are trying to characterize (in this case the superchannel \mathcal{M}) is conditional on our choice of preparation procedure.

5.2.2 IC Quantum Superchannel Tomography

In order to experimentally characterize an unknown IC superchannel we can use the method of superchannel quantum process tomography detailed in [Section 4.4.1](#). For a projective preparation procedure $\mathcal{P}_{ij} \in T(\mathcal{X}_1, \mathcal{X}_2)$ which consist of an initial projection (or post-selection) onto the state ρ_i followed by a rotation to the state ρ_j and a final measurement of the output being in state ρ_k we have that probability of observing a count is given by $p_{ijk} = \langle\langle \Pi_{ijk} | \Lambda_{\mathcal{M}_{IC}} \rangle\rangle$ where $\Pi_{ijk} \equiv \rho_i \otimes \rho_j^* \otimes \rho_k$. From here one may then reconstruct the Choi-matrix using any of the methods described in [Section 3.4](#).

Since the IC superchannel \mathcal{M}_{IC} will always be CP-map under the assumptions that the state preparation channel only acts on the system, we can use constrained MLE estimation methods in good faith to ensure a CP fit. This is in contrast to the conventional QPT reconstruction of an initially correlated channel which may return a genuine non-CP result. Hence constrained MLE methods in the standard tomographic protocol cannot distinguish between a non-CP result due to initial correlations, and those from statistical error, and in the former case will return a CP description that may greatly differ from the true description of the reduced dynamics.

We note that if an experimental implementation does not have access to a projective preparation procedure, the set of unitary channels only spans a $d^4 - 2d^2 + 2$ dimensional subspace, and can only characterize the unital part of \mathcal{M}_{IC} . However, in the case of the IC superchannel the unital subspace is not suitable for characterizing initial correlations, as in the absence of initial correlations \mathcal{M}_{IC} is highly non-unital.

5.3 Initial Correlation Measures

The IC superchannel \mathcal{M}_{IC} contains information about initial SE correlations that are visible through their effect on the subsequent experiment, and how different preparation procedures may influence the resulting system dynamics. In this section we introduce two quantitative measures to extract this information: the *Initial Correlation norm* (IC norm) which quantifies the distinguishability of a IC superchannel from one without initial correlations, and *preparation fidelity* which quantifies the quality of a given preparation procedure for implementing a desired quantum channel and can be used to assess the influence of state preparation on the the reduced system dynamics.

5.3.1 Initial Correlation Norm

For an IC superchannel $\mathcal{M}_{\text{IC}} \in T(\mathcal{X} \otimes \mathcal{X}, \mathcal{X})$ where $\dim(\mathcal{X}) = d$, we can define an *average initial system state*

$$\rho_{s,\text{av}} = \frac{1}{d} \text{Tr}_{23}[\Lambda_{\mathcal{M}_{\text{IC}}}] \quad (5.22)$$

and an *average effective map* for the evolution of the system as

$$\Lambda_{\mathcal{E}_{\text{av}}} = \text{Tr}_1[\Lambda_{\mathcal{M}_{\text{IC}}}] \quad (5.23)$$

From Eq. (5.11) we see that for a product initial state ($\rho_{\text{se}} = \rho_s \otimes \rho_e$) the superchannel \mathcal{M}_{IC} is given by the tensor product

$$\Lambda_{\mathcal{M}_{\text{IC}}} = \rho_s \otimes \Lambda_{\mathcal{E}} \quad (5.24)$$

In this case $\rho_{s,\text{av}} = \rho_s$, and $\Lambda_{\mathcal{E}_{\text{av}}} = \Lambda_{\mathcal{E}}$ is the Choi matrix of the channel \mathcal{E} describing the open evolution of the system alone. In this case the channel \mathcal{E} is the one that would be reconstructed from conventional QPT.

For an IC superchannel \mathcal{M}_{IC} we can now define the corresponding separable superchannel \mathcal{M}_{sep} by

Definition 5.3.1 (Average Separable Superchannel). Let $\mathcal{M}_{\text{IC}} \in T(\mathcal{X} \otimes \mathcal{Y}, \mathcal{Y})$ be an IC superchannel. We define the *average separable superchannel* \mathcal{M}_{sep} by

$$\Lambda_{\mathcal{M}_{\text{sep}}} = \rho_{s,\text{av}} \otimes \Lambda_{\mathcal{E}_{\text{av}}} = \frac{1}{d} \text{Tr}_{23}[\Lambda_{\mathcal{M}_{\text{IC}}}] \otimes \text{Tr}_1[\Lambda_{\mathcal{M}_{\text{IC}}}] \quad (5.25)$$

In general $\mathcal{M} \neq \mathcal{M}_{\text{sep}}$ and the distance between \mathcal{M} and \mathcal{M}_{sep} can be used to quantify the strength of the initial *SE* correlations. We thus define the *initial correlation norm*:

Definition 5.3.2 (IC Norm). Let $\mathcal{M}_{\text{IC}} \in T(\mathcal{X} \otimes \mathcal{X}, \mathcal{X})$ be an IC superchannel. The *initial correlation norm* of \mathcal{M}_{IC} is given by

$$\|\mathcal{M}_{\text{IC}}\|_{\text{IC}} = \frac{1}{2} \|\mathcal{M}_{\text{IC}} - \mathcal{M}_{\text{sep}}\|_{\diamond} \quad (5.26)$$

The matrix $\Lambda_{\Delta\mathcal{M}_{\text{IC}}} = \Lambda_{\mathcal{M}_{\text{IC}}} - \Lambda_{\mathcal{M}_{\text{sep}}}$ was introduced as *correlation memory matrix* in [Mod12] since it describes how the dynamics is affected by initial correlations. Using the bounds on the CB-trace norm from Theorem 3.2.1 we have that the IC-norm is upper and lower bounded by

$$\frac{1}{2d^2} \|\Lambda_{\Delta\mathcal{M}_{\text{IC}}}\|_1 \leq \|\mathcal{M}_{\text{IC}}\|_{\text{IC}} \leq \frac{1}{2} \|\text{Tr}_3 |\Lambda_{\Delta\mathcal{M}_{\text{IC}}}| \|_{\infty} \quad (5.27)$$

where $d = \dim(\mathcal{X})$ is the dimension of the input Hilbert space. Note that the IC-norm is not technically a norm acting on the superchannel \mathcal{M}_{IC} , as it is not zero when $\mathcal{M}_{\text{IC}} = 0$. It is actually a norm acting on the difference $\mathcal{M}_{\text{IC}} - \mathcal{M}_{\text{sep}}$, and hence a technically an IC-measure comparing \mathcal{M}_{IC} to \mathcal{M}_{sep} .

Our choice of the CB trace norm allows for an operational interpretation of the IC-norm in terms of channel discrimination as discussed in [Section 3.2.1](#). When $\|\mathcal{M}_{\text{IC}}\|_{\text{IC}} = 0$, there is no operational difference between \mathcal{M}_{IC} and \mathcal{M}_{sep} , which means that there are no observable SE correlations. This can either mean that the initial SE state is indeed uncorrelated, or that the environment is Markovian and initial correlations do not affect the subsequent dynamics. The initial correlation norm thus provides a necessary and sufficient condition for the decoupling of the future state of the system from its past interactions with the environment. When $\|\mathcal{M}_{\text{IC}}\|_{\text{IC}} > 0$ there in principle exists an optimal preparation channel that can be used as a witness for initial correlations, and the specific value of the norm is related to the single shot probability of success for this witness.

For an operational interpretation we first define a renormalized version of the IC-norm, which we call the *IC-value*:

Definition 5.3.3 (IC Value). Let $\mathcal{M}_{\text{IC}} \in T(\mathcal{X} \otimes \mathcal{Y}, \mathcal{Y})$ be an IC superchannel. The *initial correlation value* of \mathcal{M}_{IC} is given by

$$\text{IC}(\mathcal{M}_{\text{IC}}) = \frac{\|\mathcal{M}_{\text{IC}}\|_{\text{IC}}}{\|\rho_{s,\text{av}}\|_{\infty}} = \frac{\|\mathcal{M}_{\text{IC}} - \mathcal{M}_{\text{sep}}\|_{\diamond}}{2\|\rho_{s,\text{av}}\|_{\infty}}. \quad (5.28)$$

Like the IC-norm, the IC-value is 0 if and only if $\mathcal{M}_{\text{IC}} = \mathcal{M}_{\text{sep}}$. Further the IC Value and IC Norm are related by the following bounds

$$\|\mathcal{M}_{\text{IC}}\|_{\text{IC}} \leq \text{IC}(\mathcal{M}_{\text{IC}}) \leq r \|\mathcal{M}_{\text{IC}}\|_{\text{IC}}. \quad (5.29)$$

where r is the rank of $\rho_{s,\text{av}}$. We can bound the probability of successful discriminating an initially correlated channel from the average uncorrelated channel in terms of the IC value by the following theorem:

Theorem 5.3.1 (IC Distinguishability). *Let $\mathcal{M}_{\text{IC}} \in T(\mathcal{X} \otimes \mathcal{X}, \mathcal{X})$ be an IC superchannel. The optimal single shot strategy for distinguishing the IC superchannel from the average separable superchannel \mathcal{M}_{sep} succeeds with probability p which is bounded by*

$$\frac{1}{2} (1 + \text{IC}(\mathcal{M}_{\text{IC}})) \leq p \leq \frac{1}{2} (1 + r \|\mathcal{M}_{\text{IC}}\|_{\text{IC}}). \quad (5.30)$$

where r is the rank of the reduced system state $\rho_{s,\text{av}}$.

The reason for the difference between [Theorem 5.3.1](#) and the usual probabilistic interpretation of the CB trace norm for channel discrimination arises because the IC superchannel is not trace preserving. We now prove this theorem.

Proof of [Theorem 5.3.1](#). From the definition of the CB trace norm we have

$$\|\mathcal{M}_{\text{IC}}\|_{\text{IC}} = \max \left\{ \frac{1}{2} \left\| (\mathcal{I} \otimes (\mathcal{M}_{\text{IC}} - \mathcal{M}_{\text{sep}})) (|V\rangle\langle\langle V|) \right\|_1 : V \in L(X \otimes \mathcal{X}), \|V\|_2 = 1 \right\}. \quad (5.31)$$

If we define operators for the output of the superchannel

$$P_0 = (\mathcal{I} \otimes \mathcal{M}_{\text{IC}}) (|V\rangle\langle\langle V|) \quad (5.32)$$

$$P_1 = (\mathcal{I} \otimes \mathcal{M}_{\text{sep}}) (|V\rangle\langle\langle V|) \quad (5.33)$$

then for a probabilistic interpretation of the IC-norm we require these operators to be renormalized to density matrices $\rho_j = P_j / \text{Tr}[P_j]$ and we may use that the best single shot strategy for distinguishing states ρ_0 and ρ_1 succeeds with probability $p = \frac{1}{2} (1 + \frac{1}{2} \|\rho_0 - \rho_1\|_1)$. Using tensor rearrangements one can show

$$(\mathcal{I} \otimes \mathcal{M}) (|V\rangle\langle\langle V|) = (V^T \otimes \mathbb{1}_{\mathcal{X}}) \Lambda_{\mathcal{M}} (\bar{V} \otimes \mathbb{1}_{\mathcal{X}}). \quad (5.34)$$

$$(5.35)$$

Now, since for the IC superchannel we assume the SE interaction is TP we have $\text{Tr}_3[\Lambda_{\mathcal{M}}] = \rho_{s,\text{av}} \otimes \mathbb{1}$, and hence

$$\text{Tr}[(\mathcal{I} \otimes \mathcal{M}) (|V\rangle\langle\langle V|)] = \text{Tr}[(V^T \otimes \mathbb{1}_{\mathcal{X}}) \Lambda_{\mathcal{M}} (\bar{V} \otimes \mathbb{1}_{\mathcal{X}})] \quad (5.36)$$

$$= \text{Tr}[\bar{V} V^T (\rho_{s,\text{av}} \otimes \mathbb{1})]. \quad (5.37)$$

Thus for the IC-norm we have that the two output operators have the same normalization

$$\text{Tr}[P_0] = \text{Tr}[P_1] = \text{Tr}[\bar{V} V^T (\rho_{s,\text{av}} \otimes \mathbb{1})] = p_V \quad (5.38)$$

and so

$$\|\mathcal{M}_{\text{IC}}\|_{\text{IC}} = \max \left\{ \frac{p_V}{2} \|\rho_0 - \rho_1\|_1 : V \in L(X \otimes \mathcal{X}), \|V\|_2 = 1 \right\}. \quad (5.39)$$

Now consider the operator V that maximizes the expression in [Eq. \(5.39\)](#). In terms of the corresponding normalized output density matrices $\rho_j = P_j / p_V$ we have

$$\frac{1}{p_V} \|\mathcal{M}_{\text{IC}}\|_{\text{IC}} = \frac{1}{2} \|\rho_0 - \rho_1\|_1 \quad (5.40)$$

and hence the probability of correctly distinguishing ρ_0 from ρ_1 using the optimal single shot strategy is given by

$$p = \frac{1}{2} \left(1 + \frac{1}{p_V} \|\mathcal{M}_{\text{IC}}\|_{\text{IC}} \right). \quad (5.41)$$

Now we look to bound the normalization factor p_V . From the definition in Eq. (5.38) we have

$$p_V = \text{Tr}[\bar{V} V^T (\rho_{s,\text{av}} \otimes \mathbb{1})] = \text{Tr}[B^\dagger (\rho_{s,\text{av}} \otimes \mathbb{1})] \quad (5.42)$$

where $B = \bar{V} V^T \geq 0$ and satisfies $\|B\|_1 = \text{Tr}[\bar{V} V^T] = 1$. Hence we have

$$p_V \leq \max \{ \text{Tr}[B^\dagger (\rho_{s,\text{av}} \otimes \mathbb{1})] : \|B\|_1 = 1 \} \quad (5.43)$$

$$= \|\rho_{s,\text{av}} \otimes \mathbb{1}\|_\infty \quad (5.44)$$

$$= \|\rho_{s,\text{av}}\|_\infty \quad (5.45)$$

For the lower bound, we consider the case where the reduced system state is a rank r , and has all r eigenvalues equal and given by $1/r$. In this case $\|\rho_{s,\text{av}}\|_\infty = \frac{1}{r}$. Any maximization over input V should beat this lower bound, as there is always another eigenvalue of the reduced system state which greater than or equal to this value. Hence we have

$$\frac{1}{r} \leq p_V \leq \|\rho_{s,\text{av}}\|_\infty \Rightarrow \|\rho_{s,\text{av}}\|_\infty^{-1} \leq \frac{1}{p_V} \leq r. \quad (5.46)$$

Thus we have that the optimal value that satisfies the IC-norm gives

$$\frac{\|\mathcal{M}_{\text{IC}}\|_{\text{IC}}}{\|\rho_{s,\text{av}}\|_\infty} \leq \frac{1}{2} \|\rho_0 - \rho_1\|_1 \leq r \|\mathcal{M}_{\text{IC}}\|_{\text{IC}} \quad (5.47)$$

So by the definition of state discrimination in terms of the trace distance (Helstrom's Theorem) we have that the optimal single shot strategy for distinguishing initial correlations succeeds with probability

$$\frac{1}{2} (1 + \text{IC}(\mathcal{M}_{\text{IC}})) \leq p \leq \frac{1}{2} (1 + r \|\mathcal{M}_{\text{IC}}\|_{\text{IC}}). \quad (5.48)$$

□

5.3.2 Preparation Fidelity

The information contained in the IC superchannel \mathcal{M}_{IC} can be used to optimize against the influence of the environment. We introduce the measure of *preparation fidelity* F_{prep} for

the case where high-fidelity projective preparation procedures are readily available, such as in photonic experiments, noting that similar measures could be defined for other scenarios. Consider a system preparation via initial post-selection on the state ρ . The subsequent evolution is then described by the effective map \mathcal{E}_ρ given by

$$\Lambda_{\mathcal{E}_\rho} = \frac{1}{p_\rho} \text{Tr}_1 \left[(\rho_1^\dagger \otimes \mathbb{1}_{23}) \Lambda_{\mathcal{M}_{\text{IC}}} \right], \quad (5.49)$$

where

$$p_\rho = \frac{1}{d} \text{Tr} \left[(\rho_1^\dagger \otimes \mathbb{1}_{23}) \Lambda_{\mathcal{M}_{\text{IC}}} \right] \quad (5.50)$$

is the probability of success for the post-selection on ρ . Studying these effective maps for different ρ allows us to find the optimal preparation procedure for any desired evolution of the system. We define the *average preparation fidelity* for post-selection in the state ρ as follows:

Definition 5.3.4 (Average Preparation Fidelity). Consider an IC superchannel $\mathcal{M}_{\text{IC}} \in T(\mathcal{X} \otimes \mathcal{X}, \mathcal{X})$, a density matrix $\rho \in D(\mathcal{X})$, and a target channel $\mathcal{U}_s \in C(\mathcal{X})$. We define the *average preparation fidelity* for implementing \mathcal{U}_s via a post-selected preparation procedure into the state ρ as

$$F_{\text{prep}}(\mathcal{M}_{\text{IC}}, \rho, U_s) = \frac{1}{d^2} F(\Lambda_{\mathcal{E}_\rho}, \Lambda_{U_s}) = \frac{1}{d^2} \text{Tr}[\Lambda_{\mathcal{E}_\rho} \Lambda_{U_s}]. \quad (5.51)$$

where \mathcal{E}_ρ is the effective channel form [Eq. \(5.49\)](#) after projective preparation of the system into the state ρ .

F_{prep} quantifies the process fidelity between the implemented effective map \mathcal{E}_ρ and the desired target unitary channel \mathcal{U}_s for initial projection onto ρ . The average preparation fidelity over all initial projections can be obtained from

$$\int d\rho \Lambda_{\mathcal{E}_\rho} = \Lambda_{\mathcal{E}_{av}} = \text{Tr}_1[\Lambda_{\mathcal{M}_{\text{IC}}}] \quad (5.52)$$

On the other hand, maximizing F_{prep} over all states ρ for a given target unitary U_s finds a preparation which allows for the highest quality implementation of U_s . Note that this is not equivalent to minimizing the impact of the environment, since the optimal preparation might harness some of the environmental correlations to improve the gate performance.

We note that we could have defined this expression analogous to average gate fidelity in [Section 3.2.3](#) by renormalizing [Eq. \(5.51\)](#) according to

$$\frac{1 + d F_{\text{prep}}(\mathcal{M}_{\text{IC}}, \rho, U_s)}{1 + d}. \quad (5.53)$$

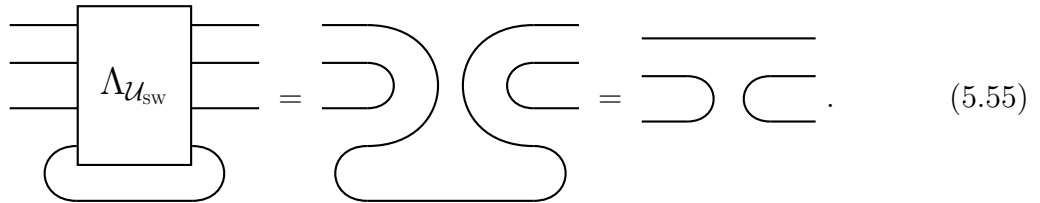
5.4 Example IC Superchannels

In this section we demonstrate the construction of IC superchannels for a classically correlated and an entangled initial system-environment state. We will also restrict ourselves to considering joint SE unitaries between the system and a subspace of the environment that is the same dimension as the the system. We note that while for full generality the dimension of the environment should be modelled as the square of the system dimension [Sch96, Nar07], the case of equal dimension case is sufficient to illustrate the IC superchannel techniques [SBC⁺11, CGM⁺12]. In particular, the subspace of the environment that may be initially correlated with the system cannot be of dimension greater than the system.

In the following we will consider two specific examples of SE interactions for arbitrary system dimensions. The first where we model the joint SE unitary as a swap-gate:

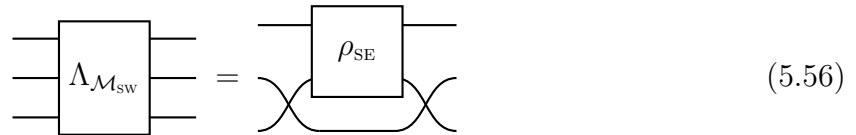
$$U_{\text{sw}} = \sum_{i,j} |z_j\rangle\langle x_i| \otimes |x_i\rangle\langle z_j|. \quad (5.54)$$

where $\{|x_j\rangle\}, \{|z_j\rangle\}$ are any orthonormal bases for the system and environment respectively. For this interaction we have that the partial trace over the environment output is given by



$$\Lambda_{U_{\text{sw}}} = \text{Diagram 1} = \text{Diagram 2} = \text{Diagram 3}. \quad (5.55)$$

Hence, from the definition in Eq. (5.11), for an arbitrary initial SE state the IC superchannel for the swap unitary \mathcal{M}_{sw} is given by the Choi-matrix



$$\Lambda_{\mathcal{M}_{\text{sw}}} = \text{Diagram 1} \quad (5.56)$$

In some sense this can be thought of as modelling the one of the worst errors that can happen to the system: no matter what preparation procedure one attempts to apply, the environment may replace the desired state with a different one. The actual state it is replaced with will depend greatly on the correlations in the initial state, and hence if properly characterized could potentially be used as a resource for state preparation.

The second interaction we consider is a controlled-unitary

$$U_{\text{CU}} = \sum_i U_i \otimes |z_i\rangle\langle z_i|. \quad (5.57)$$

for some specific choice of orthonormal basis $\{|z_i\rangle\}$ for the environment, and where we assume for simplicity that each of the unitaries U_i 's are orthogonal ($\text{Tr}[U_i^\dagger U_j] = d\delta_{ij}$). For this interaction we have the partial trace over the output joint unitary is given by the Choi-matrix

$$\Lambda_{U_{\text{CU}}} = \text{Diagram 1} = \text{Diagram 2} = \text{Diagram 3}. \quad (5.58)$$

Thus for an arbitrary initial SE state the IC superchannel \mathcal{M}_{CU} is given by

$$\Lambda_{\mathcal{M}_{\text{CU}}} = \text{Diagram 1} = \text{Diagram 2}. \quad (5.59)$$

This type of interaction can capture a Pauli-type error channel. If we wish to implement a target unitary U_1 this will succeed if the environment is in state $|z_1\rangle$. However, if the environment for some reason is in a state $|z_j\rangle$ $j \neq 1$ then a different unitary may be implemented resulting in an error. In an initially uncorrelated case setting the environment to a mixed state $\sum_j p_j |z_j\rangle\langle z_j|$, where p_j is the error probability, models this Pauli-channel type failure modality.

5.4.1 Classically Correlated Environment

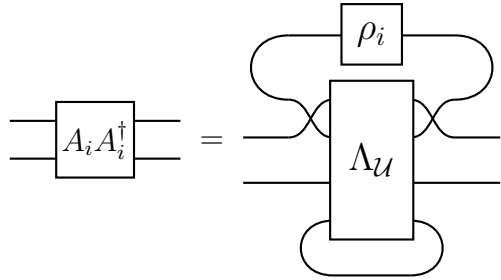
For our first example of initially correlated SE state we consider a classically correlated state of the form

$$\rho_{\text{se}} = \sum_{i=1}^r p_i |x_i\rangle\langle x_i| \otimes \rho_i, \quad 1 \leq r \leq \dim(\mathcal{X}) \quad (5.60)$$

where $1 \leq r \leq \dim(\mathcal{X})$, $p_i > 0$, $\sum_{i=1}^r p_i = 1$, and $\{|x_i\rangle\}$ is an orthonormal basis for \mathcal{X} . Note that this state is classically correlated in the sense that it is a 1-sided zero quantum discord state with respect to measurements performed on the system [MBC⁺12]. If we assume that the joint SE interaction is described by a joint unitary U_{se} then we have that the IC superchannel is given by

$$\Lambda_{\mathcal{M}_{\text{ic}}} = \sum_{i=1}^r p_i |x_i\rangle\langle x_i| \otimes A_i A_i^\dagger \quad (5.61)$$

where



$$\quad (5.62)$$

is a positive semidefinite operator satisfying $\text{Tr}_{\mathcal{Y}}[A_i A_i^\dagger] = \mathbb{1}$. Hence the average initial system state and channel is given by

$$\rho_{s,av} = \sum_{i=1}^r p_i |x_i\rangle\langle x_i| \quad (5.63)$$

$$\Lambda_{\mathcal{E},av} = \sum_{i=1}^r p_i A_i A_i^\dagger \quad (5.64)$$

and so the correlation memory matrix is given by

$$\Lambda_{\Delta M_{\text{ic}}} = \sum_{i=1}^r p_i |x_i\rangle\langle x_i| \otimes \left(A_i A_i^\dagger - \sum_{j=1}^r p_j A_j A_j^\dagger \right) \quad (5.65)$$

$$= \sum_{i=1}^r p_i |x_i\rangle\langle x_i| \otimes \left((1 - p_i) A_i A_i^\dagger - \sum_{j=1, j \neq i}^r p_j A_j A_j^\dagger \right) \quad (5.66)$$

$$= \sum_{i=1}^r p_i (1 - p_i) |x_i\rangle\langle x_i| \otimes \left(A_i A_i^\dagger - B_i B_i^\dagger \right) \quad (5.67)$$

where

$$B_i B_i^\dagger = \sum_{j=1, j \neq i}^r \frac{p_j}{1 - p_i} A_j A_j^\dagger \quad (5.68)$$

is a positive semidefinite operator satisfying $\text{Tr}_Y[B_i B_i^\dagger] = \mathbb{1}$.

Now we can compute

$$|\Lambda_{\Delta\mathcal{M}_{\text{IC}}}| = \sum_{i=1}^r p_i(1-p_i)|x_i\rangle\langle x_i| \otimes \sqrt{(A_i A_i^\dagger - B_i B_i^\dagger)^2} \quad (5.69)$$

Hence an upper bound for the IC-norm is

$$\frac{1}{2} \left\| \text{Tr}_3 |\Lambda_{\Delta\mathcal{M}_{\text{IC}}}| \right\|_\infty = \frac{1}{2} \left\| \sum_{i=1}^r p_i(1-p_i)|x_i\rangle\langle x_i| \otimes \text{Tr}_Y \sqrt{(A_i A_i^\dagger - B_i B_i^\dagger)^2} \right\|_\infty \quad (5.70)$$

$$\leq \frac{1}{2} \left\| \sum_{i=1}^r p_i(1-p_i)|x_i\rangle\langle x_i| \otimes \text{Tr}_Y [A_i A_i^\dagger + B_i B_i^\dagger] \right\|_\infty \quad (5.71)$$

$$= \frac{1}{2} \left\| \sum_{i=1}^r 2p_i(1-p_i)|x_i\rangle\langle x_i| \otimes \mathbb{1} \right\|_\infty \quad (5.72)$$

$$= \left\| \sum_{i=1}^r p_i(1-p_i)|x_i\rangle\langle x_i| \right\|_\infty \quad (5.73)$$

$$= \max \{p_i(1-p_i) : i = 1, \dots, r\}. \quad (5.74)$$

where equality holds when $A_i A_i^\dagger B_i B_i^\dagger + B_i B_i^\dagger A_i A_i^\dagger = 0$. The maximum value of Eq. (5.74) is 1/4 obtained when $r = 2$, and $p_1 = p_2 = 1/2$. Hence we have

$$\|\mathcal{M}_{\text{IC}}\|_{\text{IC}} \leq \frac{1}{4} \quad (5.75)$$

For bounding the IC-Value we note that for a maximally classically correlated state $p_i = 1/r$ for all i and hence the IC value is maximized by

$$\text{IC}(\mathcal{M}_{\text{IC}}) \leq \frac{r-1}{r} \quad (5.76)$$

We also derive an expression for the preparation fidelity for a classically correlated input states. Suppose we consider post-selection on a pure state $\rho = |\psi\rangle\langle\psi|$ where $|\psi\rangle = \sum_j \lambda_j |x_j\rangle$. Then we have that the effective channel after projective preparation is given by

$$\Lambda_{\mathcal{E}_\rho} = \frac{1}{p_\rho} \sum_{i,j,k} \lambda_i \bar{\lambda}_j p_k \langle x_j | x_k \rangle \langle x_j | x_i \rangle A_k A_k^\dagger = \frac{1}{p_\rho} \sum_i p_i |\lambda_i|^2 A_i A_i^\dagger \quad (5.77)$$

where $p_\rho = \sum_i p_i |\lambda_i|^2$. Hence we have that the average preparation fidelity for ρ with respect to a target unitary U_s is given by

$$F_{\text{prep}}(\mathcal{M}_{\text{IC}}, |\psi\rangle\langle\psi|, U_s) = \frac{\sum_i p_i |\lambda_i|^2 \langle\langle U_s | A_i A_i^\dagger | U_s \rangle\rangle}{d^2 \sum_j p_j |\lambda_j|^2}. \quad (5.78)$$

Let us now consider the specific example of the swap gate and controlled-unitary SE interactions for a fully classically correlated initial state. By fully classically correlated we mean that the state has zero quantum discord with respect to measurements on both the system and environment [MBC⁺12]. This is achieved by setting $\rho_i = |z_i\rangle\langle z_i|$ for some orthonormal basis $\{|z_i\rangle\}$ for the environment. For the two interactions we have that the operator $A_i A_i^\dagger$ in Eq. (5.61) is given by

$$U_{\text{sw}} : A_i A_i^\dagger = \mathbb{1} \otimes |z_i\rangle\langle z_i| \quad (5.79)$$

$$U_{\text{CU}} : A_i A_i^\dagger = |U_i\rangle\langle U_i| \quad (5.80)$$

Hence by Theorem 3.2.2 we have that for this case the upper bound of the CB trace norm is tight and in both cases

$$\|\mathcal{M}_{\text{IC}}\|_{\text{IC}} = \max \{p_i(1 - p_i) : i = 1, \dots, r\} \quad (5.81)$$

$$\text{IC}(\mathcal{M}_{\text{IC}}) = \frac{\max \{p_i(1 - p_i) : i = 1, \dots, r\}}{\max \{p_i : i = 1, \dots, r\}} \quad (5.82)$$

Thus if we have a maximally classically correlated initial state with $r = d$ and $p_i = 1/d$ for all i , we have

$$\text{IC}(\mathcal{M}_{\text{IC}}) = \frac{d - 1}{d} \quad (5.83)$$

which approaches 1 for $d \gg 1$. We see here that this is tight with the upper bound for the one-sided discord state originally considered in Eq. (5.60), and hence the additional correlations that one may have when there is non-zero discord with respect to measurement on the systems do not have any operational difference with respect to the IC value. The fact that we can still experience different reduced system dynamics, as quantified by the IC-value, even with classical initial correlations is not surprising as it has been shown previously that having a zero quantum discord state is not sufficient for obtaining CP description reduced dynamics from a standard tomographic experiment [Woo08, BGTW11].

If we consider preparation fidelity for these examples we have that for the swap interaction $\langle\langle U_s | A_i A_i^\dagger | U_s \rangle\rangle = \langle z_i | z_i \rangle = 1$, and hence

$$F_{\text{prep}}(\mathcal{M}_{\text{sw}}, |\psi\rangle\langle\psi|, U_s) = \frac{1}{d^2}. \quad (5.84)$$

so changing our initial projection does not change the average fidelity of the resulting channel for subsequent local rotations for preparing other states. For the controlled gate, suppose the target unitary we are trying to implement is U_1 , then we have

$$F_{\text{prep}}(\mathcal{M}_{\text{CU}}, |\psi\rangle\langle\psi|, U_s) = \frac{p_1 |\lambda_1|^2}{\sum_j p_j |\lambda_j|^2} \quad (5.85)$$

hence in this case we can obtain perfect average preparation fidelity by selecting $\lambda_1 = 1$ and performing the initial post-selection in the state $\rho = |x_1\rangle\langle x_1|$, which also post-selects the correlated environment into the state that ensures U_1 is the reduced system operation.

5.4.2 Maximally Entangled Environment

Now we consider the swap gate and controlled unitary SE interactions in the case of a maximally entangled initial SE state

$$\rho_{\text{se}} = \sum_{i,j=1}^d \frac{1}{d} |x_i\rangle\langle x_j| \otimes |z_i\rangle\langle z_j| \quad (5.86)$$

for some choice of orthonormal bases $\{|x_j\rangle\}, \{|z_j\rangle\}$ for the system and environment respectively. Note that here we assume that the basis for the environment is the same as the control basis for the controlled unitary in [Eq. \(5.57\)](#). Using the swap gate SE interaction in [Eq. \(5.54\)](#) we have the Choi-matrix for the IC superchannel is given by

$$\Lambda_{\mathcal{M}_{\text{sw}}} = \sum_{i,j=1}^d \frac{1}{d} |x_i\rangle\langle x_j| \otimes \mathbb{1} \otimes |z_i\rangle\langle z_j| \quad (5.87)$$

Hence the reduced average system state and average channel are given by

$$\rho_{s,av} = \frac{\mathbb{1}}{d}, \quad \Lambda_{\mathcal{E},av} = \frac{1}{d} \mathbb{1} \otimes \mathbb{1} \Rightarrow \Lambda_{\mathcal{M}_{\text{sep}}} = \frac{1}{d^2} \mathbb{1} \otimes \mathbb{1} \otimes \mathbb{1}. \quad (5.88)$$

This gives correlation memory matrix of the form

$$\Lambda_{\Delta\mathcal{M}_{\text{sw}}} = \sum_{i,j=1}^d \frac{1}{d} |x_i\rangle\langle x_j| \otimes \mathbb{1} \otimes |z_i\rangle\langle z_j| - \frac{1}{d^2} \mathbb{1} \otimes \mathbb{1} \otimes \mathbb{1} \quad (5.89)$$

$$= V \left[\left(|B_1\rangle\langle B_1| - \sum_{\alpha=1}^{d^2} \frac{1}{d^2} |B_\alpha\rangle\langle B_\alpha| \right) \otimes \mathbb{1} \right] V^\dagger \quad (5.90)$$

$$(5.91)$$

where V is a unitary that swaps the second and third subsystems, and $\{|B_\alpha\rangle\rangle\}$ is an orthogonal basis for $\mathcal{X} \otimes \mathcal{Y}$ consisting of vectorized unitaries with $B_1 = \sum_{j=1}^d |z_j\rangle\langle x_j|$. This basis satisfies

$$\mathrm{Tr}_2[|U_\alpha\rangle\rangle\langle\langle U_\alpha|] = \mathbb{1}, \quad \sum_{\alpha=1}^{d^2} |U_\alpha\rangle\rangle\langle\langle U_\alpha| = d\mathbb{1} \otimes \mathbb{1} \quad (5.92)$$

and so we have that

$$\mathrm{Tr}_Y |\Lambda_{\Delta\mathcal{M}_{\mathrm{sw}}}| = \frac{1}{d} \mathrm{Tr}_2 \left[\left(\frac{d^2-1}{d^2} |B_1\rangle\rangle\langle\langle B_1| + \sum_{\alpha=2}^{d^2} \frac{1}{d^2} |B_\alpha\rangle\rangle\langle\langle B_\alpha| \right) \right] \otimes \mathbb{1} \quad (5.93)$$

$$= \frac{1}{d} \left(\frac{d^2-1}{d^2} + \sum_{\alpha=2}^{d^2} \frac{1}{d^2} \right) \mathbb{1} \otimes \mathbb{1} \quad (5.94)$$

$$= \frac{2}{d} \left(\frac{d^2-1}{d^2} \right) \mathbb{1} \otimes \mathbb{1} \quad (5.95)$$

Now since the correlation memory matrix for the swap gate satisfies the conditions for [Theorem 3.2.2](#) we have

$$\|\mathcal{M}_{\mathrm{sw}}\|_{\mathrm{IC}} = \frac{d^2-1}{d^3} \quad (5.96)$$

$$\mathrm{IC}(\mathcal{M}_{\mathrm{sw}}) = \frac{d^2-1}{d^2} \quad (5.97)$$

For the case of the controlled unitary SE interaction in [Eq. \(5.57\)](#) we have

$$\Lambda_{\mathcal{M}_{\mathrm{cu}}} = \sum_{i=1}^d \frac{1}{d} |x_i\rangle\langle x_i| \otimes |U_i\rangle\rangle\langle\langle U_i| \quad (5.98)$$

which gives an average separable IC superchannel of

$$\rho_{s,av} = \frac{\mathbb{1}}{d}, \quad \Lambda_{\mathcal{E},av} = \sum_{i=1}^d \frac{1}{d} |U_i\rangle\rangle\langle\langle U_i| \Rightarrow \Lambda_{\mathcal{M}_{\mathrm{sep}}} = \sum_{i=1}^d \frac{1}{d^2} \mathbb{1} \otimes |U_i\rangle\rangle\langle\langle U_i| \quad (5.99)$$

Hence the correlation memory matrix is given by

$$\Lambda_{\Delta\mathcal{M}_{\text{cu}}} = \sum_i^d \frac{1}{d} \left(|x_i\rangle\langle x_i| - \frac{\mathbb{1}}{d} \right) \otimes |U_i\rangle\langle U_i| \quad (5.100)$$

$$= \sum_i^d \frac{1}{d} \left[\left(\frac{d-1}{d} \right) |x_i\rangle\langle x_i| - \sum_{j=1, j \neq i}^d \frac{1}{d} |x_j\rangle\langle x_j| \right] \otimes |U_i\rangle\langle U_i| \quad (5.101)$$

$$= \sum_i^d \frac{1}{d} \left[\left(\frac{d-1}{d} \right) |x_i\rangle\langle x_i| - \sum_{j=1, j \neq i}^d \frac{1}{d} |x_j\rangle\langle x_j| \right] \otimes |U_i\rangle\langle U_i| \quad (5.102)$$

$$(5.103)$$

Hence

$$\text{Tr}_{\mathcal{Y}} |\Lambda_{\Delta\mathcal{M}_{\text{cu}}}| = \sum_i^d \frac{1}{d} \left[\left(\frac{d-1}{d} \right) |x_i\rangle\langle x_i| + \sum_{j=1, j \neq i}^d \frac{1}{d} |x_j\rangle\langle x_j| \right] \otimes \text{Tr}_{\mathcal{Y}} [|U_i\rangle\langle U_i|] \quad (5.104)$$

$$= \sum_i^d \frac{1}{d} \left[\left(\frac{d-1}{d} \right) |x_i\rangle\langle x_i| + \frac{1}{d} (\mathbb{1} - |x_i\rangle\langle x_i|) \right] \otimes \mathbb{1} \quad (5.105)$$

$$= \frac{2}{d} \left(\frac{d-1}{d} \right) \mathbb{1} \otimes \mathbb{1} \quad (5.106)$$

Since $\Lambda_{\Delta\mathcal{M}_{\text{cu}}}$ satisfies the conditions of [Theorem 3.2.2](#) we have

$$\|\mathcal{M}_{\text{cu}}\|_{\text{IC}} = \frac{d-1}{d^2} \quad (5.107)$$

$$\text{IC}(\mathcal{M}_{\text{cu}}) = \frac{d-1}{d} \quad (5.108)$$

Here we see that for a maximally entangled initial SE state the IC-norm and IC-value for the swap interaction is strictly greater than for the controlled interaction. This is in contrast to the classically correlated initial state where the IC-norm and value were equal for the swap and controlled-unitary SE interactions.

If we consider preparation fidelity for the two interaction with an entangled input state we find that the effective channels after post-selection on a pure state $\rho = |\psi\rangle\langle\psi|$ where $|\psi\rangle = \sum_j \lambda_j |x_j\rangle$ are given by

$$\Lambda_{\mathcal{E}_{\rho, \text{sw}}} = \sum_i \bar{\lambda}_i \lambda_j \mathbb{1} \otimes |z_i\rangle\langle z_j| \quad (5.109)$$

$$\Lambda_{\mathcal{E}_{\rho, \text{cu}}} = \sum_i |\lambda_i|^2 |U_i\rangle\langle U_i|. \quad (5.110)$$

Hence we have that the average preparation fidelity for ρ with respect to a target unitary U_s is given by

$$F_{\text{prep}}(\mathcal{M}_{\text{sw}}, |\psi\rangle\langle\psi|, U_s) = \frac{1}{d^2} \quad (5.111)$$

$$F_{\text{prep}}(\mathcal{M}_{\text{cu}}, |\psi\rangle\langle\psi|, U_s) = \sum_i |\lambda_i|^2 |\langle\langle U_s | U_i \rangle\rangle|^2. \quad (5.112)$$

This means that we can maximize the average preparation fidelity to be 1 for the controlled-unitary interaction so long as $U_s = U_i$ for some i . The average preparation fidelity for the swap interaction is, however, independent of our choice of initial projection. These results agree with those for the classically correlated initial state. We make the comment that while the swap interaction leads to terrible average preparation fidelity, if one didn't care about the average case, but only on the fidelity of returning a given output state ρ from the channel, then the swap interaction could actually be used as a resource if the system was sufficiently correlated with the environment. In this sense one could engineer the output state by projecting the system into the state that leaves the environment in the state closest to the desired output.

5.4.3 Qubit System

We now consider an explicit example of $d = 2$ qubit case where we parameterize the strength of the initial correlations for two classes of initial SE states considered in [Sections 5.4.1](#) and [5.4.2](#). We consider a class of initial states which may display purely classical correlations defined by

$$\rho_{\text{SE}}^{\text{C}}(\theta) = \cos^2(\theta) |0\rangle\langle 0| \otimes |0\rangle\langle 0| + \sin^2(\theta) |1\rangle\langle 1| \otimes |1\rangle\langle 1| \quad (5.113)$$

The class of initial states we consider which may display entanglement are given by

$$\rho_{\text{SE}}^{\text{Q}}(\theta) = |\Psi(\theta)\rangle\langle\Psi(\theta)| \quad (5.114)$$

$$(5.115)$$

$$\text{where } |\Psi(\theta)\rangle = \cos(\theta) |0\rangle \otimes |0\rangle + \sin(\theta) |1\rangle \otimes |1\rangle. \quad (5.116)$$

In both cases the parameter θ determines the strength of initial correlations, with the minimum and maximum correlated state corresponding to $\theta = 0$ and $\theta = \pi/4$ respectively. We also consider the two SE interactions given by a swap gate and a controlled-unitary in [Eqs. \(5.54\)](#) and [\(5.57\)](#) respectively.

For the classically correlated initial state using our results in Eqs. (5.82) and (5.82) we have

$$\max \{p_i(1 - p_i) : i = 1, \dots, r\} = \sin^2(\theta) \cos^2(\theta) = \frac{1}{4} \sin^2(2\theta) \quad (5.117)$$

and

$$\max \{p_i : i = 1, \dots, r\} = \max\{\cos^2(\theta), \sin^2(\theta)\} \quad (5.118)$$

Hence for both SE unitaries have the IC-norm and IC-value are given by

$$\|\mathcal{M}_{\text{sw}}^c\|_{\text{IC}} = \|\mathcal{M}_{\text{cu}}^c\|_{\text{IC}} = \frac{1}{4} \sin^2(2\theta) \quad (5.119)$$

$$\text{IC}(\mathcal{M}_{\text{IC}}^c) = \min\{\sin^2(\theta), \cos^2(\theta)\} \quad (5.120)$$

In the case of the entangled state we must solve the norm explicitly as we only considered the maximally entangled initial state in Section 5.4.2. For the swap interaction we have

$$\Lambda_{\mathcal{M}_{\text{sw}}^q}(\theta) = V(|\Psi(\theta)\rangle\langle\Psi(\theta)| \otimes \mathbb{1})V^\dagger \quad (5.121)$$

where V is a unitary which swaps the second and third subsystems of \mathcal{M} . We can compute the upper and lower bounds using Mathematica. We find them to be

$$\frac{1}{8} \|\Lambda_{\Delta\mathcal{M}_{\text{sw}}^q}\|_1 = \frac{1}{2} \|\text{Tr}_Y |\Lambda_{\Delta\mathcal{M}_{\text{sw}}^q}\|_\infty = \frac{1}{8} (\sin^2(2\theta) + 2|\sin(2\theta)|) \quad (5.122)$$

Hence we have that the IC-norm and IC-value are given by

$$\|\mathcal{M}_{\text{sw}}^q(\theta)\|_{\text{IC}} = \frac{1}{8} (\sin^2(2\theta) + 2|\sin(2\theta)|) \quad (5.123)$$

$$\text{IC}(\mathcal{M}_{\text{sw}}^q(\theta)) = \frac{\sin^2(2\theta) + 2|\sin(2\theta)|}{8 \max\{\cos^2(\theta), \sin^2(\theta)\}} \quad (5.124)$$

Fig. 5.2 shows the comparison between the IC-value and IC-norm for the classically correlated, and the entangled initial state as a function of the θ . Here we have normalized the IC-norm by the rank of the initial state, so that it obtains the same maximum value as the IC-value in the maximally correlated case. We see that the entangled state leads to a strictly greater measure of initial correlations for all values of θ .

For the controlled-unitary interaction we have

$$\Lambda_{\mathcal{M}_{\text{cu}}^q}(\theta) = \cos^2(2\theta)|0\rangle\langle 0| \otimes |U_0\rangle\langle U_0| + \sin^2(2\theta)|1\rangle\langle 1| \otimes |U_1\rangle\langle U_1|. \quad (5.125)$$

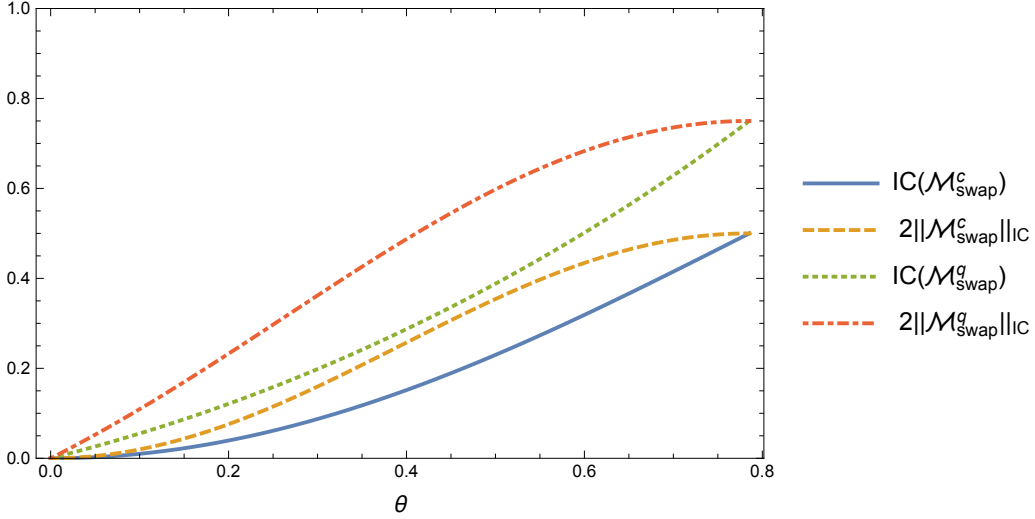


Figure 5.2: Comparison of the IC-Value and rank-normalized IC-norm for an initially correlated qubit state interacting with the environment by a swap-gate. The $\mathcal{M}_{\text{sw}}^c, \mathcal{M}_{\text{sw}}^q$ correspond to the classically correlated and entangled initial SE states in Eqs. (5.113) and (5.114) respectively. θ parameters the degree of correlation with $\theta = \pi/4$ corresponding to the maximally correlated configuration for each state. We see that the measure of initial correlations via the IC-norm and IC-value is strictly greater for the entangled state than the classically correlated state.

where, as before, we assume that $\text{Tr}[U_0^\dagger U_1] = 0$. By computing the upper and lower bounds using Mathematica and we find them to be

$$\frac{1}{8} \|\Lambda_{\Delta \mathcal{M}_{\text{cu}}^q}\|_1 = \frac{1}{2} \|\text{Tr}_y |\Lambda_{\Delta \mathcal{M}_{\text{cu}}^q}|\|_\infty = \frac{1}{4} \sin^2(2\theta) \quad (5.126)$$

Hence we have that the IC-norm and IC-value are given by

$$\|\mathcal{M}_{\text{cu}}^q(\theta)\|_{\text{IC}} = \frac{1}{4} \sin^2(2\theta) \quad (5.127)$$

$$\text{IC}(\mathcal{M}_{\text{cu}}^q(\theta)) = \min \{ \sin^2(\theta), \cos^2(\theta) \}. \quad (5.128)$$

We see here that this is the same for the classical correlated initial state. Both these values also correspond to the same values as for the classically correlated initial state with the swap interaction in Fig. 5.2.

5.5 Experimental Demonstration

We now demonstrate the characterization of initial correlations using an IC superchannel for the case of a system and environment each consisting of a single photonic qubit. The experimental setup is given in Fig. 5.3 and the experiment was implemented by M. Ringbauer, A. G. White and A. Fedrezzi at the University of Queensland [RWM⁺15]. In the following we will use polarization basis notation for the states of the system and environment. These are defined in terms of the computational basis as:

$$\begin{aligned} |H\rangle &\equiv |0\rangle & |D\rangle &\equiv \frac{1}{\sqrt{2}}(|0\rangle + |1\rangle) & |R\rangle &\equiv \frac{1}{\sqrt{2}}(|0\rangle + i|1\rangle) \\ |V\rangle &\equiv |1\rangle & |A\rangle &\equiv \frac{1}{\sqrt{2}}(|0\rangle - |1\rangle) & |L\rangle &\equiv \frac{1}{\sqrt{2}}(|0\rangle - i|1\rangle). \end{aligned}$$

where $|H\rangle, |V\rangle, |D\rangle, |A\rangle, |R\rangle, |L\rangle$ correspond to horizontal, vertical, diagonally, anti-diagonally, right-circular, and left-circular polarization respectively.

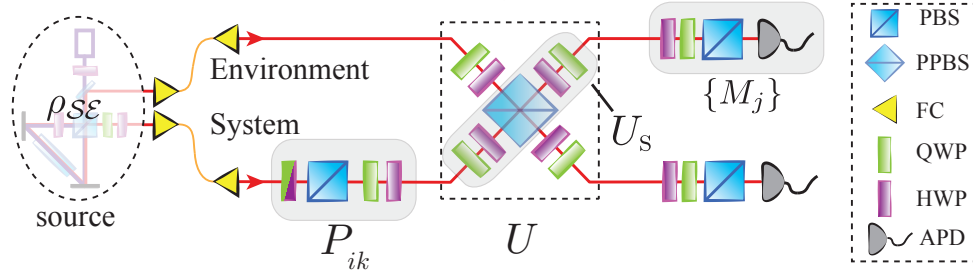


Figure 5.3: Experimental setup. System and environment photons are created in the state ρ_{SE} with controllable degree of entanglement, using the source of Ref. [FTP⁺07]. Arbitrary preparations \mathcal{P}_{ij} on the system and measurements $\{M_j\}$ are implemented by means of polarizers (PBS), quarter- and half-wave plates (QWP, HWP) and single-photon detectors (APD). The joint SE evolution U is implemented as a CZ gate between a set of HWPs and QWPs. In the case of no initial correlations this setup implements the target system evolution U_s . The CZ gate is based on non-classical interference at a partially polarizing beam splitter (PPBS) with reflectivities of $r_H = 0$ ($r_V = 2/3$) for horizontally (vertically) polarized light [LWP⁺05].

Our model for the system-environment interaction is that of a controlled unitary operation as was discussed Section 5.4. Suppose an experimenter aims to implement the target system evolution described by the unitary operator U_s which is one of the following three gates:

1. Pauli-Z gate: $U_s = Z$.
2. Hadamard gate: $U_s = H = R_Y Z R_Y^\dagger$.
3. Rotation gate: $U_s = Z R_Y$.

where $R_Y = \exp(-i\frac{\pi}{4}\sigma_y)$, denotes a $\pi/4$ -rotation around the σ_y -axis. Due to coupling to the environment the reduced dynamics of the system will in general deviate from that described by U_s . We simulate this influence by replacing the Z operations in the above decomposition of U_s by controlled- Z (CZ) operations

$$U_{\text{CZ}} = \mathbb{1}_s \otimes |H\rangle\langle H|_e + Z_s \otimes |V\rangle\langle V|_e. \quad (5.129)$$

This means the interaction of the first subsystem is switched on and off conditional on the state of the environment, which is modelled as another photonic qubit. In the case of target Z and H gates the environment might thus cause a failure of the system unitary (i.e. the identity operation is implemented), while in the case $U_s = Z R_Y$ it can introduce a phase error.

The initial SE state was generated via spontaneous parametric down conversion in the form

$$|\psi\rangle_{\text{SE}} = \cos(2\theta) |H\rangle_s \otimes |V\rangle_e + \sin(2\theta) |V\rangle_s \otimes |H\rangle_e, \quad (5.130)$$

where $|H\rangle, |V\rangle$ correspond to horizontally and vertically polarized photons respectively. In this case the strength of the initial correlations is parametrized by the *tangle*:

$$\tau = \sin^2(4\theta) \quad (5.131)$$

and can be tuned from uncorrelated ($\theta = 0$) to maximal correlation ($\theta = \pi/8$) [FTP⁺07]. We prepared states with tangle

$$\tau = \{0.012, 0.136, 0.423, 0.757, 0.908\} \quad (5.132)$$

with an average fidelity of $F = 0.96(1)$ with the corresponding ideal state in Eq. (5.130). The system was then subjected to the preparation procedure P_{ij} from Eq. (4.35), which prepared it in the state ρ_j by first projecting onto the state ρ_i followed by a unitary rotation. Here for each index we select the state from the tomographically over-complete set

$$\mathbf{P} = \{|H\rangle, |V\rangle, |D\rangle, |A\rangle, |R\rangle, |L\rangle\}. \quad (5.133)$$

State tomography of the output state was also implemented using measurement of \mathbf{P} . The description of the Choi-matrix $\Lambda_{\mathcal{M}}$ was then reconstructed using the SDP in Section 5.2.2.

Since our experiment had good counting statistics ($N_{ijk} > 5000$) we used a small value of $\beta = 0.1$ for our hedging parameter in Eq. (4.43) for the observed probabilities p_{ijk} . We note, however, that the reconstruction is not particularly sensitive to the value of β .

Further, since the total number of counts N_{ijk} is unknown a priori, we define it for our experiment by totalling the observed counts for measurement configurations that sum to identity (eg. for $|H\rangle$ and $|V\rangle$). Since the second index of Π_{ijk} corresponds to the rotated state for the initial projective preparation procedure, only the first and third indices correspond to true measurements and so we have $K = 4$ in Eq. (4.43).

5.5.1 Reconstructed IC Superchannels

Since separable IC superchannels have the product form $\rho_s \otimes \Lambda_{\mathcal{E}}$ we will express the reconstructed Choi-matrices using the polarization basis for the index corresponding to the effective initial state, and the Pauli basis for the indices corresponding to the effective channel. This is equivalent to expressing a separable IC superchannel as $\rho_s \otimes \chi_{\mathcal{E}}$, where $\chi_{\mathcal{E}}$ is the χ -matrix. This is achieved by a partial change of basis of the Choi-matrix by applying the Pauli vectorization change of basis operator to the second two indices of $\Lambda_{\mathcal{M}}$. The real part of the reconstructed Choi-matrix $\Lambda_{\mathcal{M}}$ for nominal unitary $U_s = H$ and Z for increasing tangles of the initial state are shown in Fig. 5.4a and Fig. 5.4b respectively. In both these plots as the tangle of the initial state we see the emergence of a peak corresponding to the identity operation (shown in yellow). This is characteristic for the simulated increased tendency of the nominal single-qubit operation U_s (shown in green) to fail in the presence of stronger initial correlations. For these reconstructions the imaginary part is negligible and not shown. The real and imaginary parts of the Choi-matrix for $U_s = ZR_Y$ are shown in Fig. 5.5a and Fig. 5.5b respectively. Here the emerging peaks for increased initial correlation correspond to a phase error acting on the system.

To contrast the reconstruction of the IC superchannel with standard quantum process tomography, we show the reconstruction of the χ -matrix for the channel \mathcal{U}_s for different preparation procedures and low initial correlation of $\tau = 0.136$ in for the case of nominal $U_s = H$ in Fig. 5.6. We see here how different choices of system preparation procedure can lead to vastly different reconstructed channels with fidelities varying between 0.853 and 0.683. The superchannel \mathcal{M} in Fig. 5.4a clearly illustrates the reason for this discrepancy: a term that corresponds to the identity operation and increases with the strength of initial correlations. This is exactly the simulated environment-induced failure mode of the system evolution.

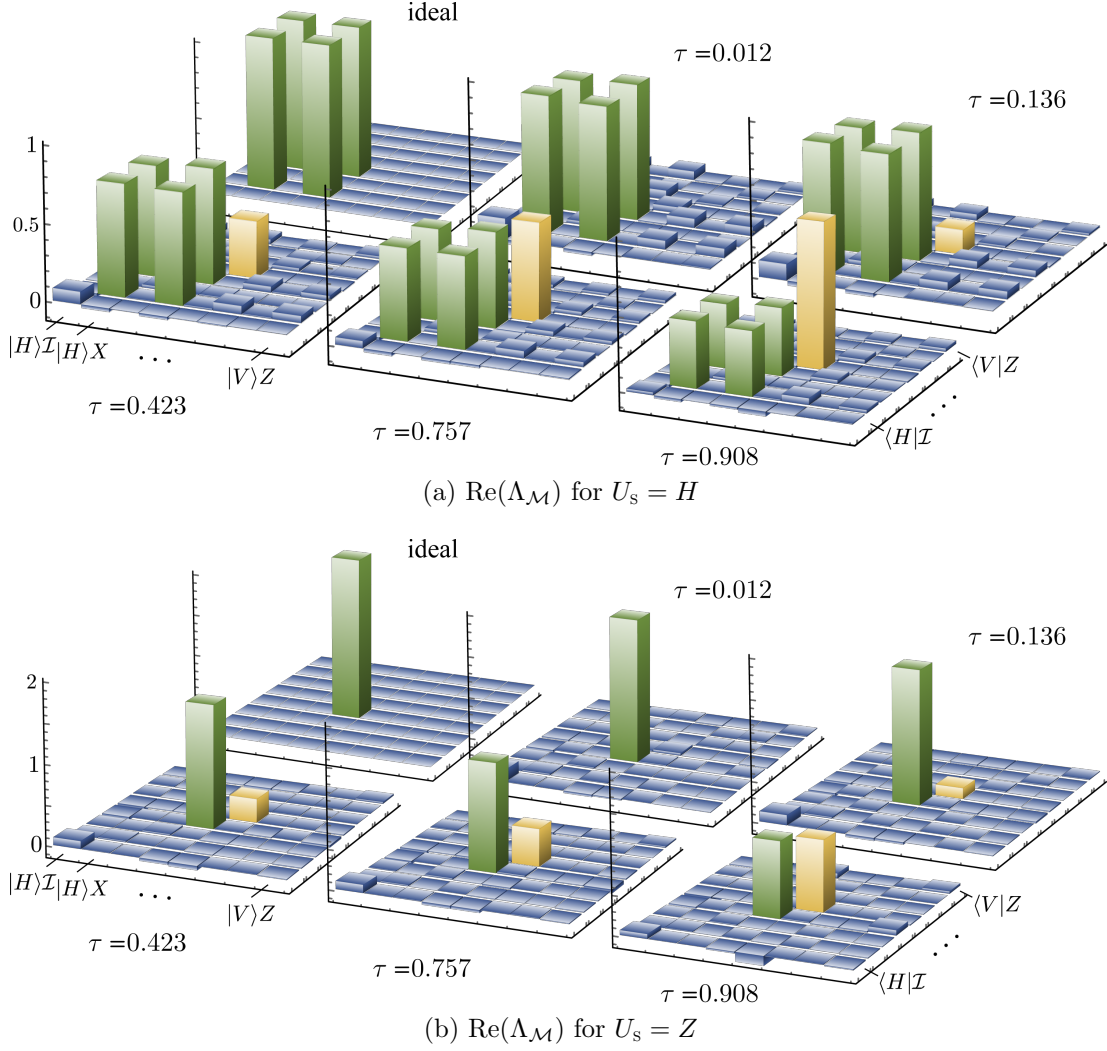


Figure 5.4: Real parts of $\Lambda_{\mathcal{M}}$ for (a) $U_s = H$ and (b) $U_s = H$ in the ideal, uncorrelated case and experimental results for increasing strength of initial correlations. The matrices $\Lambda_{\mathcal{M}}$ are shown in a polarization-Pauli basis, with the elements from left to right corresponding to $\{|H\rangle, |V\rangle\} \otimes \{\mathcal{I}, X, Y, Z\}$ and from front to back corresponding to $\{\langle H|, \langle V|\} \otimes \{\mathcal{I}, X, Y, Z\}$. The emergence of a peak corresponding to the identity operation (shown in yellow) is characteristic for the simulated increased tendency of the single-qubit operation U_s (shown in green) to fail in the presence of stronger initial correlations. The negligible imaginary parts are not shown.

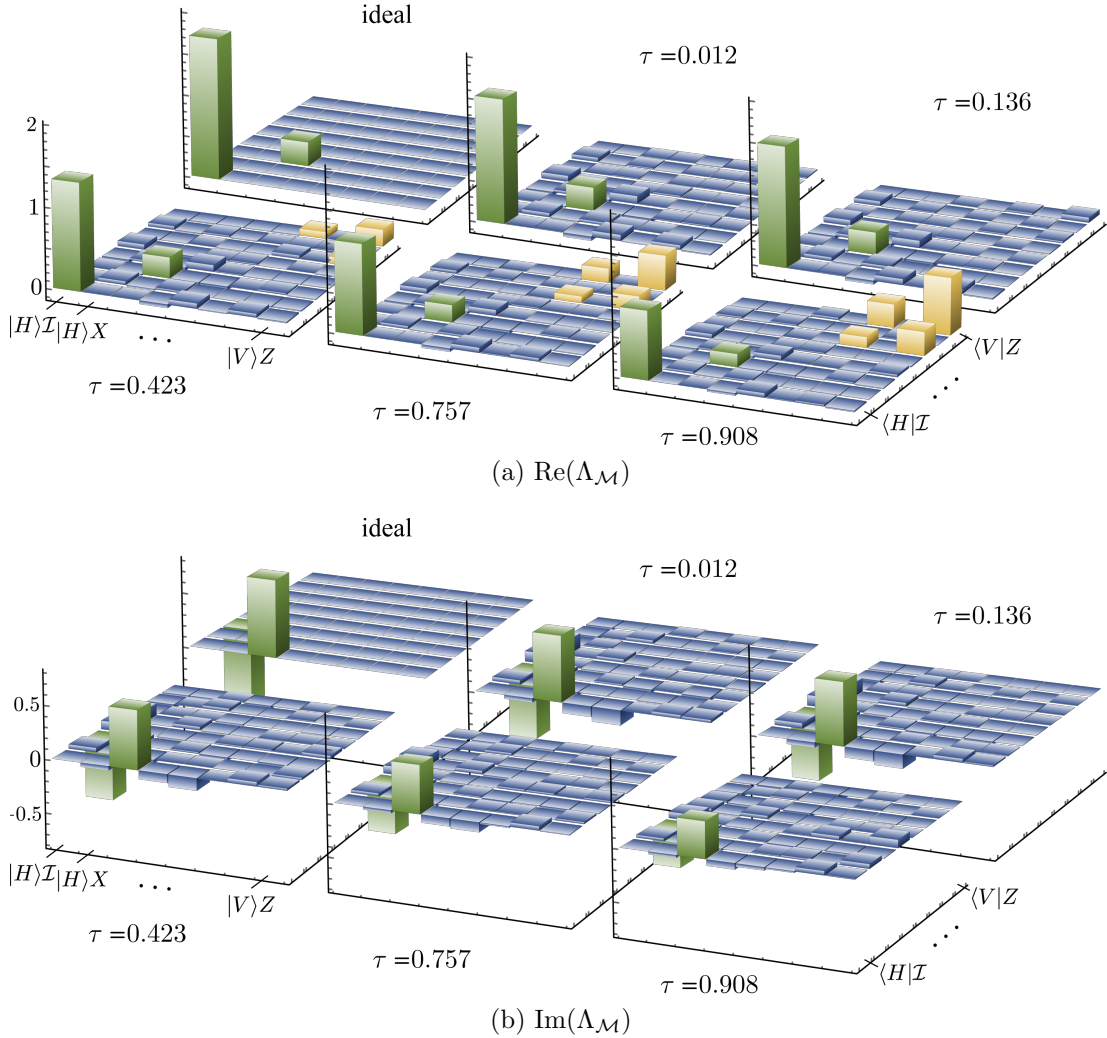


Figure 5.5: (a) Real and (b) imaginary parts of $\Lambda_{\mathcal{M}}$ for an intended $U = R_Y Z$ operation on the system in the ideal, uncorrelated case and experimentally for increasing strength of initial correlations. The matrices $\Lambda_{\mathcal{M}}$ are shown in the same basis as in Fig. 5.4.

5.5.2 Initial-Correlation Norm and Preparation Fidelity

The computed value of the IC norm $\|\mathcal{M}\|_{\text{IC}}$ for the reconstructed channels in Section 5.5.1 plotted against the correlation strength τ of the simulated initial SE states are shown in Fig. 5.7, here we have rank-normalized the IC-norm by multiplying by rank of the input state. For all three SE interactions U_s the maximum obtained value of $2\|\mathcal{M}\|_{\text{IC}}$ is

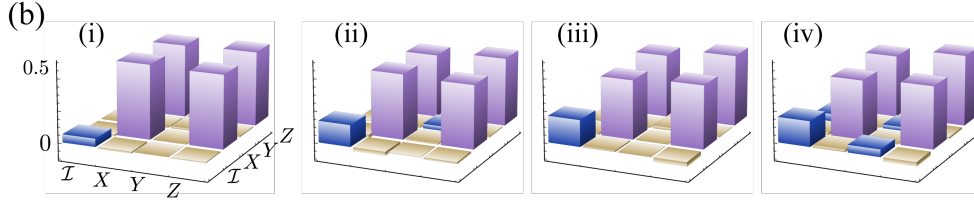


Figure 5.6: Real parts of the χ -matrices $\chi_{\mathcal{U}}$ (shown in the Pauli basis) for U_s obtained via QPT for different choices of preparation procedure in the case of low initial correlation $\tau = 0.136$. Cases (i) and (ii) correspond to a fixed ρ_k in Fig. 5.1(b), (iii) corresponds to $\rho_k = \rho_i$, and (iv) is the case where $1 \leq k \leq 4$. The information contained in the superchannel \mathcal{M} can be used to identify the optimal preparation procedure.

approximately 0.5, which is in agreement with theoretical expectations in Section 5.4.2. For a maximally correlated initial state the simulated SE coupling would cause a failure of the evolution with probability 1/2.

To demonstrate the use of preparation fidelity to optimize an experiment we use our experimentally obtained \mathcal{M} to optimise for maximum fidelity for the target $U_s = Z$ and target $U_s = R_Y Z$. The computed preparation fidelity for these targets are shown in Fig. 5.8(a) and Fig. 5.8(b) respectively. The measured value for the IC-norm in these cases was $\|\mathcal{M}\|_{\text{IC}} = 0.062(5)$ and $\|\mathcal{M}\|_{\text{IC}} = 0.034(2)$ for Fig. 5.8(a) and Fig. 5.8(b) respectively. In Fig. 5.8(a), the effect of the environment is minimized for initial projection onto the state

$$\cos(\phi)|H\rangle + e^{i\varphi} \sin(\phi)|V\rangle \quad (5.134)$$

with $\phi \approx 0.658$ and $\varphi \approx 0.252$. This demonstrates that even for nearly uncorrelated SE states, the chosen preparation procedure affects the achieved fidelity. In this example, the projection on the optimal state instead of the basis state $|H\rangle$ improved the fidelity by 0.2%. Similarly, minimizing F_{prep} finds the worst-case preparation, which could give insight into where and why an experimental setup fails.

5.6 Summary

In this chapter we demonstrated that the superchannel formalism can be used to characterize the dynamics of a system that is initially correlated with its environment, a situation which in general cannot be described as a CP-map in the standard quantum channel formalism. In contrast to some previous work on open system dynamics in the presence of

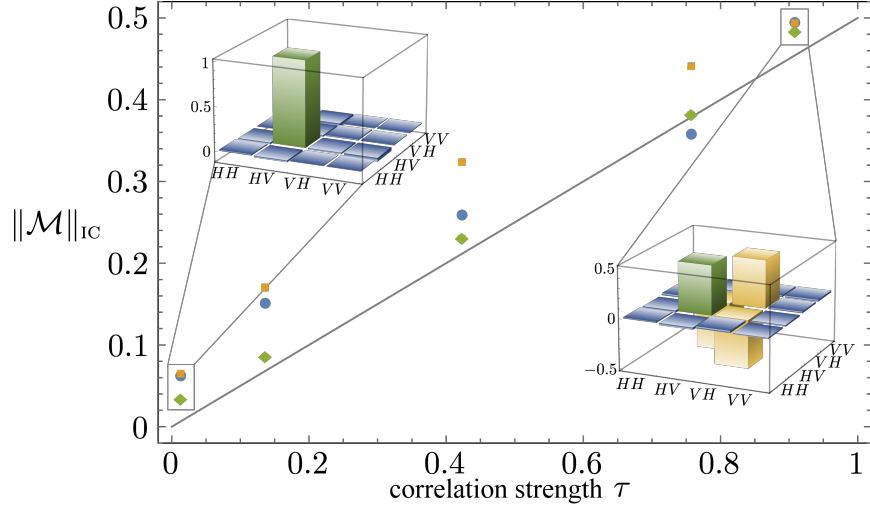


Figure 5.7: Initial correlation norm of the IC superchannel \mathcal{M}_{IC} vs correlation strength of the initial state ρ_{SE} . The correlation strength is given by the tangle τ from Eq. (5.131) and \mathcal{M}_{IC} is formed from a controlled unitary SE interaction with target unitary $U = \sigma_z$ (blue circles), $U = H$ (yellow squares) and $U = R_Y$ (green diamonds). The values of τ were obtained from state tomography of ρ_{SE} for each experiment. The measured real parts of the states with weakest and strongest initial correlations are shown in the respective insets. The solid line corresponds to the IC-norm in the ideal case. Error bars from Poissonian counting statistics are on the order of the symbol size.

initial correlations these characterization techniques operationally significant and experimentally accessible. Notably, the reconstruction of \mathcal{M}_{IC} is a direct generalization of QPT, based on subjecting the system to d^4 , rather than d^2 linearly-independent preparation procedures \mathcal{P}_{ij} , which need not all be fully de-correlating. Therefore, tools developed to improve the efficiency of QPT, such as compressive sensing [SKM⁺11, FGLE12] can also be applied to the reconstruction of \mathcal{M} .

Since the output of a channel is determined solely by the input and the channel itself, it can be thought of as a Markovian two-point connection. The presence of initial correlations, however, demands the use of the superchannel approach, which is thus an important step towards understanding non-Markovian quantum processes. Along these lines, the superchannel approach has recently been used to derive the lower bound on entropy production in a generic quantum process [VM14].

Our technique is most useful in quantum architectures which are strongly coupled to their environment, such as spins in local spin baths. Another application is in quantum

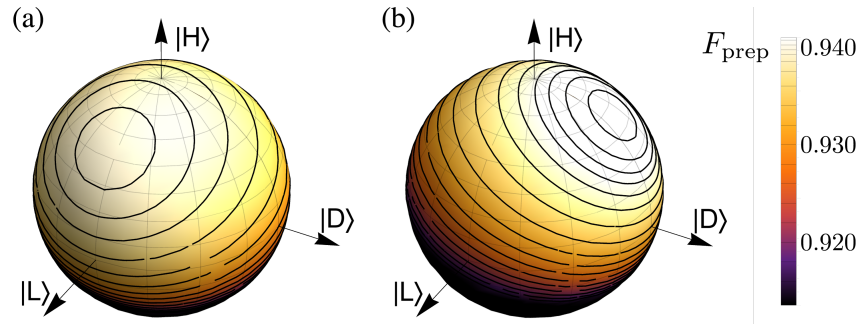


Figure 5.8: Optimization of the preparation procedure. The average preparation fidelity $F_{\text{prep}}(\mathcal{M}, \rho_1, Z)$ for (a) $U_s = Z$ and (b) $U_s = R_Y Z$ is shown as a density plot on the surface of the Bloch sphere of the initial-projection state ρ_1 . In both cases, we chose the lowest strength of initial correlation realized in the experiment to visualize the effect even for very weak SE correlations.

control, where control timescales can be much faster than environmental reset times. Finally, it has been suggested that non-Markovianity can be exploited as a resource [BCM13]; we showed how the superchannel formalism can be used to that extent in our gate optimization.

Chapter 6

Parallel State Initialization in an Ensemble Quantum System

6.1 Introduction

A fundamental challenge to implementing quantum information processing on a physical device is the efficient removal of entropy to rapidly and repeatably initialize a quantum system in a high purity state. This is necessary for the application of quantum error correcting codes to suppress and mitigate the effects of noise and errors that naturally occur in quantum information processors, sensors, and communication devices [Ter15]. Additionally, in spectroscopic applications the signal-to-noise ratio increases significantly with state purity, allowing for the detection of small spin ensembles. A variety of techniques for removing entropy from a quantum system are commonly used, including dynamic nuclear polarization [Abr61, Ram08], strong projective measurements or filters [KLM01], algorithmic cooling [SV99, SMW05, BMR⁺05, RMBL08], optical pumping [WKV⁺10, GPR⁺14], laser cooling [WI79, MMK⁺95, VC00], microwave cooling [VOB⁺06, WRLZ08], and dissipative state engineering [VWC09, LGR⁺13] among others.

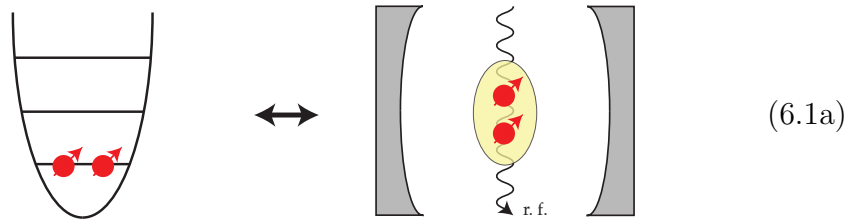
As the number of qubits in quantum devices increases initialization methods that can be implemented in parallel across many qubits are necessary to enable scalability. In the case of spin ensembles, thermal relaxation processes at low temperature naturally initialize all spins in parallel and can be used for state preparation, however to achieve high purity at thermal equilibrium very low temperatures and strong magnetic fields are needed so that thermal excitations are not energetic enough to cause significant transitions out of the ground state. A significant issue is that at low temperatures the time required for the

spin system to reach thermal equilibrium with the environment – the energy relaxation time, T_1 – often becomes very long thus limiting the rate at which spin resets and signal averaging may be applied [Abr61].

Recently, it was demonstrated that a superconducting qubit may be prepared in an arbitrary pure state through sideband cooling by a high quality factor (high-Q) cavity [MVZ+12, GLP+13]. In Sections 6.3 and 6.4 we discuss how similar microwave cooling techniques should also be applicable to ensemble spin systems in magnetic resonance, despite the relatively small coupling between the cavity and a single spin. In particular, we present a theoretical model for how a high-Q resonator (cavity) may be used to actively drive the coupled angular momentum subspaces of an ensemble spin system to a highly pure, non-thermal equilibrium state on a timescale that is significantly shorter than the thermal T_1 . The ability to reduce the effective T_1 time of a spin ensemble by simply applying a detuned microwave drive provides an important tool for error correcting spin-based quantum information processors (for example [BGC12, CJHL09, MTB+08] and references therein), and should also find applications in spectroscopy by permitting faster signal averaging. Similar work has also been considered for cooling an ensemble of nuclear spins by coupling to a mechanical resonator [BW11].

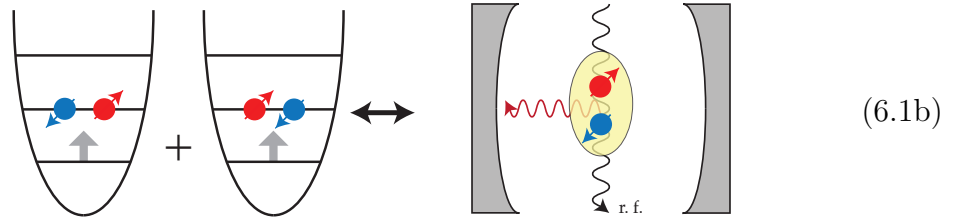
We now provide an illustrative sketch of how a spin ensemble may be actively driven to its ground state via interaction with a dissipative cavity. The coupling of the spins to the cavity allows for the exchange of excitations between the spin and cavity system, and if the cavity is placed in a refrigerator, excitations transferred to the cavity will be dissipated into a larger environment ensuring that the cavity stays close to its ground state, so long as the excitations don't exceed the cooling power of the fridge.

1. Driving the spins with an r.f. field stimulates coherent exchange interactions between the spins and the cavity. If the cavity is initially cold there are no photons.

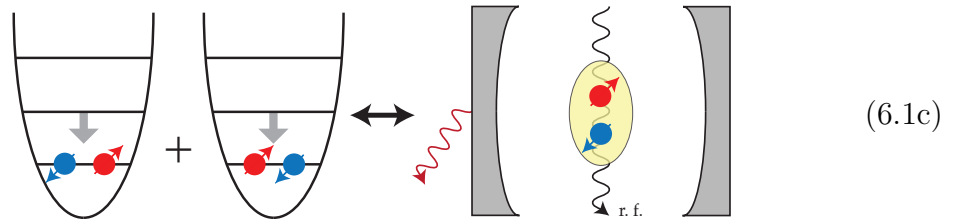


2. An excited spin may flip releasing a photon into the cavity. Due to the two spins being identical particles, the spins will be in a superposition of the cases where either

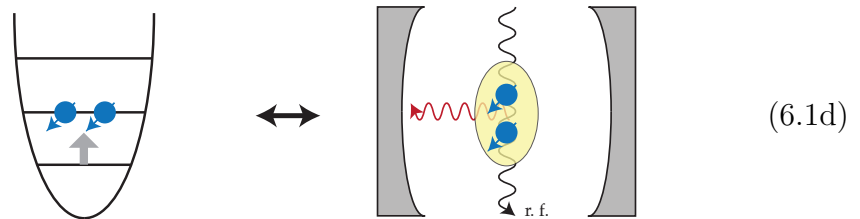
spin flips.



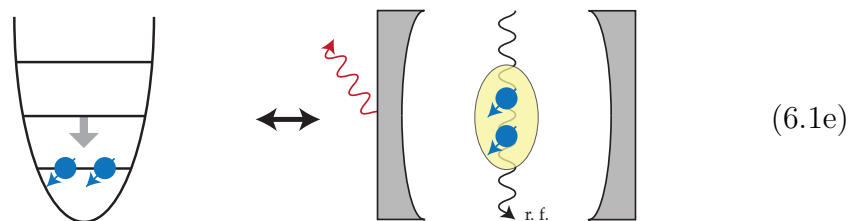
3. If the interaction strength between the spins and cavity is weaker than the dissipation rate of the cavity, the photon will be removed from the cavity before it can interact with the spins again.



4. The remaining excited spin may then flip releasing another photon to the cavity and putting both spins in the ground state.



5. Once this photon is dissipated, both the spins and cavity are both cooled to their ground states and no further exchange interactions take place.



The cavity induced dissipation of a spin ensemble can be engineered to actively cool all the coupled angular momentum subspaces of a spin ensemble to their corresponding low temperatures states. In this context the coupled angular momentum subspaces are the irreducible representations of the Hilbert space for N two-dimensional systems which preserve a global $SU(2)$ symmetry which we discuss in more detail in [Section 6.2.1](#). Under this symmetry the state space has a block diagonal structure with each block corresponding to an irreducible representation of $SU(2)$ that behaves as an effective spin- J particle [[WM02](#), [BCG10](#)]. The largest effective spin subspace has $J = N/2$ and is called the totally symmetric subspace, or Dicke subspace [[Dic54](#)], and contains all states which are permutation invariant across all N spins. The cavity cooling dissipator we derive in [Section 6.4](#) acts collectively on the spin ensemble, hence preserving the global $SU(2)$ symmetry, and so does not couple these subspaces. Hence if cooling from a maximally mixed initial state, the population in each subspace will be trapped in that subspace's ground state and the final state will be a mixed state consisting of the sum of all ground states, which we discuss in [Section 6.4.4](#). A challenge with solving the dynamics of cavity cooling an ensemble quantum system while including an interaction which couples between these subspaces, the goal of which is to enable cooling to the true ground state of the Dicke subspace. The challenge is that once such an interaction is included, in principle the full Hilbert space for N spin-1/2 particles must be included which has dimension 4^N .

In [Section 6.5](#) we present a new method for cavity cooling to the ground state of the Dicke subspace by introducing a local dephasing (T_2) noise process on each spin in the ensemble. This acts to distinguish each spin and thus breaks the collective symmetry of the ensemble. In practice dephasing is a phenomenological description that can arise from several different physical processes, for example in solid-state spin systems it can be caused by inhomogeneous broadening, or spurious coupling to neighbouring spins or a local spin bath [[Abr61](#)]. The effects of local dissipation on collective dynamics has been considered previously, where it was shown to rapidly decohere coherent states in the Dicke subspace [[BCG10](#)]. Decoherence of the Dicke subspace has also been studied for an inhomogeneously broadened ensemble of qubits coupled to a cavity [[KWM11](#)]. Unlike these previous studies we are able to use local dissipation as a resource for dissipative quantum state engineering to the ground state of the Dicke subspace – the dissipative term breaks the $SU(2)$ symmetry of the spin ensemble and couples subspaces, but does not inhibit the cavity cooling dissipative process.

Similar theoretical results have been presented for cooling an ensemble of nuclear spins by coupling to the motion of a nanoscale mechanical resonator [[BW11](#)]. There it was shown that the addition of a chemical shift to each spin was in principle sufficient to break the symmetry and achieve exponential relaxation to the ground state and this was numerically

demonstrated for five spins. A similar result was found numerically in [TW05] where superradiant (Dicke subspace) and subradiant (non-Dicke subspaces) relaxation rates were simulated for 10 inhomogeneous broadened qubits. Our work differs from the approaches in [TW05, BW11] by including the symmetry breaking mechanism as a dissipative term leading to a Lindblad master equation that we can solve perturbatively to derive an analytic expression for the cooling dynamics in the regime where the first order perturbation term dominates the dynamics. In order to solve the master equation we develop a novel new perturbation theory technique for dissipative evolution. This involves applying the Magnus expansion [Mag54, BCOR09], or Average Hamiltonian Theory [HW68, Hae76, EBW+87], in an imaginary time dissipative interaction frame to the superoperators describing the open system evolution. There is a long history in nuclear magnetic resonance of applying average Hamiltonian theory in the superoperator picture, called Average Liouvillian Theory [LDB92, GEB99, Gho00], and also the related cumulant expansion approach for stochastic noise processes [Kub62, CHHC06]. In both these cases the relevant interaction frame is defined by a Hamiltonian, which typically corresponds to a sequence of control pulses, and the average affect of a dissipative term in this frame is assessed. Our approach extends these formalisms by providing a procedure to apply these techniques in purely dissipative (non-periodic) interaction frame. The utility of this method is that in a system with multiple decoherence mechanisms one can find the average effective dissipation of one mechanism in the presence of another. As we demonstrate, this may then be used for finding the equilibrium state of the system for dissipative state engineering applications.

To analytically solve the lowest order term in the Magnus expansion we use recently introduced techniques for describing the local superoperators on spin ensembles using $SU(4)$ algebra generators [Har12, XTH13]. In this representation the dynamics preserve a global $SU(4)$ symmetry and any thermal state of Hamiltonian with $SU(2)$ symmetry will be entirely contained in the totally symmetric subspace of $SU(4)$. By considering the explicit matrix representation of the N qubit totally symmetric subspace of $SU(4)$, this approach allows us to numerically simulate the reduced dynamics of the spin ensemble for up to $N = 100$ spins on a desktop computer.

6.2 Collective Representations of Ensemble Quantum Systems

In many applications of quantum information we are interested in the dynamics of moderate to large size ensembles of identical quantum systems. For example, in NMR and ESR we have ensembles of the order of 10^3 to 10^{18} identical spins. In other proposals based on

solid-state spin ensembles, such as nitrogen-vacancy centres in diamond and phosphorus or bismuth defects in silicon, experiments are interested in ensembles of tens to hundreds of spins. Even for a modest number of spins it becomes impossible to model or simulate the full range of dynamics on the ensemble system on a classical computer due to the dimension of the Hilbert space. For an ensemble of N d -dimensional quantum systems, pure states are specified by $d^N - 1$ complex numbers, while density matrices require $d^{2N} - 1$ complex numbers.

In many situations it is not necessary to consider the full Hilbert space of the system. In particular, if the Hamiltonians or dissipators generating the system dynamics have a global symmetry the state space can be decomposed into a direct sum structure where each of these subspaces can be treated independently. In this case for states that have a block diagonal structure with respect to the subspace decomposition, the dynamics of each block can be considered independent of the others and treated separately. In this section we detail how this may be done for the case of $SU(d)$ symmetries, and in particular for permutation invariant states and operations.

Consider an ensemble of N d -dimensional quantum systems. The joint state space of the ensemble is given by the tensor product space

$$\mathcal{X} = \bigotimes_{j=1}^N \mathcal{X}_j \cong \mathbb{C}^{d^N} \quad (6.2)$$

where each $\mathcal{X}_j \cong \mathbb{C}^d$. Note that to model the state space of density matrices rather than state vectors, we can consider the joint state space as the tensor product of N d^2 dimensional vector space by working in the vectorized representation from [Section 2.2.2](#).

The tensor product Hilbert space in [Eq. \(6.2\)](#) is isomorphic to a direct sum decomposition where each vector space in the direct sum is an *irreducible representation* (irrep) of $SU(d)$. This decomposition is given by

$$\mathcal{X} \cong \bigoplus_J \left(\bigoplus_{k=1}^{m_J} \mathcal{Z}_k^J \right) \quad (6.3)$$

where $J = J_1 \dots J_{d-1}$ is a list of $d - 1$ quantum numbers which index the irrep $\mathcal{Z}^J \cong \mathbb{C}^{d_J}$, and k is a label to index the m_J copies of a given irrep. For a given N and d , the values of J , the dimension d_J , and number of copies m_J of the irrep \mathcal{Z}^J can be computed using *Young tableaux* [[Ful97](#)]. We discuss specific cases for two and four-dimensional subsystems in [Sections 6.2.1](#) and [6.2.2](#) respectively. The largest dimension representation is called

the *totally symmetric subspace*. This subspace consists of all permutation invariant states on the N systems and in quantum optics and atomic physics is often called the *Dicke subspace* [Dic54, Gar11]. We discuss representations of this subspace in Section 6.2.3.

The utility of this decomposition is that any operators acting on \mathcal{X} which preserve a *global $SU(d)$ symmetry* also preserves the structure of the direct-sum representation in Eq. (6.3). Let Π_k^J be the projector onto the subspace \mathcal{Z}_k^J . If we have a density matrix $\rho \in D(\mathcal{X})$ which is block diagonal with respect to the decomposition in Eq. (6.3) then it may be written as

$$\rho = \bigoplus_J \left(\bigoplus_{k=1}^{m_J} p_k^J \rho_k^J \right) \quad (6.4)$$

where

$$\rho_j^{\bar{A}} = \Pi_k^J \rho \Pi_k^J, \quad p_k^J = \text{Tr}[\Pi_k^J \rho]. \quad (6.5)$$

Any operator acting on \mathcal{X} with a $SU(d)$ symmetry will not couple between the different \mathcal{Z}_k^J subspaces. Hence each ρ_k^J can be treated as an isolated system on $D(\mathbb{C}^{d_J})$ thus greatly reducing the dimensionality of calculations and computations.

In practice the presence of a $SU(d)$ symmetry can be inferred from the Hamiltonians and dissipators of the system. This is done by determining whether or not they can be written in terms of the collective operators of $SU(d)$. Recall that the group $SU(d)$ is generated by $d^2 - 1$ traceless and hermitian matrices $\{t_\alpha : \alpha = 1, \dots, d^2 - 1\}$. In the case of $d = 2$ these are the familiar Pauli matrices $t_\alpha \equiv \sigma_\alpha$. Let $t_\alpha^{(j)}$ be the generator matrix acting on subsystem j . We define the collective operator T_α acting on the joint N -partite system by

$$T_\alpha = \sum_{j=1}^N t_\alpha^{(j)} \quad (6.6)$$

Any Hamiltonian or Lindblad dissipator acting on the joint system that can be written in terms of collective operators

$$H = \sum_{\alpha} \omega_{\alpha} T_{\alpha}, \quad \mathcal{D} = \sum_{\alpha} \gamma_{\alpha} D [T_{\alpha}], \quad (6.7)$$

will preserve the subspace structure of Eqs. (6.3) and (6.4). In this case we can simulate each subspace independently by choosing an explicit d_J dimensional representation for the operators T_α acting on the irrep \mathcal{Z}_J . In the following sections we give explicit constructions for these representations for the case of the totally symmetric subspace of a d -dimensional system, and for $SU(2)$.

6.2.1 Dicke Model

We now discuss the so-called *Dicke model*, which arises when one considers an ensemble of N identical spin-half particles. In this case the Hilbert space $(\mathbb{C}^2)^{\otimes N}$ can be decomposed as per Eq. (6.3) into irreps of $SU(2)$, which may be indexed by a single quantum number J . This number can be thought of as the *total spin* of the irrep, and thus each irrep is isomorphic to the state space of a spin- J particle. Hence the dimension of the irrep is given by $d_J = 2J + 1$, with allowed values of J

$$J \in \left\{ \frac{N}{2}, \frac{N}{2} - 1, \frac{N}{2} - 2, \dots, J_{min} \right\} \quad (6.8)$$

where $J_{min} = 0$ if N is even, and $J_{min} = \frac{1}{2}$ if N is odd.

The multiplicity of a given irrep is given by

$$m_J = \frac{2J + 1}{\frac{N}{2} + J + 1} \binom{N}{\frac{N}{2} + J} \quad (6.9)$$

For example, we list the decomposition of $\mathcal{X}_N = (\mathbb{C}^2)^{\otimes N}$ for $N = 2, 3, 4, 5$ using shorthand $\mathcal{Z}^J = J$ we have

$$\mathcal{X}_2 \cong 1 \oplus 0 \quad (6.10)$$

$$\mathcal{X}_3 \cong \frac{3}{2} \oplus \frac{1}{2} \oplus \frac{1}{2} \quad (6.11)$$

$$\mathcal{X}_4 \cong 2 \oplus 1 \oplus 1 \oplus 1 \oplus 0 \oplus 0 \quad (6.12)$$

$$\mathcal{X}_5 \cong \frac{5}{2} \oplus \frac{3}{2} \oplus \frac{3}{2} \oplus \frac{3}{2} \oplus \frac{3}{2} \oplus \frac{1}{2} \oplus \frac{1}{2} \oplus \frac{1}{2} \oplus \frac{1}{2} \quad (6.13)$$

The total spin operators on the full Hilbert space are given by the sum of the spin operators on individual subsystems:

$$J_\alpha = \sum_{j=1}^N \frac{1}{2} \sigma_\alpha^{(j)}, \quad J_\pm = \sum_{j=1}^N \sigma_\pm^{(j)} \quad (6.14)$$

for $\alpha = x, y, z$ where σ_α are Pauli matrices. These operators preserve the $SU(2)$ symmetry of the state space and hence do not couple irreps in the decomposition of the state space given by Eq. (6.3). In this case we may consider each irrep separately and represent these as the spin operators on that subspace.

Consider an irrep \mathcal{Z}_k^J , we can label the eigenstates of this subspace in terms of the eigenstates of $2J + 1$ dimensional representation of the Spin-Z operator $|J, m_z\rangle$ where $m_z = -J, \dots, J$. When we need to distinguish between multiple copies of irreps with the same J value we can add the subscript k to the basis: $|J, m_z, k\rangle$. An explicit representation for spin operators is then given by

$$J_z = \sum_{m_z=-J}^J m_z |J, m_z\rangle\langle J, m_z| \quad (6.15)$$

$$J_{\pm} = \sum_{m_z=-J}^J \sqrt{J(J+1) - m_z(m_z \pm 1)} |J, m_z \pm 1\rangle\langle J, m_z|. \quad (6.16)$$

Sometimes it is convenient to work in the *excitation number* basis of the system. In this case we define the excitation number to be $s \equiv m_z + J$, $|s\rangle \equiv |J, s - J\rangle$, and we have that $|s\rangle$ is the computational basis for a $2J + 1$ dimensional system:

$$J_z = \sum_{s=0}^{2J} (-J + s) |s\rangle\langle s| \quad (6.17)$$

$$J_+ = \sum_{s=0}^{2J} \sqrt{(2J - s)(s + 1)} |s + 1\rangle\langle s| \quad (6.18)$$

$$J_- = \sum_{s=0}^{2J} \sqrt{(2J - s + 1)s} |s - 1\rangle\langle s|. \quad (6.19)$$

One may also write the action of these operators on the full Hilbert space, however to do so involves finding an explicit representation for the basis vectors $|J, m_z, k\rangle$ in terms of the computational basis $|i_1, \dots, i_N\rangle$. This is rather tedious for all but the largest of the subspaces and involves computing Clebsch-Gordan coefficients. However in many cases it is not necessary as we are interested in dynamics which preserve the irrep subspace structure.

In the case of the Dicke subspace one may construct a basis for this subspace by considering all symmetric combinations of states with m excitations in the spin ensemble.

For example, in the case of $N = 4$ we have

$$|2, -2\rangle = |\downarrow\downarrow\downarrow\downarrow\rangle \quad (6.20)$$

$$|2, -1\rangle = \frac{1}{\sqrt{4}} \left(|\uparrow\downarrow\downarrow\downarrow\rangle + |\downarrow\uparrow\downarrow\downarrow\rangle + |\downarrow\downarrow\uparrow\downarrow\rangle + |\downarrow\downarrow\downarrow\uparrow\rangle \right) \quad (6.21)$$

$$|2, 0\rangle = \frac{1}{\sqrt{6}} \left(|\uparrow\uparrow\downarrow\downarrow\rangle + |\uparrow\downarrow\uparrow\downarrow\rangle + |\uparrow\downarrow\downarrow\uparrow\rangle + |\downarrow\uparrow\uparrow\downarrow\rangle + |\downarrow\uparrow\downarrow\uparrow\rangle + |\downarrow\downarrow\uparrow\uparrow\rangle \right) \quad (6.22)$$

$$|2, 1\rangle = \frac{1}{\sqrt{4}} \left(|\uparrow\uparrow\uparrow\downarrow\rangle + |\uparrow\uparrow\downarrow\uparrow\rangle + |\uparrow\downarrow\uparrow\uparrow\rangle + |\downarrow\uparrow\uparrow\uparrow\rangle \right) \quad (6.23)$$

$$|2, 2\rangle = |\uparrow\uparrow\uparrow\uparrow\rangle \quad (6.24)$$

We discuss the symmetric subspace in more detail (and for arbitrary dimension) in [Section 6.2.3](#).

Open System Evolution in the Dicke Model

If we consider mixed states of a spin ensemble which are block diagonal with respect to the $SU(2)$ decompositions then we have

$$L(\mathcal{X}_N) \mapsto \bigoplus_J \left(\bigoplus_{k=1}^{m_J} L(\mathcal{Z}_k^J) \right) \quad (6.25)$$

where in this representation we have thrown out all elements of the density matrix corresponding to coherences between different \mathcal{Z}_k^J . In this case a block diagonal density matrix may be written as

$$\rho = \bigoplus_J \left(\bigoplus_{k=1}^{m_J} p_k^J \rho_k^J \right) \quad (6.26)$$

where

$$\rho_k^J = \sum_{n,m=-J}^J \langle J, m, k | \rho_k^J | J, n, k \rangle |J, m, k\rangle \langle J, n, k| \quad (6.27)$$

$$p_k^J = \sum_{m=-J}^J \langle J, m, k | \rho_k^J | J, m, k \rangle. \quad (6.28)$$

Suppose we have a system with internal Hamiltonian given by $H_0 = \omega J_z$ at some temperature T so that the thermal state of a single spin is given by

$$\rho_{th} = p|\uparrow\rangle\langle\uparrow| + (1-p)|\downarrow\rangle\langle\downarrow|, \quad \text{where } p = \frac{1}{1 + e^{\frac{\omega}{k_b T}}}. \quad (6.29)$$

The thermal state of the ensemble will be block diagonal with respect to the $SU(2)$ subspace structure. The population of a state $\rho_{J,m_z} = |J, m_z\rangle\langle J, m_z|$ is given by [WM02]

$$p_{J,m} = M_J p^{\frac{N}{2}+m_z} (1-p)^{\frac{N}{2}-m_z} \quad (6.30)$$

where M_j is the multiplicity of a subspace with total spin J given in Eq. (6.9). In particular, for a given value of p the population contained in the Dicke subspace is

$$p_{\frac{N}{2}} = \frac{1}{2p-1} (p^{N+1} - (1-p)^{N+1}). \quad (6.31)$$

Using the probabilities $p_{J,m}$, and the decomposition of the spin operators in Eq. (6.16) we can simulate the evolution of each subspace independently for operators that preserve the irrep structure. This includes any Lindblad equations of the form

$$\frac{d\rho(t)}{dt} = \sum_{\alpha} \left(-i\omega_{\alpha}[J_{\alpha}, \rho(t)] + \gamma_{\alpha} \left(J_{\alpha}\rho(t)J_{\alpha}^{\dagger} - \frac{1}{2}(\{J_{\alpha}^{\dagger}J_{\alpha}, \rho(t)\}) \right) \right)$$

Note that local Lindblad dissipators, such as each individual spin experiencing a local T_1 or T_2 process, do *not* satisfy this form. In these cases we have dissipators of the form $\sum_{j=1}^N D[\sigma_{\alpha}^{(j)}]$, which do not preserve the Block diagonal structure of the $SU(2)$ decomposition. In fact, these operators preserve an $SU(4)$ symmetry in terms of the vectorized states space that superoperators act on which we discuss in Section 6.2.2.

6.2.2 Generalized Dicke Model

Just as the Dicke model is the decomposition of the state space of N spin-half particles into irreps of $SU(2)$, we can consider a generalized model where we do this for the Liouville space of the density matrices for N spin-half particles. Recall from Section 2.2 that we can vectorize density matrices $\rho \in L(\mathbb{C}^d)$ to form column vectors $|\rho\rangle\rangle \in \mathbb{C}^{d^2}$. In the case of $d = 2$ this corresponds to $|\rho\rangle\rangle \in \mathbb{C}^4$. Hence we can decompose the tensor product space $(\mathbb{C}^4)^{\otimes N}$ into irreps of $SU(4)$. This decomposition is more complicated than the case of $SU(2)$ as

there are 3 quantum numbers which specify each of the irreducible representations, rather than the single number for $SU(2)$ irreps. Rather than explicitly derive this in terms of irreps, we simply give a constructive method for representing each of the generators of $SU(4)$ in terms of superoperators acting on the N subsystems. This approach has been used by [Har12, XTH13] and we use their notation for the $SU(4)$ algebra here.

There are 15 generators for $SU(4)$, which each belong to one of 6 $SU(2)$ subalgebras. Let $\mathbf{O} = \{\mathcal{Q}, \Sigma, \mathcal{M}, \mathcal{N}, \mathcal{U}, \mathcal{V}\}$ be the set of sub algebra operators. For $\mathcal{O} \in \mathbf{O}$:

$$[\mathcal{O}_+, \mathcal{O}_-] = 2\mathcal{O}_3, \quad [\mathcal{O}_3, \mathcal{O}_\pm] = \pm\mathcal{O}_\pm \quad (6.32)$$

Also the pairs of operators (\mathcal{Q}, Σ) , $(\mathcal{M}, \mathcal{N})$, and $(\mathcal{U}, \mathcal{V})$ each commute:

$$[\mathcal{Q}_\alpha, \Sigma_\beta] = [\mathcal{M}_\alpha, \mathcal{N}_\beta] = [\mathcal{U}_\alpha, \mathcal{V}_\beta] = 0 \quad \forall \alpha, \beta \in \{\pm, 3\} \quad (6.33)$$

The remainder of the $SU(4)$ commutation relations are shown in Table 1. Note that only 15 of these operators are linearly independent. In particular $\mathcal{N}_3 = \mathcal{Q}_3 + \Sigma_3 - \mathcal{M}_3$, $\mathcal{U}_3 = \mathcal{M}_3 - \Sigma_3$ and $\mathcal{V}_3 = \mathcal{Q}_3 - \mathcal{M}_3$.

The $SU(4)$ generators may be given an explicit matrix representation in the full Hilbert space in terms of superoperators acting on the vectorized density matrices of N two-dimensional subsystems. This is analogous to how the spin operators, or $SU(2)$ algebra, can be given a $2J + 1$ dimensional matrix representation in terms of the spin operators acting on a spin- J particle. A superoperator acting on N subsystems can be written as $\mathcal{S}_\mathcal{E} = \sum_j B^{(j)*} \otimes A^{(j)}$, and in particular the $SU(4)$ operators may be represented in terms of our single spin operators $\sigma_\pm^{(j)}, \sigma_z^{(j)}, E_\pm^{(j)} = \frac{1}{2}(\mathbb{1} + \sigma_z^{(j)})$, and collective spin operator

	\mathcal{M}_+	\mathcal{M}_-	\mathcal{M}_3	\mathcal{N}_+	\mathcal{N}_-	\mathcal{N}_3
\mathcal{Q}_+	0	$-\mathcal{V}_+$	$-\frac{1}{2}\mathcal{Q}_+$	0	\mathcal{U}_+	$-\frac{1}{2}\mathcal{Q}_+$
\mathcal{Q}_-	\mathcal{V}_-	0	$\frac{1}{2}\mathcal{Q}_-$	$-\mathcal{U}_-$	0	$\frac{1}{2}\mathcal{Q}_-$
\mathcal{Q}_3	$\frac{1}{2}\mathcal{M}_+$	$-\frac{1}{2}\mathcal{M}_-$	0	$\frac{1}{2}\mathcal{N}_+$	$-\frac{1}{2}\mathcal{N}_-$	0
Σ_+	0	\mathcal{U}_-	$-\frac{1}{2}\Sigma_+$	0	$-\mathcal{V}_-$	$-\frac{1}{2}\Sigma_+$
Σ_-	$-\mathcal{U}_+$	0	$\frac{1}{2}\Sigma_-$	\mathcal{V}_+	0	$\frac{1}{2}\Sigma_-$
Σ_3	$\frac{1}{2}\mathcal{M}_+$	$-\frac{1}{2}\mathcal{M}_-$	0	$\frac{1}{2}\mathcal{N}_+$	$-\frac{1}{2}\mathcal{N}_-$	0

	\mathcal{U}_+	\mathcal{U}_-	\mathcal{U}_3	\mathcal{V}_+	\mathcal{V}_-	\mathcal{V}_3
\mathcal{Q}_+	0	$-\mathcal{N}_+$	$-\frac{1}{2}\mathcal{Q}_+$	0	\mathcal{M}_+	$-\frac{1}{2}\mathcal{Q}_+$
\mathcal{Q}_-	\mathcal{N}_-	0	$\frac{1}{2}\mathcal{Q}_-$	$-\mathcal{M}_-$	0	$\frac{1}{2}\mathcal{Q}_-$
\mathcal{Q}_3	$\frac{1}{2}\mathcal{U}_+$	$-\frac{1}{2}\mathcal{U}_-$	0	$\frac{1}{2}\mathcal{V}_+$	$-\frac{1}{2}\mathcal{V}_-$	0
Σ_+	$-\mathcal{M}_+$	0	$\frac{1}{2}\Sigma_+$	\mathcal{N}_+	0	$\frac{1}{2}\Sigma_+$
Σ_-	0	\mathcal{M}_-	$-\frac{1}{2}\Sigma_-$	0	$-\mathcal{N}_-$	$-\frac{1}{2}\Sigma_-$
Σ_3	$-\frac{1}{2}\mathcal{U}_+$	$\frac{1}{2}\mathcal{U}_-$	0	$-\frac{1}{2}\mathcal{V}_+$	$\frac{1}{2}\mathcal{V}_-$	0

	\mathcal{U}_+	\mathcal{U}_-	\mathcal{U}_3	\mathcal{V}_+	\mathcal{V}_-	\mathcal{V}_3
\mathcal{M}_+	0	$-\Sigma_+$	$-\frac{1}{2}\mathcal{M}_+$	\mathcal{Q}_+	0	$\frac{1}{2}\mathcal{M}_+$
\mathcal{M}_-	Σ_-	0	$\frac{1}{2}\mathcal{M}_-$	0	$-\mathcal{Q}_-$	$-\frac{1}{2}\mathcal{M}_-$
\mathcal{M}_3	$\frac{1}{2}\mathcal{U}_+$	$-\frac{1}{2}\mathcal{U}_-$	0	$-\frac{1}{2}\mathcal{V}_+$	$\frac{1}{2}\mathcal{V}_-$	0
\mathcal{N}_+	$-\mathcal{Q}_+$	0	$\frac{1}{2}\mathcal{N}_+$	0	Σ_+	$-\frac{1}{2}\mathcal{N}_+$
\mathcal{N}_-	0	\mathcal{Q}_-	$-\frac{1}{2}\mathcal{N}_-$	$-\Sigma_-$	0	$\frac{1}{2}\mathcal{N}_-$
\mathcal{N}_3	$-\frac{1}{2}\mathcal{U}_+$	$\frac{1}{2}\mathcal{U}_-$	0	$\frac{1}{2}\mathcal{V}_+$	$-\frac{1}{2}\mathcal{V}_-$	0

Table 6.1: Commutation relations for $SU(4)$ algebra.

$J_z = \sum_{j=1}^N \frac{1}{2} \sigma_z^{(j)}$, as follows:

$$\begin{aligned}
\mathcal{Q}_\pm &= \sum_{j=1}^N \left(\sigma_\pm^{(j)} \otimes \sigma_\pm^{(j)} \right) & \mathcal{Q}_3 &= \frac{1}{2} (\mathbb{1} \otimes J_z + J_z \otimes \mathbb{1}) \\
\mathcal{S}_\pm &= \sum_{j=1}^N \left(\sigma_\mp^{(j)} \otimes \sigma_\pm^{(j)} \right) & \mathcal{S}_3 &= \frac{1}{2} (\mathbb{1} \otimes J_z - J_z \otimes \mathbb{1}) \\
\mathcal{M}_\pm &= \sum_{j=1}^N \left(E_+^{(j)} \otimes \sigma_\pm^{(j)} \right) & \mathcal{M}_3 &= \frac{1}{2} \sum_{j=1}^N \left(E_+^{(j)} \otimes \sigma_z^{(j)} \right) \\
\mathcal{N}_\pm &= \sum_{j=1}^N \left(E_-^{(j)} \otimes \sigma_\pm^{(j)} \right) & \mathcal{N}_3 &= \frac{1}{2} \sum_{j=1}^N \left(E_-^{(j)} \otimes \sigma_z^{(j)} \right) \\
\mathcal{U}_\pm &= \sum_{j=1}^N \left(\sigma_\pm^{(j)} \otimes E_+^{(j)} \right) & \mathcal{U}_3 &= \frac{1}{2} \sum_{j=1}^N \left(\sigma_z^{(j)} \otimes E_+^{(j)} \right) \\
\mathcal{V}_\pm &= \sum_{j=1}^N \left(\sigma_\pm^{(j)} \otimes E_-^{(j)} \right) & \mathcal{V}_3 &= \frac{1}{2} \sum_{j=1}^N \left(\sigma_z^{(j)} \otimes E_-^{(j)} \right)
\end{aligned}$$

Note that care must be taken when ordering the tensor product of subsystems in this sum to keep track of whether we are working in the vectorized Hilbert space which acts on $d|\rho^{(1)} \otimes \dots \otimes \rho^{(n)}\rangle$ or the unravelled Hilbert space which acts on $d|\rho^{(1)}\rangle \otimes \dots \otimes |\rho^{(n)}\rangle$ as discussed in [Section 2.5.1](#).

The utility of this approach is we can express many useful open system dynamics of collective spins in this representation, and it allows us to analytically compute certain properties that may be difficult otherwise. In particular local T_2 and T_1 process dissipators may be expressed as

$$\sum_{j=1}^N D[S_z^{(j)}] = \mathcal{M}_3 - \frac{1}{2} \mathcal{Q}_3 - \frac{1}{2} \mathcal{S}_3 - \frac{N}{4} \mathcal{I} \tag{6.34}$$

$$\sum_{j=1}^N D[S_\pm^{(j)}] = 2\mathcal{Q}_\pm + 2\mathcal{Q}_3 - N\mathcal{I} \tag{6.35}$$

and collective T_2 and T_1 process may be expressed as

$$D[J_z] = -2\Sigma_3^2 \quad (6.36)$$

$$\begin{aligned} D[J_\pm] &= (\mathcal{U}_\pm + \mathcal{V}_\pm)(\mathcal{M}_\pm + \mathcal{N}_\pm) \\ &\quad - \frac{1}{2}(\mathcal{U}_\mp + \mathcal{V}_\mp)(\mathcal{U}_\pm + \mathcal{V}_\pm) \\ &\quad - \frac{1}{2}(\mathcal{M}_\mp + \mathcal{N}_\mp)(\mathcal{M}_\pm + \mathcal{N}_\pm). \end{aligned} \quad (6.37)$$

Further, to do simulations we can write the reduced dimensional representations of these operators acting only on a given irrep. In practice we will only need to consider the totally symmetric subspace, as all thermal state of operators with drift Hamiltonians of the form $H = \sum_\alpha \omega_\alpha J_\alpha$ are within this subspace. This is useful as the dimension of the totally symmetric subspace for N four-dimensional subsystems is

$$d_{sym} = \frac{1}{6}(N+1)(N+2)(N+3) \quad (6.38)$$

and hence scales as $O(N^3)$ rather than 4^N , which enables simulation of open system dynamics preserving this symmetry with up to 100 spins on a typical desktop computer. We describe how one can construct explicit representations of these operators in [Section 6.2.3](#).

6.2.3 Symmetric Subspace

The totally symmetric subspace consists of all permutation invariant states on the N d -dimensional vector spaces. In this section we describe how one may construct a basis for this subspace in terms of the basis for the full tensor product Hilbert space, and also how to construct projections of operators on the tensor product Hilbert space into operators acting only on the subspace.

Let $\mathcal{X} = \bigotimes_{j=1}^N \mathcal{X}_j \cong \mathbb{C}^{Nd}$ be the joint vector space of N copies of $\mathcal{X}_j = \mathbb{C}^d$, and let $\pi \in S_N$ be a permutation of a set with N elements. We can define a unitary $W_\pi \in L(\mathcal{X})$ which swaps the subsystems of \mathcal{X} according to the permutation π . These permutations have an intuitive represented in terms of the graphical calculus in [\[CGB⁺10\]](#) as a tensor network of N wires which connect the respective input and output locations of the vector spaces. For example for $N = 2$ we have that $\pi \in \{12, 21\}$ which corresponds to either an identity (12) or a SWAP gate (21)

$$W_{12} = \begin{array}{c} \text{---} \\ \text{---} \end{array}, \quad W_{21} = \begin{array}{c} \diagdown \quad \diagup \\ \diagup \quad \diagdown \end{array} \quad (6.39)$$

The projector into the symmetric subspace is given by the sum over all permutations π :

$$\Pi_{sym} = \frac{1}{N!} \sum_{\pi \in S_N} W_\pi \quad (6.40)$$

Let \mathcal{X}_{sym} be the totally symmetric subspace of \mathcal{X} . We can construct a basis for \mathcal{X}_{sym} by constructing all permutation invariant combinations of the basis vectors for \mathcal{X} .

For example, if we have $N = 3, d = 2$, we can construct a permutation invariant vector from the multipartite basis element $|0, 0, 1\rangle$ by

$$\Pi_{sym}|0, 0, 1\rangle = \frac{1}{\sqrt{3}} (|0, 0, 1\rangle + |0, 1, 0\rangle + |1, 0, 0\rangle) \quad (6.41)$$

In particular we have that the ordering of the index $\{0, 0, 1\}$ doesn't matter:

$$\Pi_{sym}|0, 0, 1\rangle = \Pi_{sym}|0, 1, 0\rangle = \Pi_{sym}|1, 0, 0\rangle \quad (6.42)$$

Hence we can label the basis elements $|s_j\rangle$ of \mathcal{X}_{sym} by the sorted sequences of the basis element to be projected:

$$|s_0\rangle \equiv \Pi_{sym}|0, 0, 0\rangle = |0, 0, 0\rangle \quad (6.43)$$

$$|s_1\rangle \equiv \Pi_{sym}|0, 0, 1\rangle = \frac{1}{\sqrt{3}} (|0, 0, 1\rangle + |0, 1, 0\rangle + |1, 0, 0\rangle) \quad (6.44)$$

$$|s_2\rangle \equiv \Pi_{sym}|0, 1, 1\rangle = \frac{1}{\sqrt{3}} (|0, 1, 1\rangle + |1, 0, 1\rangle + |1, 1, 0\rangle) \quad (6.45)$$

$$|s_3\rangle \equiv \Pi_{sym}|1, 1, 1\rangle = |1, 1, 1\rangle \quad (6.46)$$

In general we can construct the basis elements \mathcal{X}_{sym} by $\Pi_{sym}|j_1, \dots, j_N\rangle$. Let \underline{j} denote the sequence j_1, \dots, j_N , and let I_N be the set of all sequences \underline{i} satisfying $j_1 \leq j_2 \leq \dots \leq j_N$. The requirement that $j_k \leq j_{k+1}$ ensures that we don't over count sequences that project to the same basis element. To write out a specific function for converting a sequence \underline{j} to the integer $\zeta(\underline{j}) \in \mathbb{N}$ for the symmetric subspace basis vector we can count the number of orderings to arrive at the formula

$$\zeta(j_1, \dots, j_N) = \sum_{i=1}^N \left[\sum_{k=1}^{j_i - j_{i-1}} \binom{N - i + d - j_{i-1} - k}{d - j_{i-1} - k} \right] \quad (6.47)$$

$$= \sum_{i=1}^N \frac{1}{(N - i + 1)!} \left[\frac{(N - i + d - j_{i-1})!}{(d - j_{i-1} - 1)!} - \frac{(N - i + d - j_i)!}{(d - j_i - 1)!} \right] \quad (6.48)$$

$$(6.49)$$

where we define $j_0 \equiv 0$.

With this ordering we can define the basis vector for the sequence $\underline{j} = j_1, \dots, j_N$ as

$$|s_{\zeta(\underline{j})}\rangle = \frac{1}{\sqrt{n(\underline{j})}} \left(\sum_{\text{perm}(\underline{j})} |\underline{j}\rangle \right) \quad (6.50)$$

where we sum over all unique permutations of the indicies j_1, \dots, j_N . The number of elements of these permutations is given by the multinomial

$$n(\underline{j}) = \binom{N}{\underline{j}(0), \dots, \underline{j}(d-1)} = \frac{N!}{\underline{j}(0)! \dots \underline{j}(d-1)!} \quad (6.51)$$

where $\underline{j}(l)$ is the number of l 's in the sequence j_1, \dots, j_N where $l = 0, \dots, d-1$.

Using this ordering we can calculate the dimension of the symmetric subspace by

$$d_{sym} = \dim(\mathcal{X}_{sym}) \quad (6.52)$$

$$= 1 + \zeta(d-1, \dots, d-1) \quad (6.53)$$

$$= \frac{(N+d-1)!}{N!(d-1)!} \quad (6.54)$$

$$= \binom{N+d-1}{d-1} \quad (6.55)$$

States and Operators in the Symmetric Subspace

Let $|\psi\rangle \in \mathcal{X}$ be a state in \mathcal{X} , and let $A \in L(\mathcal{X})$ be an operator acting on \mathcal{X} . $|\psi\rangle$ is a state in the symmetric subspace if and only if $\Pi_{sym}|\psi\rangle = |\psi\rangle$, and the operator A is in the symmetric operator subspace $A \in D(\mathcal{X}_{sym})$ if and only if $\Pi_{sym} A \Pi_{sym} = \rho$.

Two useful classes of operators are permutation invariant tensor products of the form $|\psi_N\rangle = |\psi\rangle^{\otimes N}$ and $A_N = A^{\otimes N}$. We now describe how to construct explicit representations for these operators in \mathcal{X}_{sym} in terms of the basis $\{|s_j\rangle\}$ without explicitly computing the

projection of the full vector or matrix with Π_{sym} . We have

$$|\psi_N\rangle_{sym} = \sum_{j=0}^{d_{sym}-1} |j\rangle \langle s_j | \Psi \rangle \quad (6.56)$$

$$= \sum_{\underline{j} \in I_N} |\zeta(\underline{j})\rangle \langle \underline{j} | \Pi_{sym} | \psi_N \rangle \quad (6.57)$$

$$= \sum_{\underline{j} \in I_N} |\zeta(\underline{j})\rangle \langle \underline{j} | \psi_N \rangle \quad (6.58)$$

$$= \sum_{\underline{j} \in I_N} |\zeta(\underline{j})\rangle \left(\prod_{k=1}^N \langle j_k | \psi \rangle \right) \quad (6.59)$$

$$(6.60)$$

Similarly for tensor product operators $A_N = A^{\otimes N}$ we have

$$A_{N,sym} = \sum_{i,j=0}^{d_{sym}-1} |i\rangle \langle j | \langle s_i | \rho_N | s_j \rangle \quad (6.61)$$

$$= \sum_{\underline{i}, \underline{j} \in I_N} |\zeta(\underline{i})\rangle \langle \zeta(\underline{j}) | \langle \underline{i} | \Pi_{sym} \rho_N \Pi_{sym} | \underline{j} \rangle \quad (6.62)$$

$$= \sum_{\underline{i}, \underline{j} \in I_N} |\zeta(\underline{i})\rangle \langle \zeta(\underline{j}) | \langle \underline{i} | \rho_N | \underline{j} \rangle \quad (6.63)$$

$$= \sum_{\underline{i}, \underline{j} \in I_N} |\zeta(\underline{i})\rangle \langle \zeta(\underline{j}) | \left(\prod_{k=1}^N \langle i_k | \rho | j_k \rangle \right) \quad (6.64)$$

$$(6.65)$$

Symmetric Subspace Preserving Operators

For an arbitrary operator $A \in L(\mathcal{X})$, we have that A preserves the symmetric subspace structure of \mathcal{X} if and only if it does not take any state in the symmetric subspace out of the subspace. This is true if and only if A commutes with the projector onto \mathcal{X}_{sym} . To see this let $|\psi\rangle \in \mathcal{X}_{sym}$, then

$$[A, \Pi_{sym}] = 0 \Rightarrow A|\psi\rangle = A \Pi_{sym} |\psi\rangle \quad (6.66)$$

$$= \Pi_{sym} A |\psi\rangle \in \mathcal{X}_{sym} \quad (6.67)$$

and

$$A|\psi\rangle \in \mathcal{X}_{sym} \Rightarrow A\Pi_{sym}|\psi\rangle = A|\psi\rangle = \Pi_{sym}A|\psi\rangle \quad (6.68)$$

$$\Rightarrow [A, \Pi_{sym}] = 0. \quad (6.69)$$

One useful type of operator which preserves the symmetric subspace is the symmetric sum operator $\Sigma[A] = \sum_{k=1}^N A^{(k)}$, where $A^{(1)} \equiv A \otimes \mathbb{1} \otimes \dots \otimes \mathbb{1}$, and similarly for $A^{(2)}, \dots, A^{(N)}$. Unlike the symmetric product operators, this operator also has support on the other subspaces. Hence if we wish to simulate the evolution of the symmetric subspace alone we must project this operator onto this subspace. The projection of this operator onto $L(\mathcal{X}_{sym})$ is given by

$$\Sigma[A]_{sym} = \sum_{k=1}^N \sum_{i,j=0}^{d_{sym}-1} |i\rangle\langle j| \langle s_i | A^{(k)} | s_j \rangle \quad (6.70)$$

$$(6.71)$$

Since the basis vectors $|s_j\rangle$ consist of all permutations of a sequence of indices \bar{j} they act on each $A^{(k)}$ identically. Hence

$$\Sigma[A]_{sym} = N \sum_{i,j=0}^{d_{sym}-1} |i\rangle\langle j| \langle s_i | A^{(k)} | s_j \rangle \quad (6.72)$$

for any $k = 1, \dots, N$. Hence we have

$$\Sigma[A]_{sym} = N \sum_{i,j=0}^{d_{sym}-1} |i\rangle\langle j| \langle s_i | A^{(k)} | s_j \rangle \quad (6.73)$$

$$= N \sum_{\underline{i}, \underline{j} \in I_N} \langle \underline{i} | \Pi_{sym} A^{(k)} \Pi_{sym} | \underline{j} \rangle |\zeta(\underline{i})\rangle\langle\zeta(\underline{j})| \quad (6.74)$$

$$= N \sum_{\underline{i}, \underline{j} \in I_N} \sum_{k,l=1}^N \left(\frac{\langle \underline{i}_{\bar{k}} | \underline{j}_{\bar{l}} \rangle}{\sqrt{n(\underline{i})n(\underline{j})}} \langle i_k | A | j_l \rangle \right) |\zeta(\underline{i})\rangle\langle\zeta(\underline{j})| \quad (6.75)$$

Symmetric Subspace Representation of $SU(4)$

Of particular use will be the symmetric subspace representations of the $SU(4)$ generators given in [Section 6.2.2](#). We have that

$$\begin{aligned}
 \mathcal{Q}_\pm &= \Sigma [S_\pm \otimes S_\pm] & \mathcal{Q}_3 &= \Sigma \left[\frac{1}{2} (\mathbb{1} \otimes S_z + S_z \otimes \mathbb{1}) \right] \\
 S_\pm &= \Sigma [S_\mp \otimes S_\pm] & \Sigma_3 &= \Sigma \left[\frac{1}{2} (\mathbb{1} \otimes S_z - S_z \otimes \mathbb{1}) \right] \\
 \mathcal{M}_\pm &= \Sigma [E_+ \otimes S_\pm] & \mathcal{M}_3 &= \Sigma [E_+ \otimes S_z] \\
 \mathcal{N}_\pm &= \Sigma [E_- \otimes S_\pm] & \mathcal{N}_3 &= \Sigma [E_- \otimes S_z] \\
 \mathcal{U}_\pm &= \Sigma [S_\pm \otimes E_+] & \mathcal{U}_3 &= \Sigma [S_z \otimes E_+] \\
 \mathcal{V}_\pm &= \Sigma [S_\pm \otimes E_-] & \mathcal{V}_3 &= \Sigma [S_z \otimes E_-]
 \end{aligned}$$

Note that by construction these operators act on the *unravalled* vectorized density matrices.

6.3 Spin-Cavity Master Equation

Now that we have described the state space structure of ensemble quantum systems we return to the physical system of interest, that of an ensemble spin system coupled to a single mode cavity. The mathematical model we use to describe this quantum system is called the Tavis-Cummings (TC) Hamiltonian [[TC68](#), [TC69](#)]

$$H_{TC} = \sum_{j=1}^N g(\sigma_+^{(j)} a + \sigma_-^{(j)} a^\dagger) = g(J_+ a + J_- a^\dagger) \quad (6.76)$$

which is the N -spin generalization of the Jaynes-Cummings Hamiltonian. Where $J_\pm = \sum_{j=1}^N \sigma_\pm^{(j)}$ are the total angular momentum spin raising and lowering operators for an ensemble of N spins. The TC Hamiltonian is also known as the Dicke model [[Dic54](#)] and has been studied extensively for quantum optics (for a recent review see [[Gar11](#)]). In the following section we show how one may derive an effective TC-interaction for an inductively driven ensemble of non-interacting spin-1/2 particles quantized in a large static magnetic field and magnetically coupled to a high-Q cavity as illustrated in [Fig. 6.1](#).

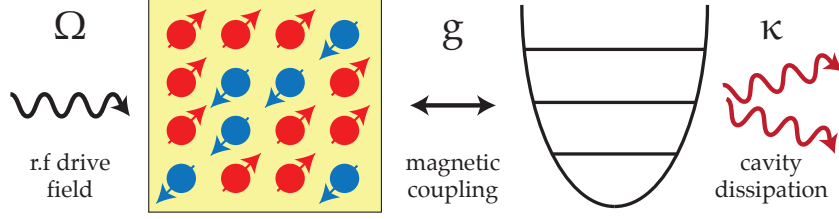


Figure 6.1: Illustration of the driven Tavis-Cummings Hamiltonian for an ensemble of spins identically coupled to a single mode cavity represented as a quantum harmonic oscillator.

In the presence of the drive the ensemble of spins interact with the cavity via coherent radiative processes and may be treated quantum mechanically as a single collective magnetic dipole coupled to the cavity [BP70]. The driven spin-cavity system is described by the Hamiltonian $H = H_0 + H_R(t) + H_I$, with

$$H_0 = \omega_c a^\dagger a + \omega_s J_z \quad (6.77)$$

$$H_R(t) = 2\Omega \cos(\omega_s t) J_x \quad (6.78)$$

$$H_I = 2g(a^\dagger + a) J_x, \quad (6.79)$$

where $a^\dagger(a)$ are the creation (annihilation) operators describing the cavity, Ω is the strength of the drive field (Rabi frequency), ω_c is the resonant frequency of the cavity, ω_s is the Larmor resonance frequency the spins, and g is the coupling strength of the cavity to a single spin in the ensemble in units of $\hbar = 1$.

The Hamiltonian H preserves a global $SU(2)$ symmetry on the spin-ensemble, and hence preserves the subspace structure of any block diagonal state with respect to the decomposition in terms of the irreps of $SU(2)$ discussed in Section 6.2.1. Under this decomposition we may treat each subspace separately as an effective spin- J particle independently interacting with the cavity as illustrated in Fig. 6.2. Let us consider a single irrep of the $SU(2)$ decomposition with a total spin value of J . The eigenstates of H_0 are the tensor product of photon-number states for the cavity and spin states of collective angular momentum in the J_z direction:

$$|n\rangle_c |J, m_z\rangle_s \quad \text{where} \quad n = 0, 1, 2, \dots \quad (6.80)$$

$$m_z = -J, -J + 1, \dots, J - 1, J$$

The collective excitation number of the joint system is given by $N_{ex} = a^\dagger a + (J_z + J)$.

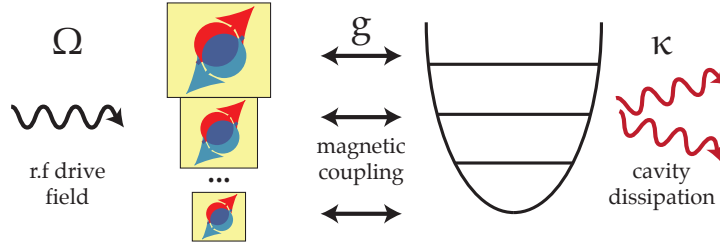


Figure 6.2: Illustration of the effective interaction of the Tavis-Cummings Hamiltonian after the state space of the spin ensemble has been decomposed into irreducible representations of $SU(2)$.

The interaction term H_I commutes with N_{ex} , and hence preserves the total excitation number of the system. It drives transitions between the state $|n\rangle_c |J, m_z\rangle_s$ and states $|n+1\rangle_c |J, m_z-1\rangle_s$ and $|n-1\rangle_c |J, m_z+1\rangle_s$ at a rate of $\sqrt{(n+1)(J(J+1) - m_z(m_z-1))}$ and $\sqrt{n(J(J+1) - m_z(m_z+1))}$, respectively.

If we move into a rotating frame defined by $H_1 = \omega_s(a^\dagger a + J_z)$, the spin-cavity Hamiltonian is transformed to

$$\tilde{H}^{(1)} = e^{itH_1} H_{sc} e^{-itH_1} - H_1 \quad (6.81)$$

$$= \delta\omega a^\dagger a + \Omega J_x + g(a^\dagger J_- + a J_+) \quad (6.82)$$

where $\delta\omega = \omega_c - \omega_s$ is the detuning of the drive from the cavity resonance frequency, and we have made the standard rotating wave approximation (RWA) to remove any time-dependent terms in the Hamiltonian [Abr61]. We note that the interaction term $g(a^\dagger J_- + a J_+)$ in Eq. (6.82) is what is typically referred to as the TC Hamiltonian. The rotating-wave approximation (RWA) used here is valid when the resonant frequencies of the cavity and spin ensemble, ω_c, ω_s , are larger than the inverse time scale we are interested in. In our case this time scale will be dictated by the dissipation rate for the cavity κ , and the interaction Hamiltonian frequency g . Hence we require $\omega_c, \omega_s \gg \kappa, g$.

By appropriately choosing the detuning $\delta\omega$, we can define an effective interaction that in the drive frame is equivalent to the exchange interaction in Eq. (6.82), but in the J_x eigenbasis rather than the J_z eigenbasis. To do this we first move into the interaction frame of $H_2 = \delta\omega a^\dagger a + \Omega J_x$, the Hamiltonian transforms to

$$\tilde{H}^{(2)}(t) = e^{itH_2} H^{(1)} e^{-itH_2} - H_2 \quad (6.83)$$

$$= g e^{itH_2} (a^\dagger J_- + a J_+) e^{-itH_2} \quad (6.84)$$

$$= g e^{i\delta\omega t} a^\dagger \left(e^{i\Omega J_x t} J_- e^{-i\Omega J_x t} \right) + h.c. \quad (6.85)$$

$$= g e^{i\delta\omega t} a^\dagger \left[J_x - i \left(e^{i\Omega J_x t} J_y e^{-i\Omega J_x t} \right) \right] + h.c. \quad (6.86)$$

Next we use the Baker-Campbell-Hausdorff expansion with

$$\text{Ad}_{J_x}^0(J_y) \equiv J_y \quad (6.87)$$

$$\text{Ad}_{J_x}^1(J_y) \equiv [J_x, J_y] \quad (6.88)$$

$$\text{Ad}_{J_x}^n(J_y) \equiv [J_x, \text{Ad}_{J_x}^{n-1}(J_y)] \quad (6.89)$$

to get

$$e^{i\Omega J_x t} J_y e^{-i\Omega J_x t} = \sum_{n=0}^{\infty} \frac{(i\Omega t)^n}{n!} \text{Ad}_{J_x}^n(J_y) \quad (6.90)$$

$$= \sum_{n=0}^{\infty} \frac{(i\Omega t)^{2n}}{(2n)!} \text{Ad}_{J_x}^{2n}(J_y) + \frac{(i\Omega t)^{2n+1}}{(2n+1)!} \text{Ad}_{J_x}^{2n+1}(J_y) \quad (6.91)$$

$$= \cos(\Omega t) J_y + i \sin(\Omega t) J_z \quad (6.92)$$

$$= \frac{1}{2} \left[\cos(\Omega t) (J_+^{(x)} + J_-^{(x)}) + i \sin(\Omega t) (J_+^{(x)} - J_-^{(x)}) \right] \quad (6.93)$$

$$= \frac{1}{2} \left(e^{i\Omega t} J_+^{(x)} + e^{-i\Omega t} J_-^{(x)} \right) \quad (6.94)$$

where $J_{\pm}^{(x)} \equiv J_y \pm iJ_z$ are the spin-ladder operators in the x -basis. Hence we have

$$\tilde{H}^{(2)}(t) = g e^{i\delta\omega t} a^\dagger \left[J_x - \frac{i}{2} \left(e^{i\Omega t} J_+^{(x)} + e^{-i\Omega t} J_-^{(x)} \right) \right] + h.c. \quad (6.95)$$

which may be broken up in terms of frequency components

$$\tilde{H}^{(2)}(t) = H_{0\Omega}(t) + H_{-\Omega}(t) + H_{+\Omega}(t) \quad (6.96)$$

$$H_{0\Omega}(t) = g \left(e^{-i\delta\omega t} a + e^{i\delta\omega t} a^\dagger \right) J_x \quad (6.97)$$

$$H_{-\Omega}(t) = \frac{i g}{2} \left(e^{-i\Delta-t} a J_+^{(x)} - e^{i\Delta-t} a^\dagger J_-^{(x)} \right) \quad (6.98)$$

$$H_{+\Omega}(t) = \frac{i g}{2} \left(e^{-i\Delta+t} a J_-^{(x)} - e^{i\Delta+t} a^\dagger J_+^{(x)} \right) \quad (6.99)$$

where $\Delta_{\pm} = \delta\omega \pm \Omega$.

In analogy to Hartmann-Hahn matching in magnetic resonance cross-relaxation experiments [BS58, HH62, BBGP⁺13] for $\delta\omega > 0$ we may set the cavity detuning to be close to the Rabi frequency of the drive, so that $\Delta = \delta\omega - \Omega$ is small compared to $\delta\omega$. By making a second RWA in the interaction frame of H_2 , the interaction Hamiltonian reduces to the $H_{-\Omega}$ flip-flop exchange interaction between the cavity and spins in the x -basis:

$$H_I(t) = \frac{ig}{2} \left(e^{-i\Delta t} a J_+^{(x)} - e^{i\Delta t} a^\dagger J_-^{(x)} \right). \quad (6.100)$$

This allows for the transfer of population between eigenstates of constant excitation value, while cavity dissipation will act to reduce the excitation number by removing cavity excitations from the system as illustrated in Fig. 6.3. The RWA is valid in the regime where the detuning and Rabi drive strength are large compared to the inverse time scale, t_c , of interest ($\delta\omega, \Omega \gg 1/t_c$). From here we will drop the (x) superscript and just note that we are working in the J_x eigenbasis.

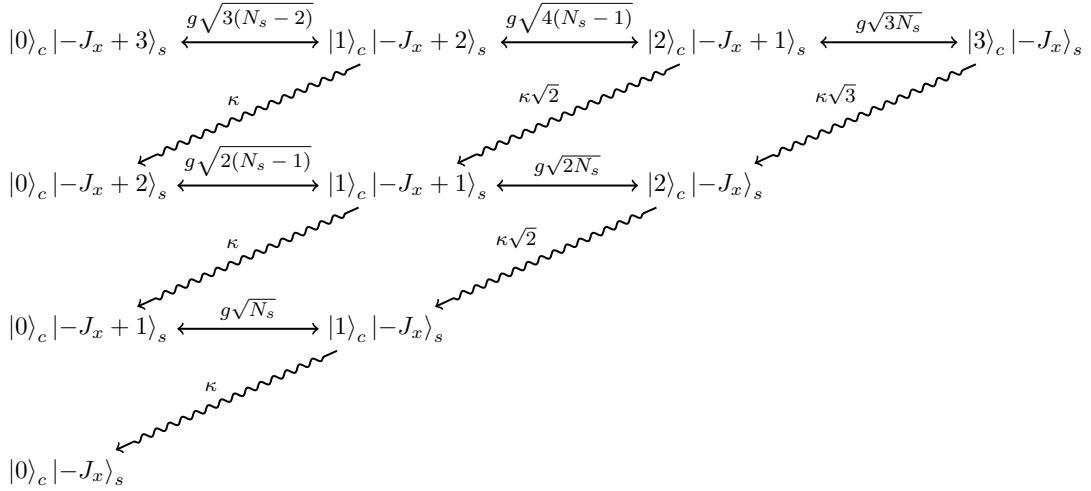


Figure 6.3: Energy level diagram of the joint spin-cavity system with coherent transitions denoted by a solid line and cavity dissipation rates denoted by a curved line. States are labelled as $|n\rangle_c | -J_x + m \rangle_s$, where m is the number of spin excitations and n is the number of cavity excitations.

This resonant coupling is analogous to a collective version of the Purcell effect, which has been previously noted for magnetic resonance systems and is normally small enough to

be neglected. Isolating the spin-cavity exchange interaction allows efficient energy transfer between the two systems, permitting them to relax to a joint equilibrium state in the interaction frame of the control field. The coherent enhancement of the ensemble spin-cavity coupling – similar to the enhancement of the vacuum Rabi frequency for atomic ensembles, but not restricted to the single-excitation manifold [YI99] – enhances spin polarization at a rate that may exceed the thermal relaxation rate.

We note that the spin-cavity exchange coupling also exists in the absence of the Rabi drive, and theoretically permits cooling of the spin system by matching the resonance frequency of the spin system to the cavity resonance. However, in the low-Q case typical in traditional magnetic resonance this process is thermally driven, and thus corresponds to a set of incoherent radiative processes that may not be described by a single Hamiltonian [BP70]. This Purcell effect in magnetic resonance systems has been previously noted and is normally small enough to be neglected [Pur46, Mol69]. However, if one can treat this process using cavity QED [CGB⁺10, MSM⁺12] one may consider Eq. (6.82) in the absence of the drive term ($\Omega = 0$), where our Hamiltonian is then given by the spin-cavity Hamiltonian is given by

$$H_I(t) = \tilde{H}^{(1)} = g \left(e^{-i\Delta t} a J_+ e^{i\Delta t} a^\dagger J_- \right) \quad (6.101)$$

so our exchange interaction occurs in the Z eigenbasis rather than the X eigenbasis of the driven model.

6.3.1 Adiabatic Elimination of a Dissipative Cavity

In this section we consider the master equation of an arbitrary system exchange coupled to a dissipative cavity and show how one may derive an effective master equation for the system by adiabatically eliminating the cavity. This effective master equation can be thought of as the reduced description of a system that is coupled to an environment consisting of a single mode cavity. This follows the approach originally used by [Aga74], however, we phrase it in the modern language of a *time-convolutionless* (TCL) master equation [PB02]. The cavity cooling master equation introduced in Section 6.4 will then be a special case of this general description.

Suppose that the unitary part of the joint system-cavity dynamics is described by the Hamiltonian $H_0 + H_I$, where

$$H_0 = \omega_s H_{s,0} + \omega_c a^\dagger a \quad (6.102)$$

$$H_I = H_{s,1}(a + a^\dagger) \quad (6.103)$$

and that after moving into the interaction frame defined by H_0 we may make a rotating wave approximation so that our interaction frame Hamiltonian is of the form

$$H_I(t) \approx \sum_{\alpha} H_{\alpha}(t) \quad (6.104)$$

$$H_{\alpha}(t) = A_{\alpha}(t)^{\dagger} a + A_{\alpha}(t) a^{\dagger} \quad (6.105)$$

We define the following superoperator for the Liouvillian \mathcal{L} describing the unitary portion of the system evolution

$$\mathcal{L}[H] = -i(\mathbb{1} \otimes H + H^T \otimes \mathbb{1}) \quad (6.106)$$

$$\mathcal{L}[H] : \rho(t) \mapsto -i[H, \rho] \quad (6.107)$$

and Lindblad dissipator $D[A]$ describing the non-unitary evolution:

$$D[A] = \bar{A} \otimes A - \frac{1}{2}(\mathbb{1} \otimes A^{\dagger}A + A^T \bar{A} \otimes \mathbb{1}) \quad (6.108)$$

$$D[A] : \rho \mapsto A \rho A^{\dagger} - \frac{1}{2}\{A^{\dagger}A, \rho\}, \quad (6.109)$$

The open system evolution of the system and cavity is then described by the Lindblad master equation

$$\frac{d}{dt}|\rho(t)\rangle\rangle = (\mathcal{L}_I(t) + \mathcal{D}_c)|\rho(t)\rangle\rangle \quad (6.110)$$

where $\mathcal{L}_I(t) = \mathcal{L}[H_I(t)]$ is the superoperator describing evolution under the Hamiltonian $H(t)$, and \mathcal{D}_c describes the quality factor of the cavity phenomenologically as a photon amplitude damping channel[Aga74]:

$$\mathcal{D}_c = \frac{\kappa}{2} \left((1 + \bar{n}) D[a] + \bar{n} D[a^{\dagger}] \right), \quad (6.111)$$

where

$$\bar{n} = \text{Tr}[a^{\dagger} a \rho_{eq}] \quad (6.112)$$

characterizes the equilibrium temperature of the bath, and κ is the cavity dissipation rate ($\propto 1/Q$). The expectation value of the number operator at equilibrium is related to the temperature, T_c , of the bath by

$$\bar{n} = (e^{\omega_c/k_B T} - 1)^{-1} \Leftrightarrow T_c = \frac{\omega_c}{k_B} \left[\ln \left(\frac{1 + \bar{n}}{\bar{n}} \right) \right]^{-1} \quad (6.113)$$

where k_B is the Boltzmann constant.

We now wish to eliminate the cavity from our description of the dynamics and find an effective Lindblad master equation describing the reduced evolution of the system alone. We may do this using the TCL master equation formalism [PB02]. First we move into the interaction frame defined by the dissipator \mathcal{D}_c . The interaction superoperators and density matrices in this frame are given by

$$\tilde{\mathcal{S}}(t) = e^{-t\mathcal{D}_c} \mathcal{S}(t) e^{t\mathcal{D}_c} \quad (6.114)$$

$$|\tilde{\rho}(t)\rangle\rangle = e^{-t\mathcal{D}_c} |\rho(t)\rangle\rangle. \quad (6.115)$$

Hence we have that the joint system-cavity master equation in the dissipator interaction frame is

$$\frac{d}{dt} |\tilde{\rho}(t)\rangle\rangle = \tilde{\mathcal{L}}_I(t) |\tilde{\rho}(t)\rangle\rangle \quad (6.116)$$

We then define a projection operator \mathcal{P} onto the relevant degrees of freedom for our reduced system

$$\mathcal{P} |\rho(t)\rangle\rangle \equiv |\rho_s(t) \otimes \rho_{eq}\rangle\rangle \quad (6.117)$$

where

$$\rho_s(t) = \text{Tr}_c [\rho(t)] \quad (6.118)$$

$$\Rightarrow |\rho_s(t)\rangle\rangle = \mathcal{S}_{\text{Tr}_c} |\rho(t)\rangle\rangle \quad (6.119)$$

$$\mathcal{D}_c |\rho_{eq}\rangle\rangle = 0. \quad (6.120)$$

ρ_{eq} is the equilibrium state of the cavity under the dissipator \mathcal{D}_c . In the case of weak coupling between the system and cavity with respect to the cavity dissipation rate, the second order TCL master equation is given by [PB02]

$$\frac{d}{dt} \mathcal{P} |\tilde{\rho}(t)\rangle\rangle = \int_0^{t-t_1} d\tau \mathcal{P} \tilde{\mathcal{L}}_I(t) \tilde{\mathcal{L}}_I(t-\tau) \mathcal{P} |\tilde{\rho}(t)\rangle\rangle. \quad (6.121)$$

We now explicitly consider the interaction frame of the dissipator. To do this we use the definition of the adjoint channel \mathcal{D}_c^\dagger :

$$\mathcal{D}[A]^\dagger = A^T \otimes A^\dagger - \frac{1}{2} (\mathbb{1} \otimes A^\dagger A + A^T \bar{A} \otimes \mathbb{1}). \quad (6.122)$$

The adjoint channel has the following useful properties:

$$\mathcal{D}_c^\dagger |\mathbb{1}\rangle\rangle = 0, \quad \mathcal{D}_c^\dagger |a\rangle\rangle = -\frac{\kappa}{2} |a\rangle\rangle, \quad \mathcal{D}_c^\dagger |a^\dagger\rangle\rangle = -\frac{\kappa}{2} |a^\dagger\rangle\rangle \quad (6.123)$$

$$e^{t\mathcal{D}_c^\dagger} |\mathbb{1}\rangle\rangle = |\mathbb{1}\rangle\rangle, \quad e^{t\mathcal{D}_c^\dagger} |a\rangle\rangle = e^{-\frac{\kappa}{2}t} |a\rangle\rangle, \quad e^{t\mathcal{D}_c^\dagger} |a^\dagger\rangle\rangle = e^{-\frac{\kappa}{2}t} |a^\dagger\rangle\rangle. \quad (6.124)$$

Hence we have that

$$e^{-t\mathcal{D}_c^\dagger} \mathcal{S}_{\text{Tr}_c}^\dagger = \mathcal{S}_{\text{Tr}_c}^\dagger \Rightarrow \mathcal{P}|\tilde{\rho}(t)\rangle\rangle = \mathcal{S}_{\text{Tr}_c} e^{-it\mathcal{D}_c} |\rho(t)\rangle\rangle \quad (6.125)$$

$$= \mathcal{S}_{\text{Tr}_c} |\rho(t)\rangle\rangle \quad (6.126)$$

$$= \mathcal{P}\rho(t). \quad (6.127)$$

In addition, $\mathcal{D}_c \mathcal{P}|\rho(t)\rangle\rangle = 0$, and so the reduced dynamics of the spin-ensemble in the interaction frame of the dissipator, Eq. (6.111), is given to 2nd order by the TCL master equation [Bul87]:

$$\frac{d}{dt} |\rho_s(t)\rangle\rangle = \int_0^{t-t_0} d\tau \mathcal{S}_{\text{Tr}_c} \mathcal{L}_I(t) e^{\tau\mathcal{D}_c} \mathcal{L}_I(t-\tau) |\rho_s(t) \otimes \rho_{eq}\rangle\rangle. \quad (6.128)$$

The partial trace superoperator acting on the joint Hamiltonian has cavity terms proportional to $\langle\langle a |, \langle\langle a^\dagger |$. Hence

$$\mathcal{S}_{\text{Tr}_c} L_I(t) e^{\tau\mathcal{D}_c} \cong \left(e^{\tau\mathcal{D}_c^\dagger} L_I(t)^\dagger | \mathbb{1}_c \rangle\rangle \right)^\dagger \quad (6.129)$$

has terms $e^{\tau\mathcal{D}_c^\dagger} |a\rangle\rangle = e^{-\kappa\tau/2} |a\rangle\rangle$, and similarly for $|a^\dagger\rangle\rangle$. Thus, we have

$$\frac{d}{dt} |\rho_s(t)\rangle\rangle = \int_0^{t-t_0} d\tau e^{-\kappa\tau/2} \mathcal{S}_{\text{Tr}_c} \mathcal{L}_I(t) \mathcal{L}_I(t-\tau) |\rho_s(t) \otimes \rho_{eq}\rangle\rangle \quad (6.130)$$

$$\Leftrightarrow \frac{d}{dt} \rho_s(t) = - \int_0^{t-t_0} d\tau e^{-\kappa\tau/2} \text{Tr}_c \left[H_I(t), \left[H_I(t-\tau), \rho_s(t) \otimes \rho_{eq} \right] \right]. \quad (6.131)$$

We now expand this in terms of the component Hamiltonians $H_\alpha(t)$. Define

$$\mathcal{C}_{\alpha,\beta}(t, s) = \text{Tr}_c \left[H_\alpha(t), [H_\beta(s), \rho_s(t) \otimes \rho_{eq}] \right], \quad (6.132)$$

then we have

$$\frac{d}{dt} \rho_s(t) = - \sum_{\alpha,\beta} \int_0^{t-t_0} d\tau e^{-\kappa\tau/2} \mathcal{C}_{\alpha,\beta}(t, t-\tau). \quad (6.133)$$

Using the properties of our cavity equilibrium state

$$\text{Tr}[aa^\dagger \rho_{eq}] = \bar{n} + 1, \quad \text{Tr}[a^\dagger a \rho_{eq}] = \bar{n}, \quad \text{Tr}[a^2 \rho_{eq}] = \text{Tr}[a^{\dagger 2} \rho_{eq}] = 0, \quad (6.134)$$

we have only two contributing terms for each \mathcal{C} . $\mathcal{C}_{\alpha,\beta}(t, s)$ is then given by

$$\mathcal{C}_{\alpha,\beta}(t, s) = \text{Tr}_c \left[A_\alpha(t)^\dagger a, [A_\beta(s) a^\dagger, \rho \otimes \rho_{eq}] \right] + \text{Tr}_c \left[A_\alpha(t) a^\dagger, [A_\beta(s)^\dagger a, \rho \otimes \rho_{eq}] \right] \quad (6.135)$$

$$\begin{aligned} &= (\bar{n} + 1) \left(A_\alpha(t)^\dagger A_\beta(s) \rho + \rho A_\beta(s)^\dagger A_\alpha(t) - A_\beta(s) \rho A_\alpha(t)^\dagger - A_\alpha(t) \rho A_\beta(s)^\dagger \right) \\ &+ \bar{n} \left(A_\alpha(t) A_\beta(s)^\dagger \rho + \rho A_\beta(s) A_\alpha(t)^\dagger - A_\beta(s)^\dagger \rho A_\alpha(t) - A_\alpha(t)^\dagger \rho A_\beta(s) \right) \end{aligned} \quad (6.136)$$

To calculate the dissipator for these terms we assume the Markovian limit and take the upper limit of the integral to infinity $\int_0^{t-t_0} d\tau \rightarrow \int_0^\infty d\tau$. We also define the superoperator generators

$$\mathcal{G}_\alpha(t) = - \int_0^\infty d\tau e^{-\kappa\tau/2} \mathcal{C}_{\alpha,\alpha}(t, t - \tau) \quad (6.137)$$

$$\mathcal{G}_{\alpha,\beta}(t) = - \int_0^\infty d\tau e^{-\kappa\tau/2} \left(\mathcal{C}_{\alpha,\beta}(t, t - \tau) + \mathcal{C}_{\beta,\alpha}(t, t - \tau) \right) \quad (6.138)$$

Our reduced system master equation is then given by

$$\frac{d}{dt} \rho_s(t) = \left(\sum_\alpha \mathcal{G}_\alpha(t) + \sum_{\alpha < \beta} \mathcal{G}_{\alpha,\beta}(t) \right) \rho_s(t). \quad (6.139)$$

We refer to $\mathcal{G}_\alpha(t)$ and $\mathcal{G}_{\alpha,\beta}(t)$ as the diagonal and cross-terms of the master equation, respectively. Generally we need not consider the cross-terms explicitly as we may remove them under certain parameter regimes with an appropriate RWA.

Suppose that the time dependence of the operators $A_\alpha(t)$ is such that $A_\alpha(t) = e^{i\omega_\alpha t} A_\alpha$. Hence $A_\alpha(t - \tau) = e^{-i\tau\omega_\alpha} A_\alpha(t)$, and we have

$$\int_0^\infty d\tau e^{-\kappa\tau/2} e^{\pm i\tau\omega_\alpha} = \frac{2}{\kappa \mp i2\omega_\alpha} = 2 \left(\frac{\kappa \pm 2i\omega_\alpha}{\kappa^2 + 4\omega_\alpha^2} \right) = \gamma_\alpha \pm i\lambda_\alpha \quad (6.140)$$

where

$$\gamma_\alpha = \frac{2\kappa}{\kappa^2 + 4\omega_\alpha^2} \quad \lambda_\alpha = \frac{4\omega_\alpha}{\kappa^2 + 4\omega_\alpha^2}. \quad (6.141)$$

In this case the cross-terms $\mathcal{G}_{\alpha,\beta}$ will still have time dependence of $e^{\pm i(\omega_\alpha - \omega_\beta)t}$. Thus, if we have that $|\omega_\alpha - \omega_\beta| \gg \kappa$ for all α, β , we may make a RWA and disregard these high

frequency terms, yielding

$$\mathcal{G}_\alpha(t)\rho = - \int_0^\infty d\tau e^{-\kappa\tau/2} \mathcal{C}_{\alpha,\beta}(t, t-\tau)\rho \quad (6.142)$$

$$= - \int_0^\infty d\tau e^{-\kappa\tau/2} \left[(\bar{n} + 1) \left(e^{-i\omega_\alpha\tau} (A_\alpha^\dagger A_\alpha \rho - A_\alpha \rho A_\alpha^\dagger) + e^{i\omega_\alpha\tau} (\rho A_\alpha^\dagger A_\alpha - A_\alpha \rho A_\alpha^\dagger) \right) \right. \\ \left. + \bar{n} \left(e^{i\omega_\alpha\tau} (A_\alpha A_\alpha^\dagger \rho - A_\alpha^\dagger \rho A_\alpha) + e^{-i\omega_\alpha\tau} (\rho A_\alpha A_\alpha^\dagger - A_\alpha^\dagger \rho A_\alpha) \right) \right] \quad (6.143)$$

$$= (\bar{n} + 1) \left((\gamma_\alpha - i\lambda_\alpha) (A_\alpha^\dagger A_\alpha \rho - A_\alpha \rho A_\alpha^\dagger) + (\gamma_\alpha + i\lambda_\alpha) (\rho A_\alpha^\dagger A_\alpha - A_\alpha \rho A_\alpha^\dagger) \right) \\ + \bar{n} \left((\gamma_\alpha + i\lambda_\alpha) (A_\alpha A_\alpha^\dagger \rho - A_\alpha^\dagger \rho A_\alpha) + (\gamma_\alpha - i\lambda_\alpha) (\rho A_\alpha A_\alpha^\dagger - A_\alpha^\dagger \rho A_\alpha) \right) \quad (6.144)$$

$$= \gamma_\alpha \left((\bar{n} + 1) \mathcal{D}[A_\alpha] + \bar{n} D[A_\alpha^\dagger] \right) - \lambda_\alpha \mathcal{L} \left[(\bar{n} + 1) A_\alpha^\dagger A_\alpha - \bar{n} A_\alpha A_\alpha^\dagger \right]. \quad (6.145)$$

Since the time dependence of $A_\alpha(t)$ drops out in the dissipator and Liouvilians we finally obtain the master equation generator

$$\mathcal{G}_\alpha(t) = \mathcal{G}_\alpha = \gamma_\alpha \left((\bar{n} + 1) \mathcal{D}[A_\alpha] + \bar{n} D[A_\alpha^\dagger] \right) - \lambda_\alpha \mathcal{L} \left[(\bar{n} + 1) A_\alpha^\dagger A_\alpha - \bar{n} A_\alpha A_\alpha^\dagger \right]. \quad (6.146)$$

6.4 Cavity Cooling Master Equation

We now apply the general derivation of the 2nd order TCL reduced system master equation in [Section 6.3.1](#) to the Rabi-driven TC-Hamiltonian. The Rabi-driven TC-Hamiltonian has three spectral components, a "center-band" rotating with angular frequency $\delta\omega$, and sidebands with angular frequencies $\Delta_\pm = \delta\omega \pm \Omega$:

$$H_I(t) = H_{0\Omega}(t) + H_{-\Omega}(t) + H_{+\Omega}(t) \quad (6.147)$$

$$H_{0\Omega}(t) = g \left(e^{-i\delta\omega t} a + e^{i\delta\omega t} a^\dagger \right) J_x \quad (6.148)$$

$$H_{-\Omega}(t) = \frac{i g}{2} \left(e^{-i\Delta_- t} a J_+^{(x)} - e^{i\Delta_- t} a^\dagger J_-^{(x)} \right) \quad (6.149)$$

$$H_{+\Omega}(t) = \frac{i g}{2} \left(e^{-i\Delta_+ t} a J_-^{(x)} - e^{i\Delta_+ t} a^\dagger J_+^{(x)} \right). \quad (6.150)$$

In this case the cross-terms for the 2nd order TCL master equation will be of frequencies Ω and 2Ω , hence our RWA is valid in the regime where the Rabi drive strength is much

stronger than the dissipation rate ($\Omega \gg \kappa$). In this case we have three contributions to the master equation for the spin-ensemble:

$$\frac{d}{dt}\rho_s(t) = \left(\mathcal{G}_{0\Omega} + \mathcal{G}_{-\Omega} + \mathcal{G}_{+\Omega} \right) \rho_s(t) \quad (6.151)$$

where

$$\mathcal{G}_{0,\pm\Omega} = -\Omega_{0,\pm}\mathcal{L}_{0,\pm} + \frac{\Gamma_{0,\pm}}{2}\mathcal{D}_{0,\pm} \quad (6.152)$$

and

$$\begin{aligned} \Gamma_0 &= \frac{4g^2\kappa}{\kappa^2+4\delta\omega^2} & \Omega_0 &= \frac{4g^2\delta\omega}{\kappa^2+4\delta\omega^2} \\ \Gamma_- &= \frac{g^2\kappa}{\kappa^2+4\Delta_-^2} & \Omega_- &= \frac{g^2\Delta_-}{\kappa^2+4\Delta_-^2} \\ \Gamma_+ &= \frac{g^2\kappa}{\kappa^2+4\Delta_+^2} & \Omega_+ &= \frac{g^2\Delta_+}{\kappa^2+4\Delta_+^2} \\ \mathcal{D}_0 &= (2\bar{n}+1)\mathcal{D}[J_x] & \mathcal{L}_0 &= \mathcal{L}[J_x^2] \\ \mathcal{D}_- &= (\bar{n}+1)\mathcal{D}[J_-^{(x)}] + \bar{n}\mathcal{D}[J_+^{(x)}] & \mathcal{L}_- &= \mathcal{L}[(\bar{n}+1)J_+^{(x)}J_-^{(x)} - \bar{n}J_-^{(x)}J_+^{(x)}] \\ \mathcal{D}_+ &= (\bar{n}+1)\mathcal{D}[J_+^{(x)}] + \bar{n}\mathcal{D}[J_-^{(x)}] & \mathcal{L}_+ &= \mathcal{L}[(\bar{n}+1)J_-^{(x)}J_+^{(x)} - \bar{n}J_+^{(x)}J_-^{(x)}] \end{aligned}$$

To achieve cooling to the ground state of each irrep subspace of the ensemble we require that the $\mathcal{G}_{-\Omega}$ term be dominant, which implies $\Gamma_- \gg \Gamma_+, \Gamma_0$. If we assume that our Rabi drive and cavity detuning are matched, ($\Omega \approx \delta\omega$), then in the regime where our RWA is valid ($\Omega \gg \kappa$), we have

$$\frac{\Gamma_+}{\Gamma_-} = \frac{\kappa^2 + 4\Delta_-^2}{\kappa^2 + 4\Delta_+^2} \approx \frac{\kappa^2}{\kappa^2 + 16\Omega^2} \ll 1 \quad (6.153)$$

and

$$\frac{\Gamma_0}{\Gamma_-} = 4\frac{\kappa^2 + 4\Delta_-^2}{\kappa^2 + 4\delta\omega^2} \approx 4\frac{\kappa^2}{\kappa^2 + 4\Omega^2} \ll 1 \quad (6.154)$$

and so $\mathcal{G}_{-\Omega}$ will be the dominant dissipative term. We also have that $\Omega_0 \approx g^2/\delta\omega \ll 1$, $\Omega_+ \approx g^2/\Delta_+ \ll 1$ since $\delta\omega, \Delta_+ \gg g$.

Thus, we arrive at the Markovian master equation

$$\frac{d}{dt}\rho_s(t) = \left(\Omega_s \mathcal{L}_s + \frac{\Gamma_s}{2} \mathcal{D}_s \right) \rho_s(t) \quad (6.155)$$

where

$$\Omega_s = -\frac{g^2\Delta}{\kappa^2 + 4\Delta^2}, \quad \Gamma_s = \frac{g^2\kappa}{\kappa^2 + 4\Delta^2}. \quad (6.156)$$

Here Ω_s is the frequency the effective Hamiltonian, and Γ_s is the effective dissipation rate of the spin-system.

6.4.1 Subspace Equilibrium State

Let us consider the evolution of an initial spin state which is block diagonal in the $SU(2)$ state space decomposition of the Dicke model discussed in [Section 6.2.1](#). Since we can consider each block of the density matrix separately, suppose for a given subspace with total spin J , that the spin state is diagonal in the angular momentum basis:

$$\rho_J(t) = \sum_m P_{J,m}(t) \rho_{J,m} \quad (6.157)$$

$$P_{J,m}(t) = \langle J, m | \rho(t) | J, m \rangle, \quad (6.158)$$

where $P_{J,m}(t)$ is the probability of measuring the system in the collective spin state $\rho_{J,m} = |J, m\rangle\langle J, m|$ at time t . In this case the master equation, [Eq. \(6.155\)](#), reduces to a rate equation for the populations of the angular momentum states:

$$\frac{d}{dt} P_{J,m}(t) = \Gamma_s \left(A_{J,m+1} P_{J,m+1}(t) + B_{J,m} P_{J,m}(t) + C_{J,m-1} P_{J,m-1}(t) \right) \quad (6.159)$$

where

$$A_{J,m} = (1 + \bar{n}) [J(J+1) - m(m-1)] \quad (6.160)$$

$$C_{J,m} = \bar{n} [J(J+1) - m(m+1)] \quad (6.161)$$

$$B_{J,m} = -(A_{J,m} + C_{J,m}) \quad (6.162)$$

and Γ_s is the cavity cooling rate given in [Section 6.4](#). Defining $\vec{P}_J(t) = (P_{J,-J}(t), \dots, P_{J,J}(t))$, we obtain a matrix differential equation

$$\frac{d}{dt} \vec{P}_J(t) = \Gamma_s M_J \vec{P}_J(t), \quad (6.163)$$

where M is the tridiagonal matrix

$$M_J = \begin{pmatrix} B_{J,-J} & A_{J,-J+1} & 0 & 0 & 0 & \dots & 0 \\ C_{J,-J} & B_{J,-J+1} & A_{J,-J+2} & 0 & 0 & \dots & 0 \\ 0 & C_{J,-J+1} & B_{J,-J+2} & A_{J,-J+3} & 0 & \dots & 0 \\ \vdots & & & \ddots & & & \vdots \\ 0 & & \dots & & 0 & C_{J,J-1} & B_{J,J} \end{pmatrix} \quad (6.164)$$

For a given subspace initial state specified by populations $\vec{P}_J(0)$, [Eq. \(6.163\)](#) has the solution

$$\vec{P}_J(t) = \exp(t \Gamma_s M_J) \vec{P}_J(0). \quad (6.165)$$

The equilibrium state of the spin subspace satisfies $M_J \cdot \vec{P}_J(\infty) = 0$, and is given by $\rho_{J,eq} = \sum_{m=-J}^J P_{J,m}(\infty) \rho_{J,m}$, where

$$P_{J,m}(\infty) = \frac{\bar{n}^{J+m}(1+\bar{n})^{J-m}}{(1+\bar{n})^{2J+1} - \bar{n}^{2J+1}}. \quad (6.166)$$

It corresponds to the non-thermal state

$$\rho_{eq} = \frac{1}{\mathcal{Z}} \exp\left(-\frac{\omega_s}{k_B T_{eq}} J_x\right), \quad T_{eq} = \frac{\omega_s}{k_B} \left[\ln\left(\frac{1+\bar{n}}{\bar{n}}\right) \right]^{-1}. \quad (6.167)$$

where J_x is the spin- J representation of the X -spin operator. By comparing with [Eq. \(6.113\)](#), we have that the final effective temperature reached by the spin ensemble in the interaction frame of the Rabi drive is

$$T_{eq} = \frac{\omega_s}{\omega_c} T_c \quad (6.168)$$

where T_c is the temperature of the cavity. The total spin expectation value for the equilibrium state of spin-ensemble subspace is

$$\langle J_x \rangle_{eq} = -J + \bar{n} - \frac{(2J+1)\bar{n}^{2J+1}}{(1+\bar{n})^{2J+1} - \bar{n}^{2J+1}}. \quad (6.169)$$

In the limit of $N_s \gg \bar{n}$, we have that the ground state population at equilibrium is given by

$$P_{J,-J} \approx \frac{1}{1+\bar{n}} \quad (6.170)$$

and the final expectation value is approximately

$$\langle J_x \rangle_{eq} \approx -J + \bar{n}. \quad (6.171)$$

Thus, the final spin polarization in the spin- J subspace will be roughly equivalent to the thermal cavity polarization.

We note that if the detuning $\delta\omega$ were negative, matching $\Omega = \delta\omega$ would result in the $H_{+\Omega}$ term being dominant in [Eq. \(6.151\)](#), leading to a master equation of the form in [Eq. \(6.155\)](#) but with the operators J_- and J_+ interchanged. The dynamics of this master equation would drive the spin ensemble subspaces towards the $\langle J_x \rangle = J$ state rather than to the $-J$ state. Thus, the detuning must be larger than the cavity linewidth to prevent competition between the $H_{-\Omega}$ and $H_{+\Omega}$ terms, which would drive the spin system to a high entropy thermally mixed state.

6.4.2 Simulation of Subspace Cooling Rate

The tridiagonal nature of the rate matrix in Eq. (6.164) allows Eq. (6.165) to be efficiently simulated for relatively large numbers of spins. For simplicity we will consider the ideal case where the cavity is cooled to its ground state ($\bar{n} = 0$), and the spin-ensemble is taken to be maximally mixed:

$$P_{J,m}(0) = \frac{1}{2J+1} \quad \text{for } m = -J, \dots, J. \quad (6.172)$$

The simulated expectation value of $\langle J_x(t) \rangle$ for the Dicke subspace of ensemble sizes ranging from $N_s = 10^3$ to $N_s = 10^5$ is shown in Fig. 6.4, normalized by $-N_s/2$ to obtain a maximum value of 1. At a value of $-\langle J_x(t) \rangle/J = 1$ the total angular momentum subspace of the spin ensemble is completely polarized to the J_x ground eigenstate $|\frac{N_s}{2}, -\frac{N_s}{2}\rangle$. The

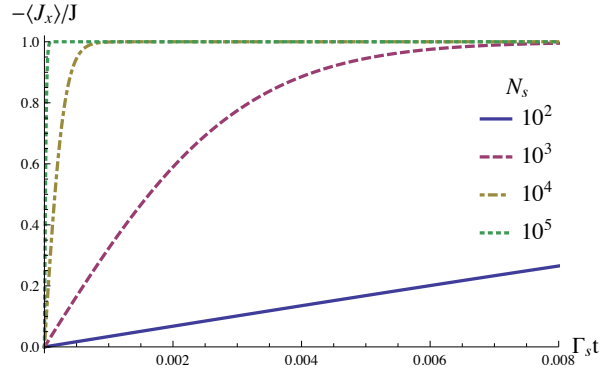


Figure 6.4: Simulated evolution of the normalized expectation value of $-\langle J_x(t) \rangle/J$ for the Dicke subspace of a cavity-cooled spin ensemble containing N_s spins. The time axis is scaled by the effective dissipation rate, Γ_s , for the spin-ensemble given in Section 6.4.

expectation value $\langle J_x(t) \rangle$ may be fit to an exponential to derive an effective cooling time-constant, $T_{1,\text{eff}}$, analogous to the thermal spin-lattice relaxation time, T_1 . A fit to a model given by

$$-\frac{\langle J_x(t) \rangle}{J} = 1 - \exp\left(-\frac{t}{T_{1,\text{eff}}}\right) \quad (6.173)$$

yields the parameters $T_{1,\text{eff}} = \lambda N_s^\gamma / \Gamma_s$ with $\lambda = 2.0406$ and $\gamma = -0.9981$. An approximate expression for the cooling time-constant of the spin- J subspace as a function of J is given

by

$$T_{1,\text{eff}}(J) \approx \frac{1}{\Gamma_s J} = \frac{\kappa^2 + 4\Delta^2}{g^2 \kappa J}, \quad (6.174)$$

Hence, for the symmetric Dicke subspace with $J = N_s/2$, we have

$$T_{1,\text{eff}}^{\text{sym}}(N_s) \approx \frac{2}{\Gamma_s N_s} = \frac{2(\kappa^2 + 4\Delta^2)}{g^2 \kappa N_s}. \quad (6.175)$$

From [Eq. \(6.174\)](#) we find that the cooling efficiency is maximized when the Rabi drive strength is matched to the cavity detuning ($\Delta = 0$). In this case the cooling rate and time-constant simplify respectively to

$$\Gamma_s = \frac{g^2}{\kappa}, \quad T_{1,\text{eff}}(J) = \frac{\kappa}{g^2 J}. \quad (6.176)$$

To achieve this result experimentally, one must choose parameters that adhere to the two RWA's used to isolate the spin-cavity exchange term in [Eq. \(6.100\)](#). Under the condition that $\delta\omega \approx \Omega$, this requires that

$$g\sqrt{N} \ll \kappa \ll \Omega, \delta\omega \ll \omega_c, \omega_s. \quad (6.177)$$

For example, assuming an implementation using X-band pulsed electron spin resonance (ESR) ($\omega_c/2\pi \approx \omega_s/2\pi = 10$ GHz), with samples that typically contain from roughly $N_s = 10^6$ spins to $N_s = 10^{17}$ spins [[BSS+12](#), [EEQ+10](#)], experimentally reasonable values are $\Omega/2\pi = 100$ MHz, $Q = 10^4$ ($\kappa/2\pi = 1$ MHz)[[BC12](#), [BMT+13](#)], and $g/2\pi = 1$ Hz [[BMT+13](#)]. For these parameters, the range of validity of the Markovian master equation is $N_s \ll \kappa^2/g^2 = 10^{12}$ and the Dicke subspace of an ensemble containing roughly 10^{11} electron spins may be polarized with an effective T_1 of $3.18 \mu\text{s}$. This polarization time is significantly shorter than the thermal T_1 for low-temperature electron spin ensembles, which can range from seconds to days [[Abr61](#)].

In the case where the cavity is thermally occupied, the final spin polarization is roughly equal to the thermal cavity polarization in [Eq. \(6.167\)](#), and the fitted effective cooling constant of the Dicke subspace of the spin ensemble as a function of the thermal cavity occupation number is shown in [Fig. 6.5](#).

We find that for cavity temperatures corresponding to $\bar{n} < \sqrt{N_s}$ the effective cooling constant $T_{1,\text{eff}}$ is approximately equal to the zero temperature value. However when $\bar{n} > \sqrt{N_s}$ the effective cooling constant is reduced, and approaches a value of

$$T_{1,\text{eff}} = \frac{1}{2\bar{n}\Gamma_s} \quad (6.178)$$

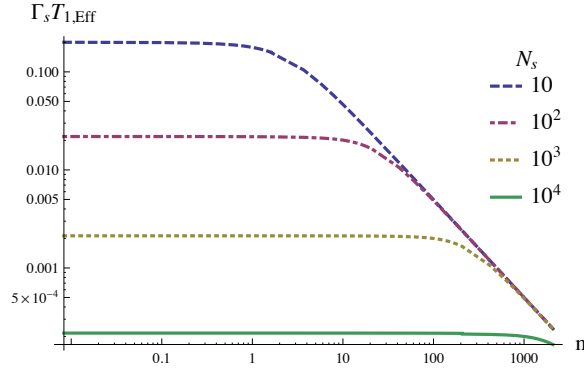


Figure 6.5: Effective cooling time-constant, $T_{1,\text{eff}}$, of the spin ensemble as a function of the equilibrium excitation number of the cavity, \bar{n} , for $N_s = 10, 10^2, 10^3, 10^4$ spins in the ensemble.

This effect appears to originate from the fact that the final spin system polarization will be equal to the cavity polarization. Thus, cooling to a spin temperature that is not fully polarized requires removing fewer photons from the spin system. Given that the spin dissipation rate is independent of the cavity temperature, it takes less time to drive the spins to a state that is not fully polarized. An example of the simulated normalized spin expectation value $-\langle J_x \rangle / J$ as a function of temperature is shown in Fig. 6.6. Here we are considering a cavity with resonant frequency of $\omega_c / 2\pi = 10$ GHz.

6.4.3 Assumptions and Validity of the Markovian Approximation

Several assumptions were made in the presented theoretical model for cavity cooling of a spin ensemble. Firstly, we have assumed that the spin ensemble has no dipolar coupling between spins. This assumption is valid for magnetically dilute samples. Dipolar coupling breaks the degeneracy in the Dicke states and reduces the cooling efficiency by reducing the number of spins effectively interacting with the cavity. Secondly, we have neglected the effects of thermal relaxation of the spin system. As the cooling effect of the cavity on the spin system relies on a coherent spin-cavity information exchange, the relaxation time of the spin system in the frame of the Rabi drive – commonly referred to as $T_{1,\rho}$ – must be significantly longer than the inverse cavity dissipation rate $1/\kappa$. Thirdly, we have assumed that the spin-cavity coupling and Rabi drive are spatially homogeneous across the spin

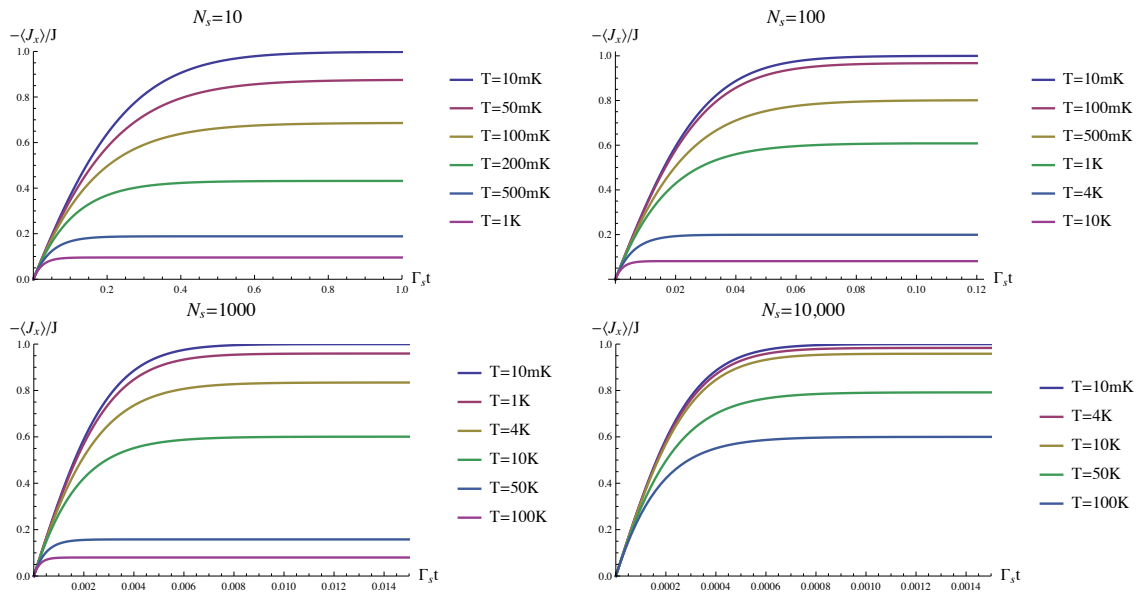


Figure 6.6: Normalized spin expectation value $-\langle J_x \rangle / J$ of the spin ensemble as a function of time for various equilibrium temperatures of the cavity. We consider the case of $N_s = 10, 100, 1000$ and $10,000$ spins in the ensemble, and a cavity with resonant frequency $\omega_c = 10$ GHz, and a Rabi drive on resonance with the detuning $\Delta = \delta\omega - \Omega = 0$.

ensemble. Inhomogeneities would lead to an incoherent distribution of cooling rates over the sample that reduces the cooling efficiency.

The derivation of the Markovian master equation (6.155) assumes that no correlations between the cavity and spin system accrue during the cooling process, such that there is no back action of the cavity dynamics on the spin system. This condition is enforced when the cavity dissipation rate, κ , exceeds the rate of coherent spin-cavity exchange in the lowest excitation manifold by at least an order of magnitude ($\kappa \gg 10g\sqrt{N_s}$). In this Markovian limit, the rate at which spin photons are added to the cavity is significantly less than the rate at which thermal photons are added, meaning the cooling power of the fridge necessary to maintain the thermal cavity temperature is sufficient to dissipate the spin photons without raising the average occupation number of the cavity.

From Eq. (6.174) we see that, in principle, the cooling efficiency could be improved by adding more spins to make κ closer to $g\sqrt{N_s}$, but in this regime the cooling power of the fridge is no longer sufficient to prevent back action from the cavity and non-Markovian effects significantly lower the cooling rate. The validity of the derived cooling rate in Eq. (6.174) depends on the validity of the Markov approximation used to derive the cavity cooling master equation in Eq. (6.155). More concretely, from the spin-cavity energy level diagram for the Dicke subspace shown in Fig. 6.3, the rate of transfer between states

$$|n\rangle_c | -J_x + m \rangle_s \longleftrightarrow |n+1\rangle_c | -J_x + m - 1 \rangle_s \quad (6.179)$$

is given by

$$g\sqrt{m(2J_x + 1 - m)(n + 1)}. \quad (6.180)$$

At the same time, the cavity dissipator of strength

$$\kappa\sqrt{n + 1} \quad (6.181)$$

is acting to drive the spin-cavity system to the state

$$|n\rangle_c | -J_x + m - 1 \rangle_s. \quad (6.182)$$

To satisfy the Markov condition, we require the cooling dynamics to always drive the spin-cavity system toward states of low excitation number (bottom left of Fig. 6.3), without significantly populating states of high excitation number (top right of Fig. 6.3). This will occur if the maximum rates for coherent transfer and cavity dissipation obey the following relationship:

$$\kappa\sqrt{n + 1} \gg g\sqrt{m(2J_x + 1 - m)(n + 1)} \iff \kappa \gg g\sqrt{m(N_s + 1 - m)}. \quad (6.183)$$

This transfer rate is greatest for a maximally excited spin system with $m = 2J_x = N_s$. In this case we recover our condition that

$$\kappa \gg g\sqrt{N_s}. \quad (6.184)$$

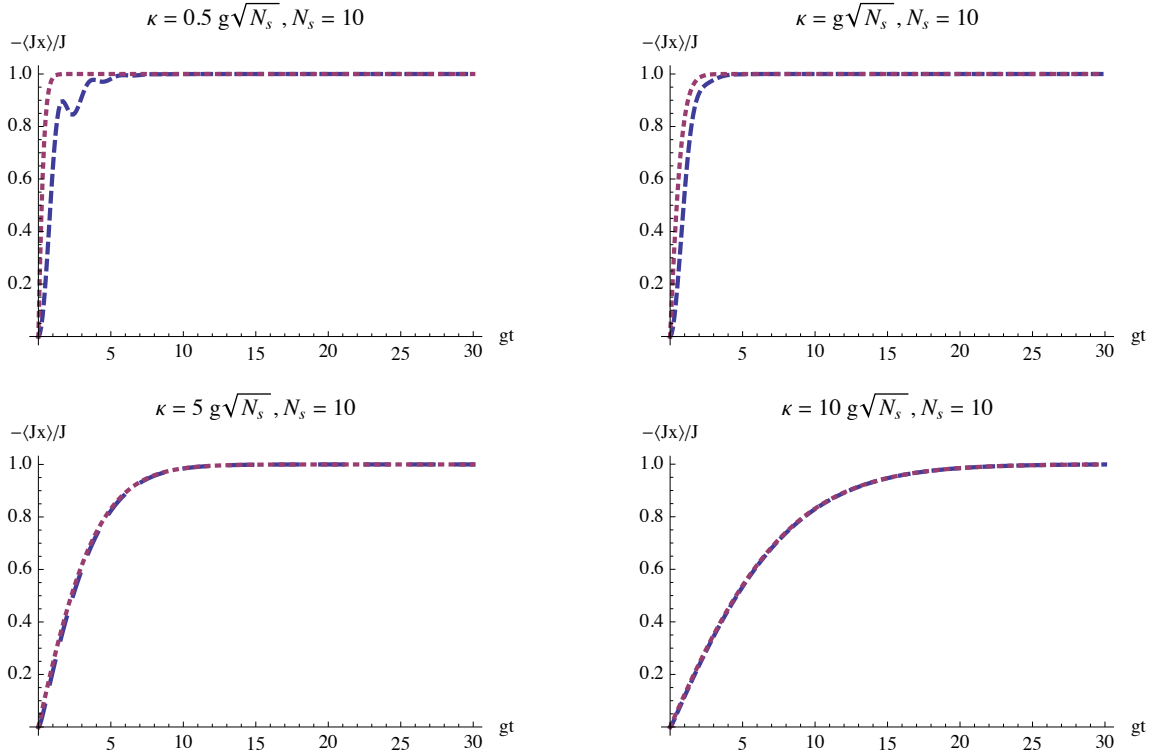


Figure 6.7: Comparison of the cooling dynamics by the Markovian master equation (pink dotted curve) and a full spin-ensemble cavity simulation with RWA (blue dashed curve). The normalized expectation value of $-\langle J_x(t) \rangle / J$ is plotted for $N_s = 10$ spins with $\kappa = 0.5, 1, 5, 10g\sqrt{N_s}$, a cavity temperature $T = 0K$, and Rabi-drive matched to the detuning $\Delta = \delta\omega - \Omega = 0$. When $\kappa \geq 10g\sqrt{N_s}$ the Markovian master equation calculation agrees very well with the full simulation. Also, as predicted by eqn. (29), the cooling rate increases for larger κ , until the point where strong coupling effects take over.

To numerically investigate where the Markovian approximation breaks down we simulated the full evolution for the spin cavity system for $N_s = 10$ spins at zero cavity temperature ($\bar{n} = 0$) with $\delta\omega = \Omega$ in the parameter regime where the RWA is valid and

compared it to the derived master equation for the spin ensemble. For the full evolution we consider

$$|\rho(t)\rangle\rangle = \exp(t(\mathcal{L}[H_I] + \kappa\mathcal{D}[a]))|\rho(0)\rangle\rangle \quad (6.185)$$

where

$$H_I = \frac{Ig}{2} \left(aJ_+^{(x)} + a^\dagger J_-^{(x)} \right) \quad (6.186)$$

$$\rho(0) = \rho_s(0) \otimes \rho_{eq} \quad (6.187)$$

$$= \frac{\mathbb{1}}{2^{N_s}} \otimes |0\rangle\langle 0| \quad (6.188)$$

where $\rho_s(0)$ is the maximally mixed state, and ρ_{eq} is the ground state of the cavity. For simulation we truncate the cavity dimension to be $N_s + 1$, the dimension of the spin-ensemble Hilbert space.

As shown in Fig. 6.7, when $\kappa = 0.5g\sqrt{N_s}$ a full non-Markovian simulation of the cooling procedure yields dynamics that are much richer than predicted by the Markovian model. In particular, coherent transfer of spin photons deposited in the cavity back to the spin system are seen as oscillations in the expectation value of J_x . These memory effects reduce the cooling efficiency such that the cooling rate is initially fast when the cavity occupation is low, then slows down significantly as higher excitations of the cavity are transferred back to the spin system. As κ becomes larger the oscillations are damped out, but the Markovian master equation still does not fully agree with the full non-Markovian simulation until we have $\kappa \approx 10g\sqrt{N_s}$. When $\kappa = 10g\sqrt{N_s}$, the oscillations are critically damped and the Markovian master equation captures the full cooling dynamics. Thus, if the cavity dissipation rate, κ , exceeds the rate of coherent spin-cavity exchange in the single excitation manifold by at least an order of magnitude — i.e. $\kappa \geq 10g\sqrt{N_s}$ — the Markovian master equation is valid.

6.4.4 Cavity Cooling in the Full State Space

The cavity cooling rate in Eq. (6.174) and final magnetization obtained in Eq. (6.169) are valid for each spin- J subspace in the decomposition of the Dicke model into the irreps of $SU(2)$ decomposition. Since the effective T_1 time of a given subspace is inversely proportional to J , larger spin subspaces will be cooled more rapidly. We now explore the average dynamics of the ensemble when the collective cooling of each subspace is taken into account for a maximally mixed initial state on the spin ensemble which has population in every subspace.

Recall that for the Dicke model we have that the state space can be factorized as

$$\mathbb{C}^{2N} \cong \bigoplus_{J=N/2}^{J_{min}} \left(\bigoplus_{k=1}^{\lambda_J} \mathbb{C}^{2J+1} \right) \quad (6.189)$$

where the multiplicity of representations with total spin J is

$$\lambda_J = \frac{2J+1}{\frac{N}{2}+J+1} \binom{N}{\frac{N}{2}+J}. \quad (6.190)$$

In the case that N is even, the lowest value of J is $J_{min} = 0$; if N is odd instead, the lowest value of J is $J_{min} = \frac{1}{2}$. For the remainder of this section we will assume that N is even so that $J_{min} = 0$ for convenience. Since the cavity cooling dissipator does not couple subspaces we may simulate each of these resulting rate equations separately, and sum them to obtain the final magnetization of the spin ensemble. In the ideal case of $\bar{n} = 0$, the final magnetization of a spin- J subspace is given by $\langle J_x \rangle_{J,eq} = -J$. We may include the cooling dynamics in this picture by using the cooling rate derived in [Section 6.4.2](#). Let us assume that the subspace is initially in a maximally mixed state. In this case the expectation of magnetization under cavity cooling is given by

$$\langle J_x(t) \rangle_J = -J [1 - e^{-t\Gamma_s J}]. \quad (6.191)$$

Hence we see that subspaces with a higher J value are cooled to their ground state at a faster rate than lower J subspaces.

To find the final magnetization and cooling rate of the full spin ensemble we may sum all the subspaces weighted by the initial population in that subspace. In this case we have

$$\langle J_x(t) \rangle = \sum_{J=0}^{N/2} -P_J J [1 - e^{-t\Gamma_s J}], \quad (6.192)$$

where for the case of a maximally mixed initial state

$$P_J = \frac{(2J+1)}{2^N} \lambda_J. \quad (6.193)$$

Hence,

$$\langle J_x(t) \rangle = \sum_{J=0}^{N/2} -\frac{J(2J+1)}{2^N} \lambda_J [1 - e^{-t\Gamma_s J}] \quad (6.194)$$

$$= \sum_{J=0}^{N/2} \frac{-J(2J+1)^2}{2^N \left(\frac{N}{2}+J+1\right)} \binom{N}{\frac{N}{2}+J} [1 - e^{-t\Gamma_s J}], \quad (6.195)$$

and the final magnetization is given by

$$\langle J_x \rangle_{eq} = \sum_{J=0}^{N/2} \frac{-J(2J+1)^2}{2^N \left(\frac{N}{2} + J + 1\right)} \binom{N}{\frac{N}{2} + J}. \quad (6.196)$$

Let us consider an example for $N = 10$ spins. In this case the allowed values of J for the subspaces are $J = 5, 4, 3, 2, 1, 0$. With corresponding multiplicities

$$\lambda_5 = 1, \quad \lambda_4 = 9, \quad \lambda_3 = 35, \quad \lambda_2 = 75, \quad \lambda_1 = 90, \quad \lambda_0 = 42. \quad (6.197)$$

In Fig. 6.8 we plot the subspace magnetization under cavity cooling from a maximally mixed initial state as given by Eq. (6.191). This illustrates that higher J value subspaces are cooled faster. We also plot the expectation value for J_x in the full Hilbert space according to Eq. (6.195) and see that the final magnetization and rate is greatly reduced from that of the $J = 5$ Dicke subspace.

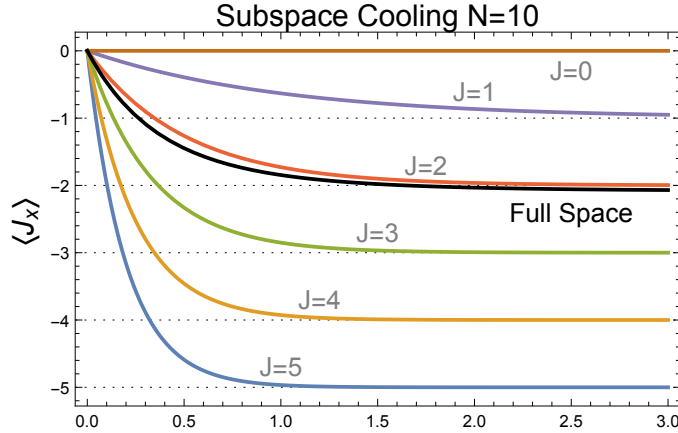


Figure 6.8: Simulation of $\langle J_x(t) \rangle$ for cavity cooling of individual coupled angular momentum subspaces for $N = 10$ spins. The subspaces have J values ranging from 5 for the Dicke subspace, to 0 for the smallest singlet subspaces. We see that the cooling time decreases as the J value of the subspace increases. The black line shows the simulation of $\langle J_x \rangle_{eq}$ in the full Hilbert space by summing over all subspace values weighted by the multiplicity of each J value.

We now derive an approximate expression for the cooling dynamics of the full Hilbert space. If N is large then the only term of Eq. (6.195) that will contribute is the one

proportional to J^3 in the numerator. In this case we have [Eq. \(6.195\)](#) may be approximated by

$$\langle J_x(t) \rangle \approx \sum_{J=0}^{N/2} \frac{-4J^3}{2^N \binom{N}{\frac{N}{2} + J}} \binom{N}{\frac{N}{2} + J} [1 - e^{-t\Gamma_s J}]. \quad (6.198)$$

We can set further bounds on this approximation by noting that $\frac{N}{2} \leq \frac{N}{2} + J + 1 \leq N + 1 \approx N$. Hence

$$\langle J_x(t) \rangle \geq \frac{-8}{2^N N} \sum_{J=0}^{N/2} J^3 \binom{N}{\frac{N}{2} + J} [1 - e^{-t\Gamma_s J}] \quad (6.199)$$

$$\langle J_x(t) \rangle \leq \frac{-4}{2^N N} \sum_{J=0}^{N/2} J^3 \binom{N}{\frac{N}{2} + J} [1 - e^{-t\Gamma_s J}]. \quad (6.200)$$

If we consider the first term, ignoring the exponential piece for now, we have

$$\sum_{J=0}^{N/2} J^3 \binom{N}{\frac{N}{2} + J} = \frac{N}{4} \binom{N}{\frac{N}{2} + 1} \binom{N}{\frac{N}{2} + 1} \approx \frac{N^2}{8} \binom{N}{\frac{N}{2}} \approx 2^N \frac{N}{4} \sqrt{\frac{N}{2\pi}} \quad (6.201)$$

where in the step we have used the asymptotic approximation of the binomial coefficient for large N :

$$\binom{N}{\frac{N}{2}} \approx 2^N \sqrt{\frac{2}{\pi N}}. \quad (6.202)$$

Hence, we have that $-\sqrt{\frac{2N}{\pi}} \leq \langle J_x \rangle_{eq} \leq -\sqrt{\frac{N}{2\pi}}$, so a lower bound on the magnetization is given by

$$\langle J_x \rangle_{eq} \leq -\sqrt{\frac{N}{2\pi}}. \quad (6.203)$$

To estimate the cooling rate we must consider the exponential decay term in [Eq. \(6.195\)](#).

Using the same approximations as before this term is given by

$$\begin{aligned}
\sum_{J=N/2}^{J_{min}} \frac{-J(2J+1)^2}{2^N \left(\frac{N}{2} + J + 1\right) \binom{N}{\frac{N}{2} + J}} e^{-t\Gamma J} &\leq -\frac{4}{2^N N} \sum_{J=0}^{N/2} J^3 \binom{N}{\frac{N}{2} + J} e^{-t\Gamma J} \\
&= \sum_{k=0}^{\infty} \frac{(-t\Gamma)^k}{k!} \left(-\frac{4}{2^N N} \sum_{J=0}^{N/2} J^{3+k} \binom{N}{\frac{N}{2} + J} \right) \\
&= -\sqrt{\frac{N}{2\pi}} \sum_{k=0}^{\infty} \frac{(-t\Gamma)^k}{k!} \left(\frac{4\sqrt{2\pi}}{2^N N^{3/2}} \sum_{J=0}^{N/2} J^{3+k} \binom{N}{\frac{N}{2} + J} \right)
\end{aligned} \tag{6.204}$$

where we have factored out the equilibrium magnetization in the last line. We now compute the sum

$$C_k = \frac{4\sqrt{2\pi}}{2^N N^{3/2}} \sum_{J=0}^{N/2} J^{3+k} \binom{N}{\frac{N}{2} + J} \tag{6.205}$$

for the first few values of k . Using the approximation in [Eq. \(6.202\)](#), and only keeping the term of the highest power of N , we have

$$C_0 \approx 1 \qquad C_1 \approx \frac{3\sqrt{\pi}}{4} \sqrt{\frac{N}{2}} \tag{6.206}$$

$$C_2 \approx 2\sqrt{\frac{N}{2}} \qquad C_3 \approx \frac{15\sqrt{\pi}}{8} \sqrt{\frac{N}{2}}^3 \tag{6.207}$$

$$C_4 \approx 6\sqrt{\frac{N}{2}}^4 \qquad C_5 \approx \frac{105\sqrt{\pi}}{16} \sqrt{\frac{N}{2}}^5 \tag{6.208}$$

$$C_6 \approx 24\sqrt{\frac{N}{2}}^6 \qquad C_7 \approx \frac{945\sqrt{\pi}}{32} \sqrt{\frac{N}{2}}^7 \tag{6.209}$$

$$C_8 \approx 120\sqrt{\frac{N}{2}}^8 \qquad C_9 \approx \frac{10395\sqrt{\pi}}{64} \sqrt{\frac{N}{2}}^9 \tag{6.210}$$

We see here that all terms are proportional to $\sqrt{\frac{N}{2}}^k$, and hence we will have a cooling rate proportional to $\sqrt{\frac{N}{2}}$. The coefficients are given by the gamma function:

$$C_k \approx \Gamma\left(\frac{k}{2} + 2\right) \sqrt{\frac{N}{2}}^k \tag{6.211}$$

where

$$\Gamma(k) = \int_0^\infty dx x^{k-1} e^{-x}. \quad (6.212)$$

Hence we have

$$\langle J_x(t) \rangle = \sum_{J=N/2}^{J_{min}} \frac{-J(2J+1)^2}{2^N \left(\frac{N}{2} + J + 1\right)} \binom{N}{\frac{N}{2} + J} (1 - e^{-t\Gamma_s J}) \quad (6.213)$$

$$\approx \langle J_x \rangle_{eq} \left[1 - \sum_{k=0}^{\infty} \frac{1}{k!} \Gamma\left(\frac{k}{2} + 2\right) \left(-t\Gamma_s \sqrt{\frac{N}{2}}\right)^k \right]. \quad (6.214)$$

Now let $\gamma = t\Gamma_s \sqrt{\frac{N}{2}}$, and define

$$\Theta(\gamma) = \sum_{k=0}^{\infty} \frac{1}{k!} \Gamma\left(\frac{k}{2} + 2\right) (-\gamma)^k \quad (6.215)$$

$$= \int_0^\infty dx \sum_{k=0}^{\infty} \frac{1}{k!} x e^{-x} \sqrt{x}^k (-\gamma)^k \quad (6.216)$$

$$= \int_0^\infty dx \sum_{k=0}^{\infty} x e^{-x} e^{-\gamma\sqrt{x}} \quad (6.217)$$

$$= \frac{1}{8} \left(2(\gamma^2 + 4) - \sqrt{\pi} e^{\frac{\gamma^2}{4}} \gamma (\gamma^2 + 6) \operatorname{erfc}\left(\frac{\gamma}{2}\right) \right) \quad (6.218)$$

where erfc is the complementary error function

$$\operatorname{erfc}(z) = 1 - \frac{2}{\pi} \int_0^z dx e^{-x^2}. \quad (6.219)$$

The complicated expression for $\Theta(\gamma)$ in Eq. (6.218) is reasonably approximated by a single exponential. We can see this in Fig. 6.9 where we compare $\Theta(\gamma)$ with $e^{-\gamma}$ and $e^{-\gamma\sqrt{\frac{\pi}{2}}}$, yielding

$$\Theta(\gamma) \approx e^{-\gamma\sqrt{\frac{\pi}{2}}}. \quad (6.220)$$

Hence we have derived an approximate expression for the cavity cooling dynamics of a maximally mixed state on the full subspace:

$$\langle J_x(t) \rangle \approx -\sqrt{\frac{N}{2\pi}} \left[1 - \exp\left(-t\Gamma_s \sqrt{\frac{\pi}{2}} \sqrt{\frac{N}{2}}\right) \right]. \quad (6.221)$$

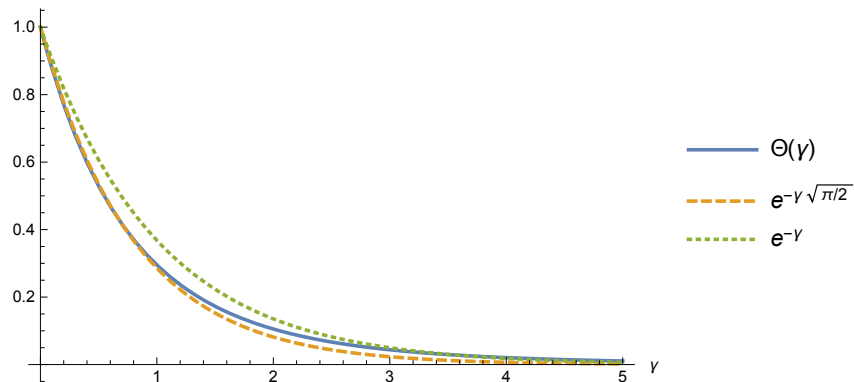


Figure 6.9: Comparison of the function $\Theta(\gamma)$ from Eq. (6.218) (solid blue line) with exponential approximations $e^{-\gamma\sqrt{\pi/2}}$ (dashed yellow line), and $e^{-\gamma}$ (dotted green line).

This shows that the effect of the cavity cooling dissipator in an initially maximally mixed state is to cool the ensemble to a final state with magnetization

$$\langle J_x \rangle \approx -\sqrt{\frac{N}{2\pi}} \sim O(\sqrt{N}) \quad (6.222)$$

with an effective cooling time constant of

$$T_{1,full} \approx \frac{2}{\Gamma_s \sqrt{\pi N}} \sim O\left(\frac{1}{\sqrt{N}}\right). \quad (6.223)$$

6.5 Cavity Cooling with Local Dephasing

The cavity cooling dissipator derived in Section 6.4 acts collectively on the spin ensemble, and hence it does not couple the irreps of $SU(2)$ in the coupled angular momentum decomposition of the state space. Thus, if the spin-ensemble is initially in a maximally mixed initial state, under ideal cavity cooling any population in a given subspace will be driven to the ground state of that subspace. The final state of the ensemble will be a mixed state consisting of the sum of all subspace ground states, as shown in Section 6.4.4. In this section we show how one may couple the subspaces by including the effect of a local dephasing dissipator on each spin in the ensemble, which in turn enables cavity cooling to the true thermal ground state of the spin ensemble.

Before we consider the general model for N spins we begin with an illustrative example of the simplest case of $N = 2$ spins. In this case dephasing can arise from a difference in the resonant frequencies of the two spins. Consider the Hamiltonian

$$H_I = \frac{\Delta\omega}{2} \frac{\sigma_z^{(1)}}{2} - \frac{\Delta\omega}{2} \frac{\sigma_z^{(2)}}{2} + g \left(\sigma_+^{(1)} a + \sigma_-^{(1)} a^\dagger \right) + g \left(\sigma_+^{(2)} a + \sigma_-^{(2)} a^\dagger \right) \quad (6.224)$$

corresponding to the TC Hamiltonian for 2 spins with the resonant frequencies of the spins given by $\omega_c + \Delta\omega/2$ and $\omega_c - \Delta\omega/2$ for spin 1 and 2 respectively, where ω_C is the resonant frequencies of the cavity. If the detuning is zero, then the two-spins may be treated as a spin-1 (triplet) and spin-0 (singlet) subspace which do not interact with each other under cavity cooling. However if $\Delta\omega > 0$ we expect coupling between the singlet and triplet, which should allow cooling to the ground state of the triplet subspace.

In [Fig. 6.10](#) we simulate the evolution under [Eq. \(6.224\)](#) in the presence of cavity dissipation at a rate κ for values of $g = 10, \kappa = 10g$, and with detuning $\Delta\omega = 0, g/10, g, 10g$. The values of g and κ were chosen so that the cavity cooling dissipation rate in the absence of detuning is $\Gamma_s = g^2/\kappa = 1$. As expected, we see that when $\Delta\omega = 0$ there is no coupling between the single state and the triplet. In this case population initially in the singlet is trapped, while population in the triplet is cooled to the triplet ground state. For $\Delta\omega > 0$ we see that the population of the singlet is cooled to the ground state of the triplet. For $\Delta\omega = g/10$ we are limited by the detuning which is shown as a strong bi-exponential in the population of the triplet ground state. The initial increase in population corresponds to cavity cooling in the triplet subspace which is quickly saturated, and then the slower second rate corresponds to the leakage from the singlet into the triplet by the detuning. Increasing the detuning to $\Delta\omega = g$ we see an increase in the transfer rate and an approximate exponential curve for the transfer of population to the triplet ground state. However, as we increase the detuning further this rate slows down again. This is because for very large detuning, both spins are very far off-resonance with the cavity, which inhibits the exchange of energy between the spins and cavity.

Treating a similar detuning mechanism for N spins is in general intractable for large N . Instead we approximate this situation by considering the phenomenological effect on each spin as local dephasing. As before, we consider an ensemble of N identical spin 1/2 particles interacting with a high-Q single mode cavity. If the spins are on-resonance with the cavity, or are driven to be on-resonance with a side-band of the cavity as was considered in [Section 6.4](#), then the spin-cavity interaction is well described by the TC Hamiltonian in [Eq. \(6.76\)](#). The cavity will also experience photon-loss dissipation at a rate inversely proportional to Q , which may be described by a Lindblad dissipator \mathcal{D}_c in [Eq. \(6.111\)](#). By adiabatically eliminating the cavity using the same technique as in [Section 6.3.1](#) we obtain

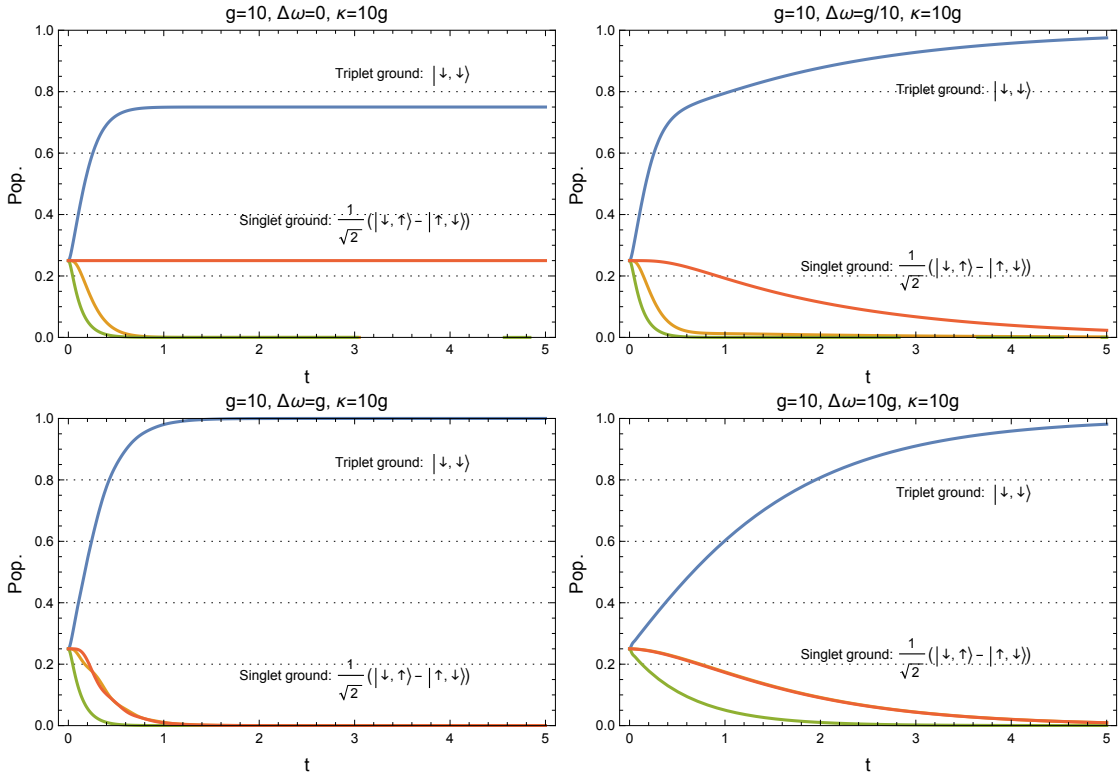


Figure 6.10: Simulation of cavity cooling under Eq. (6.224) for an ensemble of $N = 2$ spins, detuned to $\pm\Delta\omega$ from the cavity resonant frequency. We find that as the detuning increases from zero we couple the singlet and triplet subspaces allowing population transfer from the single to the triplet, and full cooling to the ground state of the triplet. However, as the detuning increases further the cooling rate is reduced, as a large detuning reduces the coupling strength between the spins and cavity thus preventing efficient transfer of energy.

the cavity cooling master equation for the spin ensemble related to Eq. (6.155), but in the Z eigenbasis

$$\frac{d}{dt}\rho_s(t) = \mathcal{D}_{cc}\rho_s(t) \quad (6.225)$$

where

$$\mathcal{D}_{cc} = \Gamma(1 + \bar{n})D[J_-] + \Gamma\bar{n}\mathcal{D}[J_+] \quad (6.226)$$

is the cavity cooling dissipator with rate $\Gamma = 4g^2/\kappa$.

To enable cooling to the ground state of the ensemble we require introducing an interaction which breaks the $SU(2)$ symmetry of the ensemble without inhibiting the cavity cooling dissipator. The model we will consider to achieve this goal is a dissipation term $\mathcal{D}_{(j)}$ which acts locally on each spin in the ensemble so that $\mathcal{D}_s = \sum_{j=1}^N \mathcal{D}_{(j)}$. The particular model of dissipation we consider is local dephasing, or T_2 dissipation, which is a phenomenological model of noise which acts on spin systems that over time causes the off diagonal elements of the density matrix to decay exponentially to zero. The intuition for this interaction is that dephasing causes a mixing of states across different spin- J subspaces that have the same J_z value. Since the cavity cooling dissipator drives each subspace to the lowest J_z value state, population trapped in the ground state of a spin- J subspace will be leaked by a T_2 process into the k^{th} excited state of a spin- $(J+k)$ subspace. This leaked population will then be cooled to the ground state of the new subspace, and in the ideal case this eventually leads to the ground true thermal state of the spin ensemble.

A T_2 process on a single spin-half system is generated by the Lindblad dissipator

$$\mathcal{D}[\sigma_z/2] = \frac{1}{4}(\sigma_z \otimes \sigma_z - \mathbb{1} \otimes \mathbb{1}) \quad (6.227)$$

$$\mathcal{D}[\sigma_z/2] : \rho \mapsto \frac{1}{4}(S_z \rho S_z^\dagger - \rho) \quad (6.228)$$

The dissipator for identical local dephasing dissipators on each spin in the ensemble is then given by

$$\mathcal{D}_{T_2} = \sum_{j=1}^N \gamma \mathcal{D}[\sigma_z^{(j)}/2]. \quad (6.229)$$

Thus, the goal is to solve the dynamics of the spin master equation

$$\frac{d}{dt}|\rho(t)\rangle\rangle = (\mathcal{D}_{cc} + \mathcal{D}_{T_2})|\rho(t)\rangle\rangle, \quad (6.230)$$

and show that this can enable cavity cooling to the ground state of the spin ensemble with magnetization $\langle J_z \rangle = -N/2$.

6.5.1 Magnus Expansion for Dissipative Interaction Frames

In order to solve the Lindblad equation Eq. (6.230) for cavity cooling in the presence of local dissipation we develop a novel new perturbation theory technique for dissipative evolution. This is an extension of Average Liouvillian Theory [LDB92, GEB99, Gho00] that involves applying the Magnus expansion in an imaginary time dissipative interaction frame to the superoperators describing the open system evolution. Consider a Lindblad master equation

$$\frac{d}{dt}|\rho(t)\rangle\rangle = (\mathcal{D}_0 + \mathcal{S}_1)|\rho(t)\rangle\rangle, \quad (6.231)$$

where the dominant term is a Lindblad dissipator

$$\mathcal{D}_0 = \sum_j \gamma_j \mathcal{D}[A_j] \quad (6.232)$$

and \mathcal{S}_1 may consist of dissipative and/or Hamiltonian evolution terms. We move into the dissipative interaction frame defined by \mathcal{D}_0 . Let

$$|\tilde{\rho}(t)\rangle\rangle \equiv e^{t\mathcal{D}_0}|\rho(t)\rangle\rangle \quad (6.233)$$

$$\tilde{\mathcal{S}}_1(t) \equiv e^{t\mathcal{D}_0}\mathcal{S}_1e^{-t\mathcal{D}_0}. \quad (6.234)$$

In this interaction frame we have that

$$\frac{d}{dt}|\tilde{\rho}(t)\rangle\rangle = \tilde{\mathcal{S}}_1|\tilde{\rho}(t)\rangle\rangle \quad (6.235)$$

Since we are in the interaction frame of a purely dissipative term, time-dependent terms in $\tilde{\mathcal{S}}_1(t)$ will be of the form $e^{\pm\omega t}$ where ω is real and positive. Thus as $t \rightarrow \infty$ the $e^{\omega t}$ terms will diverge to infinity. For the case where the original superoperator \mathcal{S}_1 is not time-dependent, we can remove these divergent terms by performing a *Wick rotation* into imaginary time [Wic54]. This is achieved by the change of variables

$$t \mapsto i\tau \quad (6.236)$$

$$dt \mapsto id\tau. \quad (6.237)$$

In the Wick rotated frame we have that

$$\tilde{\mathcal{S}}_1(i\tau) = e^{i\tau\mathcal{D}_0}\mathcal{S}_1e^{-i\tau\mathcal{D}_0} \quad (6.238)$$

which is a periodic superoperator, with period defined by \mathcal{D}_0 . If we define operators

$$G(\tau) \equiv \mathcal{S}_1(i\tau) \quad (6.239)$$

$$Q(\tau) \equiv \rho(i\tau) \quad (6.240)$$

then the Wick rotated master equation may be written as

$$\frac{d}{d\tau} |Q(\tau)\rangle\rangle = iG(\tau) |Q(\tau)\rangle\rangle \quad (6.241)$$

which has the solution

$$|Q(\tau)\rangle\rangle = \mathcal{T} \exp \left(i \int_{\tau_0}^{\tau} ds G(s) \right) |Q(0)\rangle\rangle \quad (6.242)$$

where \mathcal{T} is the time ordering operator. Since $G(\tau)$ is a periodic function with period T we can consider the average stroboscopic evolution over this period by using the Magnus expansion [Mag54]. The Magnus expansion states that

$$\mathcal{T} \exp \left(i \int_0^T ds G(s) \right) = \exp \left(\sum_{k=1}^{\infty} \Omega_k(T) \right) \quad (6.243)$$

where $\Omega_k(T)$ are the Magnus terms, with the first three given by

$$\Omega_1(T) = \int_0^T ds_1 i G(s_1) \quad (6.244)$$

$$\Omega_2(T) = -\frac{1}{2} \int_0^T ds_1 \int_0^{s_1} ds_2 [G(s_1), G(s_2)] \quad (6.245)$$

$$\Omega_3(T) = -\frac{i}{6} \int_0^T ds_1 \int_0^{s_1} ds_2 \int_0^{s_2} ds_3 \left([G(s_1), [G(s_2), G(s_3)]] + [G(s_3), [G(s_2), G(s_1)]] \right). \quad (6.246)$$

We can define an average superoperator over the period T by

$$\bar{D}_k \equiv \frac{1}{iT} \Omega_k(T), \quad (6.247)$$

and average stroboscopic evolution for a time $\tau = nT$ for integer n is given by

$$|Q(\tau)\rangle\rangle = \exp \left(i\tau \sum_{k=0}^{\infty} \bar{D}_k \right) |Q(0)\rangle\rangle. \quad (6.248)$$

Since the average dissipator is time-independent, we may inverse Wick-rotate back into real time to obtain an average description of the system dynamics over the period T , which may then be used to compute the stroboscopic evolution over integer multiples of the period T :

$$|\tilde{\rho}_s(nT)\rangle\rangle = \exp\left(nT \sum_{k=1}^{\infty} \overline{D}_k\right) |\tilde{\rho}_s(0)\rangle\rangle. \quad (6.249)$$

The stroboscopic case the evolution is only equal to the Magnus expansion when t is an integer multiple of T , however, if T is sufficiently short it is a useful course graining of the evolution. In addition, If the rate of the dissipator \mathcal{D}_0 is greater than the rate of the term \mathcal{S}_1 rate we may make a secular approximation and only need to consider the lowest order terms of the Magnus expansion. We also note that since the original interaction frame superoperator $\tilde{S}_1(t)$ had possibly divergent terms with time, so will the average dissipator \overline{D}_k , in general. Hence care must be taken to only evaluate the evolution in terms of an expectation value.

6.5.2 Average Cavity Cooling Dissipator in the T_2 Interaction Frame

We now return to finding the equilibrium state for Eq. (6.230) for cavity cooling with local dephasing. To do this we use the fact that this master equation preserves a global $SU(4)$ symmetry on the state space of vectorized density matrices. We recall that superoperators for the cavity cooling dissipator and local T_2 dissipator can be expressed in terms of $SU(4)$ generators by Eqs. (6.34) and (6.37):

$$\mathcal{D}_{cc} = \Gamma\left((1 + \bar{n})\mathcal{D}[J_-] + \bar{n}\mathcal{D}[J_+]\right) \quad (6.250a)$$

$$\begin{aligned} \mathcal{D}[J_{\pm}] = & (\mathcal{U}_{\pm} + \mathcal{V}_{\pm})(\mathcal{M}_{\pm} + \mathcal{N}_{\pm}) - \frac{1}{2}(\mathcal{U}_{\mp} + \mathcal{V}_{\mp})(\mathcal{U}_{\pm} + \mathcal{V}_{\pm}) \\ & - \frac{1}{2}(\mathcal{M}_{\mp} + \mathcal{N}_{\mp})(\mathcal{M}_{\pm} + \mathcal{N}_{\pm}) \end{aligned} \quad (6.250b)$$

$$\mathcal{D}_{T_2} = \mathcal{M}_3 - \frac{1}{2}\mathcal{Q}_3 - \frac{1}{2}\Sigma_3 - \frac{N}{4}\mathcal{I} \quad (6.250c)$$

We now consider the effective cavity cooling dissipator in the interaction frame of the T_2 dissipator using the techniques detailed in Section 6.5.1. This is given by

$$\tilde{\mathcal{D}}_{cc}(t) = e^{t\mathcal{D}_{T_2}}\mathcal{D}_{cc}e^{-t\mathcal{D}_{T_2}} \quad (6.251)$$

$$= \Gamma\left((1 + \bar{n})\tilde{\mathcal{D}}[J_-](t) + \bar{n}\tilde{\mathcal{D}}[J_+](t)\right) \quad (6.252)$$

where

$$\tilde{D}[J_{\pm}](t) = e^{t\mathcal{D}_{T_2}}\mathcal{D}[J_{\pm}]e^{-t\mathcal{D}_{T_2}}. \quad (6.253)$$

We may expand this using the BCH expansion:

$$\tilde{D}[J_{\pm}](t) = \sum_{k=0}^{\infty} \frac{t^k}{k!} \text{Ad}_{\mathcal{D}_{T_2}}^k (\mathcal{D}[J_{\pm}]) \quad (6.254)$$

where $\text{Ad}_{\mathcal{D}_{T_2}}^k (\mathcal{D}[J_{\pm}])$ are nested commutator terms with

$$\text{Ad}_{\mathcal{D}_{T_2}}^0 (\mathcal{D}[J_{\pm}]) = \mathcal{D}[J_{\pm}] \quad (6.255a)$$

$$\text{Ad}_{\mathcal{D}_{T_2}}^1 (\mathcal{D}[J_{\pm}]) = [\mathcal{D}_{T_2}, \mathcal{D}[J_{\pm}]] \quad (6.255b)$$

$$\text{Ad}_{\mathcal{D}_{T_2}}^k (\mathcal{D}[J_{\pm}]) = [\mathcal{D}_{T_2}, \mathcal{C}_{k-1}[\mathcal{D}[J_{\pm}]]] \quad (6.255c)$$

We may compute the commutator terms of the BCH expansion using the $SU(4)$ algebra. To begin, we have

$$[\mathcal{D}_{T_2}, \mathcal{M}_{\pm}] = \pm \frac{\gamma}{2} \mathcal{M}_{\pm}, \quad [\mathcal{D}_{T_2}, \mathcal{N}_{\pm}] = \mp \frac{\gamma}{2} \mathcal{N}_{\pm}, \quad (6.256)$$

$$[\mathcal{D}_{T_2}, \mathcal{U}_{\pm}] = \pm \frac{\gamma}{2} \mathcal{U}_{\pm}, \quad [\mathcal{D}_{T_2}, \mathcal{V}_{\pm}] = \mp \frac{\gamma}{2} \mathcal{V}_{\pm}. \quad (6.257)$$

Hence,

$$\begin{aligned} [\mathcal{D}_{T_2}, \mathcal{D}[J_{\pm}]] &= \frac{1}{2} \left([\mathcal{D}_{T_2}, (\mathcal{U}_{\pm} + \mathcal{V}_{\pm})(\mathcal{M}_{\pm} + \mathcal{N}_{\pm}) + (\mathcal{M}_{\pm} + \mathcal{N}_{\pm})(\mathcal{U}_{\pm} + \mathcal{V}_{\pm})] \right. \\ &\quad \left. - [\mathcal{D}_{T_2}, (\mathcal{U}_{\mp} + \mathcal{V}_{\mp})(\mathcal{U}_{\pm} + \mathcal{V}_{\pm}) + (\mathcal{M}_{\mp} + \mathcal{N}_{\mp})(\mathcal{M}_{\pm} + \mathcal{N}_{\pm})] \right) \quad (6.258a) \end{aligned}$$

$$\begin{aligned} &= \frac{1}{2} \left([\mathcal{D}_{T_2}, \mathcal{U}_{\pm} + \mathcal{V}_{\pm}](\mathcal{M}_{\pm} + \mathcal{N}_{\pm}) + (\mathcal{M}_{\pm} + \mathcal{N}_{\pm})[\mathcal{D}_{T_2}, \mathcal{U}_{\pm} + \mathcal{V}_{\pm}] \right. \\ &\quad + [\mathcal{D}_{T_2}, \mathcal{M}_{\pm} + \mathcal{N}_{\pm}](\mathcal{U}_{\pm} + \mathcal{V}_{\pm}) + (\mathcal{U}_{\pm} + \mathcal{V}_{\pm})[\mathcal{D}_{T_2}, \mathcal{M}_{\pm} + \mathcal{N}_{\pm}] \\ &\quad - [\mathcal{D}_{T_2}, \mathcal{U}_{\mp} + \mathcal{V}_{\mp}](\mathcal{U}_{\pm} + \mathcal{V}_{\pm}) - (\mathcal{U}_{\mp} + \mathcal{V}_{\mp})[\mathcal{D}_{T_2}, \mathcal{U}_{\pm} + \mathcal{V}_{\pm}] \\ &\quad \left. - [\mathcal{D}_{T_2}, \mathcal{M}_{\mp} + \mathcal{N}_{\mp}](\mathcal{M}_{\pm} + \mathcal{N}_{\pm}) - (\mathcal{M}_{\mp} + \mathcal{N}_{\mp})[\mathcal{D}_{T_2}, \mathcal{M}_{\pm} + \mathcal{N}_{\pm}] \right) \quad (6.258b) \end{aligned}$$

$$\begin{aligned} &= \pm \frac{\gamma}{2} \left[(\mathcal{M}_{\pm} \mathcal{U}_{\pm} + \mathcal{U}_{\pm} \mathcal{M}_{\pm} - \mathcal{V}_{\mp} \mathcal{U}_{\pm} - \mathcal{N}_{\mp} \mathcal{M}_{\pm}) \right. \\ &\quad \left. - (\mathcal{N}_{\pm} \mathcal{V}_{\pm} + \mathcal{V}_{\pm} \mathcal{N}_{\pm} - \mathcal{U}_{\mp} \mathcal{V}_{\pm} - \mathcal{M}_{\mp} \mathcal{N}_{\pm}) \right]. \quad (6.258c) \end{aligned}$$

Next, we define superoperators

$$\mathcal{A}_\pm = \frac{1}{2}(\mathcal{M}_\pm \mathcal{U}_\pm + \mathcal{U}_\pm \mathcal{M}_\pm - \mathcal{V}_\mp \mathcal{U}_\pm - \mathcal{N}_\mp \mathcal{M}_\pm) \quad (6.259)$$

$$\mathcal{B}_\pm = \frac{1}{2}(\mathcal{N}_\pm \mathcal{V}_\pm + \mathcal{V}_\pm \mathcal{N}_\pm - \mathcal{U}_\mp \mathcal{V}_\pm - \mathcal{M}_\mp \mathcal{N}_\pm). \quad (6.260)$$

We may then write

$$[\mathcal{D}_{T_2}, \mathcal{D}[J_\pm]] = \pm\gamma\mathcal{A}_\pm \mp \gamma\mathcal{B}_\pm. \quad (6.261)$$

If we take the commutator of \mathcal{A}_\pm and \mathcal{B}_\pm with the T_2 dissipator we find

$$[\mathcal{D}_{T_2}, \mathcal{A}_\pm] = \pm\gamma\mathcal{A}_\pm \quad [\mathcal{D}_{T_2}, \mathcal{B}_\pm] = \mp\gamma\mathcal{B}_\pm. \quad (6.262)$$

Hence for $k \geq 1$ we have that the nested commutator terms are given by

$$\text{Ad}_{\mathcal{D}_{T_2}}^k (D[J_\pm]) = (\pm\gamma)^k \mathcal{A}_\pm + (\mp\gamma)^k \mathcal{B}_\pm \quad (6.263)$$

and

$$\tilde{D}[J_\pm](t) = \sum_{k=0}^{\infty} \frac{t^k}{k!} \text{Ad}_{\mathcal{D}_{T_2}}^k (D[J_\pm]) \quad (6.264)$$

$$= \mathcal{D}[J_\pm] - \mathcal{A}_\pm - \mathcal{B}_\pm + \sum_{k=0}^{\infty} \frac{(\pm\gamma t)^k}{k!} \mathcal{A}_\pm + \sum_{k=0}^{\infty} \frac{(\mp\gamma t)^k}{k!} \mathcal{B}_\pm \quad (6.265)$$

$$= \mathcal{D}[J_\pm] + (e^{\pm\gamma t} - 1)\mathcal{A}_\pm + (e^{\mp\gamma t} - 1)\mathcal{B}_\pm. \quad (6.266)$$

In the Wick-rotated interaction frame $G(\tau) = \tilde{D}[J_\pm](i\tau)$ is periodic with period $T = 2\pi/\gamma$. Hence, if the dephasing rate is greater than the collective cavity cooling rate we may make a secular approximation for the average dissipator, and the first order Magnus term over this period is the time-independent piece of Eq. (6.266):

$$\overline{\mathcal{D}}_1 = \Gamma(1 + \bar{n})\overline{\mathcal{G}}_- + \Gamma\bar{n}\overline{\mathcal{G}}_+, \quad (6.267)$$

where

$$\begin{aligned} \overline{\mathcal{G}}_\pm &= \mathcal{D}[J_\pm] - \mathcal{A}_\pm - \mathcal{B}_\pm \\ &= \frac{1}{2}(\mathcal{U}_\pm \mathcal{N}_\pm + \mathcal{N}_\pm \mathcal{U}_\pm + \mathcal{M}_\pm \mathcal{V}_\pm + \mathcal{V}_\pm \mathcal{M}_\pm - \mathcal{U}_\mp \mathcal{U}_\pm - \mathcal{V}_\mp \mathcal{V}_\pm - \mathcal{M}_\mp \mathcal{M}_\pm - \mathcal{N}_\mp \mathcal{N}_\pm). \end{aligned} \quad (6.268)$$

6.5.3 Equilibrium Value and Cooling Rate for $\langle J_z(t) \rangle$

Using the first order average dissipator in Eq. (6.267) we are interested in computing the evolution of the expectation value of the J_z operator for an arbitrary initial state. We do this by solving

$$\begin{aligned}
\langle J_z(t) \rangle &= \langle\langle J_z | \rho(t) \rangle\rangle \\
&= \langle\langle \tilde{J}_z(t) | \tilde{\rho}(t) \rangle\rangle \\
&= \langle\langle J_z | e^{t\bar{\mathcal{D}}_1} | \rho(0) \rangle\rangle \\
&= \langle\langle \rho(0) | e^{t\bar{\mathcal{D}}_1^\dagger} | J_z \rangle\rangle,
\end{aligned} \tag{6.269}$$

where $\tilde{J}_z(t) = J_z(t)$ as $[J_z, \mathcal{D}_{T_2}] = 0$.

The solution to Eq. (6.269) is given by

$$\langle J_z(t) \rangle = e^{-t\Gamma_{cc}(1+2\bar{n})} \langle J_z(0) \rangle - \frac{N(1 - e^{-t\Gamma_{cc}(1+2\bar{n})})}{2 + 4\bar{n}}. \tag{6.270}$$

This is an exponential relaxation process with decay rate $T_1 = \Gamma_{cc}(1+2\bar{n})$ to an equilibrium state with magnetization

$$\langle J_z \rangle_{eq} = -\frac{N}{2 + 4\bar{n}}. \tag{6.271}$$

Thus in the ideal cooling limit with $\bar{n} = 0$ this is a T_1 process to the ground state of the spin ensemble with magnetization $\langle J_z \rangle_{eq} = -N/2$.

We now proceed with the derivation of Eq. (6.270). We first note that the the adjoint dissipator is given by

$$\bar{\mathcal{D}}_1^\dagger = \bar{\mathcal{G}}_1 = \Gamma(1 + \bar{n})\bar{\mathcal{G}}_-^\dagger + \Gamma\bar{n}\bar{\mathcal{G}}_+^\dagger, \tag{6.272}$$

where

$$\bar{\mathcal{G}}_\pm^\dagger = \frac{1}{2} \left(\mathcal{U}_\mp \mathcal{N}_\mp + \mathcal{N}_\mp \mathcal{U}_\mp + \mathcal{M}_\mp \mathcal{V}_\mp + \mathcal{V}_\mp \mathcal{M}_\mp - \mathcal{U}_\mp \mathcal{U}_\pm - \mathcal{V}_\mp \mathcal{V}_\pm - \mathcal{M}_\mp \mathcal{M}_\pm - \mathcal{N}_\mp \mathcal{N}_\pm \right) \tag{6.273}$$

and we have used $\mathcal{M}_\pm^\dagger = \mathcal{M}_\mp$. Similar arguments follow for $\mathcal{N}_\pm, \mathcal{U}_\pm, \mathcal{V}_\pm$.

Next, we compute the terms of $\bar{\mathcal{D}}_1^\dagger | J_z \rangle\rangle$. We have that

$$(\mathcal{U}_\mp \mathcal{N}_\mp + \mathcal{N}_\mp \mathcal{U}_\mp) | J_z \rangle\rangle = \sum_{i,j=1}^N \left(\left| E_+^{(j)} S_\mp^{(i)} J_z E_-^{(i)} S_\pm^{(j)} \right\rangle\rangle + \left| \sigma_\mp^{(i)} E_+^{(j)} J_z S_\pm^{(j)} E_-^{(i)} \right\rangle\rangle \right) \tag{6.274a}$$

$$= \sum_{i,j=1}^N \left(\left| E_+^{(j)} S_\mp^{(i)} E_-^{(i)} J_z S_\pm^{(j)} \right\rangle\rangle + \left| \sigma_\mp^{(i)} J_z E_+^{(j)} S_\pm^{(j)} E_-^{(i)} \right\rangle\rangle \right). \tag{6.274b}$$

Hence,

$$(\mathcal{U}_-\mathcal{N}_- + \mathcal{N}_-\mathcal{U}_-)|J_z\rangle\rangle = \sum_{i,j=1}^N \left| S_-^{(i)} J_z S_+^{(j)} E_-^{(i)} \right\rangle\rangle \quad (6.275a)$$

$$= \sum_{i,j=1}^N \left(\left| [S_-^{(i)}, J_z] S_+^{(j)} E_-^{(i)} \right\rangle\rangle + \left| J_z S_-^{(i)} S_+^{(j)} E_-^{(i)} \right\rangle\rangle \right) \quad (6.275b)$$

$$= \sum_{j=1}^N \left| (J_z + \mathbb{1}) E_-^{(j)} \right\rangle\rangle \quad (6.275c)$$

$$= \left| (J_z + \mathbb{1}) \left(\frac{N}{2} \mathbb{1} - J_z \right) \right\rangle\rangle \quad (6.275d)$$

and

$$(\mathcal{U}_+\mathcal{N}_+ + \mathcal{N}_+\mathcal{U}_+)|J_z\rangle\rangle = \sum_{i,j=1}^N \left| E_+^{(j)} S_+^{(i)} J_z S_-^{(j)} \right\rangle\rangle \quad (6.276a)$$

$$= \sum_{i,j=1}^N \left| E_+^{(j)} S_+^{(i)} S_-^{(j)} (J_z - \mathbb{1}) \right\rangle\rangle \quad (6.276b)$$

$$= \left| \left(\frac{N}{2} \mathbb{1} + J_z \right) (J_z - \mathbb{1}) \right\rangle\rangle, \quad (6.276c)$$

where we have made use of the relations

$$E_\pm E_\pm = E_\pm, E_\pm E_\mp = 0 \quad (6.277a)$$

$$E_\pm S_\pm = S_\pm E_\mp = S_\pm \quad (6.277b)$$

$$E_\mp S_\pm = S_\pm E_\pm = 0. \quad (6.277c)$$

Thus, we arrive at the expression

$$(\mathcal{U}_\mp \mathcal{N}_\mp + \mathcal{N}_\mp \mathcal{U}_\mp)|J_z\rangle\rangle = \left| \left(\frac{N}{2} \mathbb{1} \mp J_z \right) (J_z \pm \mathbb{1}) \right\rangle\rangle. \quad (6.278)$$

Similarly one can show

$$(\mathcal{V}_\mp \mathcal{M}_\mp + \mathcal{M}_\mp \mathcal{V}_\mp)|J_z\rangle\rangle = \left| \left(\frac{N}{2} \mathbb{1} \mp J_z \right) (J_z \pm \mathbb{1}) \right\rangle\rangle. \quad (6.279)$$

For the other terms we have

$$\mathcal{U}_\mp \mathcal{U}_\pm |J_z\rangle\rangle = \sum_{i,j=1}^N \left| E_+^{(j)} E_+^{(i)} J_z S_\mp^{(i)} S_\pm^{(j)} \right\rangle\rangle = \sum_{i,j=1}^N \left| J_z E_+^{(j)} E_+^{(i)} S_\mp^{(i)} S_\pm^{(j)} \right\rangle\rangle. \quad (6.280)$$

Hence,

$$\mathcal{U}_- \mathcal{U}_+ |J_z\rangle\rangle = \sum_{i,j=1}^N \left| J_z E_+^{(j)} E_+^{(i)} S_-^{(i)} S_+^{(j)} \right\rangle\rangle = 0 \quad (6.281)$$

$$\mathcal{U}_+ \mathcal{U}_- |J_z\rangle\rangle = \sum_{i,j=1}^N \left| J_z E_+^{(j)} S_+^{(i)} S_-^{(j)} \right\rangle\rangle = \sum_{j=1}^N \left| J_z E_+^{(j)} \right\rangle\rangle = \left| J_z \left(\frac{N}{2} \mathbb{1} + J_z \right) \right\rangle\rangle. \quad (6.282)$$

Similarly,

$$\mathcal{M}_- \mathcal{M}_+ |J_z\rangle\rangle = 0 \quad (6.283a)$$

$$\mathcal{M}_+ \mathcal{M}_- |J_z\rangle\rangle = \left| J_z \left(\frac{N}{2} \mathbb{1} + J_z \right) \right\rangle\rangle \quad (6.283b)$$

$$\mathcal{V}_- \mathcal{V}_+ |J_z\rangle\rangle = \left| J_z \left(\frac{N}{2} \mathbb{1} - J_z \right) \right\rangle\rangle \quad (6.283c)$$

$$\mathcal{V}_+ \mathcal{V}_- |J_z\rangle\rangle = 0 \quad (6.283d)$$

$$\mathcal{N}_- \mathcal{N}_+ |J_z\rangle\rangle = \left| J_z \left(\frac{N}{2} \mathbb{1} - J_z \right) \right\rangle\rangle \quad (6.283e)$$

$$\mathcal{N}_+ \mathcal{N}_- |J_z\rangle\rangle = 0. \quad (6.283f)$$

Thus, we have

$$\begin{aligned} \bar{G}_\pm^\dagger |J_z\rangle\rangle &= \left| \left(\frac{N}{2} \mathbb{1} \mp J_z \right) J_z \right\rangle\rangle \pm \left| \left(\frac{N}{2} \mathbb{1} \mp J_z \right) \right\rangle\rangle - \left| \left(\frac{N}{2} \mathbb{1} \mp J_z \right) J_z \right\rangle\rangle \\ &= -|J_z\rangle\rangle \pm \frac{N}{2} |\mathbb{1}\rangle\rangle. \end{aligned} \quad (6.284)$$

Next, we need to evaluate $\bar{G}_\pm^\dagger |\mathbb{1}\rangle\rangle$. We have that

$$\begin{aligned} \mathcal{M}_+ |\mathbb{1}\rangle\rangle &= 0 & \mathcal{M}_- |\mathbb{1}\rangle\rangle &= |J_-\rangle\rangle \\ \mathcal{U}_+ |\mathbb{1}\rangle\rangle &= 0 & \mathcal{U}_- |\mathbb{1}\rangle\rangle &= |J_+\rangle\rangle \\ \mathcal{N}_+ |\mathbb{1}\rangle\rangle &= |J_+\rangle\rangle & \mathcal{N}_- |\mathbb{1}\rangle\rangle &= 0 \\ \mathcal{V}_+ |\mathbb{1}\rangle\rangle &= |J_-\rangle\rangle & \mathcal{V}_- |\mathbb{1}\rangle\rangle &= 0 \end{aligned}$$

and so

$$\bar{G}_+^\dagger |\mathbb{1}\rangle\rangle = \frac{1}{2} \left(\mathcal{N}_- |J_+\rangle\rangle + \mathcal{V}_- |J_-\rangle\rangle \right) - \frac{1}{2} \left(\mathcal{N}_- |J_+\rangle\rangle + \mathcal{V}_- |J_-\rangle\rangle \right) = 0 \quad (6.285)$$

$$\bar{G}_-^\dagger |\mathbb{1}\rangle\rangle = \frac{1}{2} \left(\mathcal{M}_+ |J_-\rangle\rangle + \mathcal{U}_+ |J_+\rangle\rangle \right) - \frac{1}{2} \left(\mathcal{M}_+ |J_-\rangle\rangle + \mathcal{U}_+ |J_+\rangle\rangle \right) = 0. \quad (6.286)$$

Finally, we may put this all together to obtain

$$\begin{aligned}
\bar{G}^\dagger |J_z\rangle\rangle &= \Gamma \left((1 + \bar{n})\bar{G}_-^\dagger + \bar{n}\bar{G}_+^\dagger \right) |J_z\rangle\rangle \\
&= \Gamma(1 + \bar{n})(-|J_z\rangle\rangle - \frac{N}{2}|\mathbb{1}\rangle\rangle) + \Gamma\bar{n}(-|J_z\rangle\rangle + \frac{N}{2}|\mathbb{1}\rangle\rangle) \\
&= -\Gamma(1 + 2\bar{n}) \left[|J_z\rangle\rangle + \left(\frac{N}{2 + 4\bar{n}} \right) |\mathbb{1}\rangle\rangle \right].
\end{aligned} \tag{6.287}$$

In general, for $k \geq 1$

$$\bar{G}^k |J_z\rangle\rangle = (-1)^k \Gamma^k (1 + 2\bar{n})^k \left[|J_z\rangle\rangle + \left(\frac{N}{2 + 4\bar{n}} \right) |\mathbb{1}\rangle\rangle \right]. \tag{6.288}$$

Hence,

$$\begin{aligned}
e^{t\bar{G}} |J_z\rangle\rangle &= \sum_{k=0}^{\infty} \frac{t^k}{k!} \bar{G}^k |J_z\rangle\rangle \\
&= |J_z\rangle\rangle + \sum_{k=1}^{\infty} \frac{t^k}{k!} (-1)^k \Gamma^k (1 + 2\bar{n})^k \left[|J_z\rangle\rangle + \left(\frac{N}{2 + 4\bar{n}} \right) |\mathbb{1}\rangle\rangle \right] \\
&= e^{-t\Gamma(1+2\bar{n})} |J_z\rangle\rangle - (1 - e^{-t\Gamma(1+2\bar{n})}) \left(\frac{N}{2 + 4\bar{n}} \right) |\mathbb{1}\rangle\rangle.
\end{aligned} \tag{6.289}$$

For an arbitrary initial state $\rho(0)$, the expectation value of J_z under this evolution is given by

$$\langle J_z(t) \rangle = e^{-t\Gamma(1+2\bar{n})} \langle J_z(0) \rangle - (1 - e^{-t\Gamma(1+2\bar{n})}) \left(\frac{N}{2 + 4\bar{n}} \right). \tag{6.290}$$

The effective dynamics are thus described by an exponential decay process with decay rate $T_1 = \Gamma(1 + 2\bar{n})$ to an equilibrium state with magnetization

$$\langle J_z \rangle_{eq} = -\frac{N}{2 + 4\bar{n}}, \tag{6.291}$$

where in the ideal cooling limit of $\bar{n} = 0$, this is a T_1 process to the ground state of the spin ensemble.

6.5.4 Simulation Results

We now consider numerical simulations to compare the average cavity cooling dissipator derived in [Section 6.5.2](#) with evolution under the spin master equation [Eq. \(6.230\)](#) for

cavity cooling with local dephasing:

$$\frac{d}{dt}|\rho(t)\rangle\rangle = (\mathcal{D}_{cc} + \mathcal{D}_{T_2})|\rho(t)\rangle\rangle. \quad (6.292)$$

We also compare with evolution under the full spin-cavity evolution for the Tavis-Cummings master equation with cavity dissipation and local spin dephasing:

$$\frac{d}{dt}|\rho_{sc}(t)\rangle\rangle = (\mathcal{L}_{TC} + \mathcal{D}_c + \mathcal{D}_{T_2})|\rho_{sc}(t)\rangle\rangle. \quad (6.293)$$

All simulations were done using the *QuantumUtils for Mathematica* package [WHG15] with an N -spin matrix representation of the symmetric subspace projection of the $SU(4)$ algebra, as described in Section 6.2.3.

Fig. 6.11 shows the results for simulation of $\langle J_z(t) \rangle$ for cavity cooling with local dephasing using Eq. (6.292) for both $N = 10$ spins, and $N = 100$ spins with dephasing rates of $\gamma = \lambda N \Gamma$ with $\lambda = 0, 0.1, 1, 10$. In both cases we see that for $\gamma = 0$ we have cavity cooling alone and population is trapped, and as $\gamma > 0$ we break the $SU(2)$ symmetry and achieve cooling to the ground state, at a rate that increases with γ . For $\gamma = 10N\Gamma$ we have good agreement between the first order average dissipator expansion $\overline{\mathcal{D}}_1$ in Eq. (6.267), shown as the dotted black line in Fig. 6.11. This is expected, as Eq. (6.269) is only a valid approximation of the true dynamics when the dephasing strength γ is sufficiently strong to disregard higher order terms in the Magnus expansion where terms that do not commute with the dissipator rapidly average to zero. In practice this corresponds to requiring $\gamma > \Gamma N$. The parameter $C = \Gamma N / \gamma$ is also called the *cooperativity* of the spin ensemble, and hence the condition for the validity of lowest order Magnus approximation is that $C < 1$.

In Fig. 6.12 we simulated for $N = 10$ spins and a cavity truncated to 4 levels with values of the spin-cavity coupling of $g = 100$, cavity dissipation rate $\kappa = 4g^2 = 4 \times 10^4$, and spin dephasing rate $\gamma = \lambda N$ with $\lambda = 0, 0.1, 1, 10, 10^2, 10^3, 10^4, 10^5$. The values of g and κ were chosen to satisfy the Markovian condition $\kappa \gg g\sqrt{N}$ for $N = 100$, while giving an effective spin cavity cooling rate of $\Gamma = 4g^2/\kappa = 1$. In addition the strong cavity dissipation rate allows us to truncate the cavity to low dimension. We find that the spin-cavity master equation in Eq. (6.293) is in agreement with the spin cavity cooling master equation in Eq. (6.292) for dephasing rates up to $10N\Gamma$, however as the dephasing rate increases beyond the collective cavity dissipation rate, the cooling rate begins to slow down. In this strong dephasing regime the master equation in Eq. (6.293) is not an honest description of the dynamics. This is because when deriving the cavity cooling dissipator we must take into account that strong dephasing will reduce the effective coupling strength

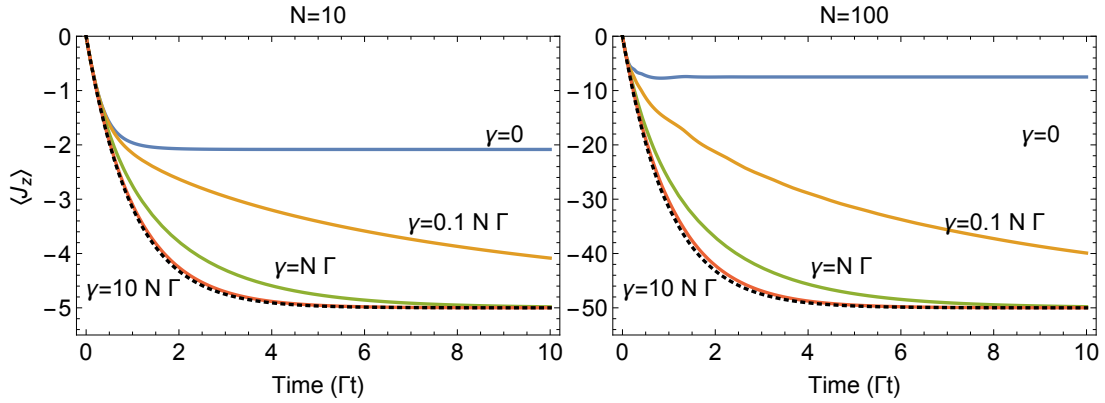


Figure 6.11: Simulations of the expectation value $\langle J_z(t) \rangle$ for a maximally mixed initial state of an ensemble of $N = 10$ spins (left), and $N = 100$ spins (right) for master equation described by first order average dissipator in Eq. (6.267) (dotted black line), and of the cavity cooling with local dephasing master equation in Eq. (6.292) for $\gamma = \lambda N \Gamma$, and $\lambda = 0, 0.1, 1, 10$. In both cases we see good agreement between the first order approximation and the full master equation for dephasing rate $\gamma = 10 N \Gamma$.

between the spin and the cavity, and hence reduce the effective dissipation rate Γ . Following the derivation in Section 6.3.1, this can be incorporated by a Lorentzian cavity cooling rate $\Gamma = 4g^2\kappa/(\kappa^2 + 4\Delta^2)$, where Δ is a parameter that depends the the physical mechanism that gives rise to the local dephasing parameter γ . For example, in the simplest case of $N = 2$, this is the expression where the dephasing mechanism arises due to the the two spins being tuned to $\pm\Delta/2$ off-resonance considered in Eq. (6.224).

6.6 Summary

In this chapter we have described how, in theory, cavity cooling techniques can be used to drive the Dicke subspace, and other coupled angular momentum subspaces of an ensemble system to their respective ground states by coupling to a high-Q cavity. We also demonstrated that these subspaces may be coupled to enable cavity cooling of all subspaces to the ensemble ground state by including a local dephasing dissipator on each spin in the ensemble.

The subspace cavity cooling we derived in Section 6.4 could prove useful as a fast reset operation of the collective subspaces of a spin-ensemble at a rate potentially much faster

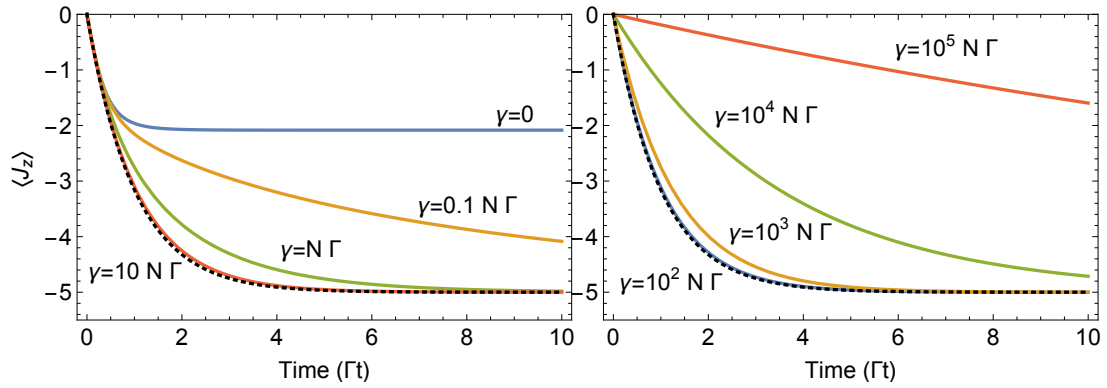


Figure 6.12: Simulations of the expectation value $\langle J_z(t) \rangle$ for the full spin-cavity master equation in Eq. (6.293) with the addition of a local dephasing dissipator for $N = 10$ spins in a maximally mixed initial state, and a cavity truncated to 4 levels initialized in the ground state. Evolution under the 1st order average dissipator in Eq. (6.267) is shown as the dotted black line in both figures. In the left figure we see good agreement with the cavity cooling master equation simulation in Fig. 6.11, however for stronger dephasing rates (right) we see that the cooling rate of the spin ensemble decreases.

than thermal relaxation. These subspaces provide a useful tool for implementing quantum error correction and other quantum information protocols in proposals for hybrid quantum systems where a Dicke subspace of a spin-ensemble is used as quantum memory for storing an excitation from a superconducting qubit or resonator [KOB+10, AKN+11, DPF+11, KGD+11]. However, since cavity cooling alone does not couple the Dicke subspace to the non-Dicke subspaces, it is less useful if the initial state is highly mixed. In Section 6.4.4 we showed that for a maximally mixed initial state, the cooling dynamics in the full Hilbert space could be approximated by an exponential relaxation process to a state with final magnetization of order \sqrt{N} , with a cooling time constant proportional to $1/\sqrt{N}$. Cavity cooling with local dephasing as discussed in Section 6.5 may be useful in this context, and this technique could prove useful for initializing an ensemble spin system in a highly pure state without requiring extremely high fields or low temperatures. In addition the method for dissipative perturbation theory that we developed to solve the cooling master equation in Section 6.5.1 could be useful for other systems where a dissipative term is dominant.

Dephasing is always present in a real physical system and this proposal uses it as a resource for dissipative state engineering. Depending on the physical system there are many possible mechanisms that give rise to dephasing. For spins systems it may arise due to inhomogeneous static fields across the ensemble, which could be engineered by introducing gradient fields. For systems such as phosphorus defects in silicon and nitrogen vacancy centres in diamond it could arise due to hyperfine coupling between each electron and a local nuclear spin or spin bath, and the strength of these interactions can be engineered to a degree using decoupling sequences. Finally it could also be engineered by stochastically applying σ_z gates to randomly chosen individual spins in the ensemble if one has the ability to implement local control operations.

For simplicity of presentation our calculation in Section 6.5 considered the case where the spin-ensemble was on resonance with the cavity and the resulting cavity cooling dissipator can be thought of in terms of the Purcell effect. One could also consider the side-band cooling approach as detailed in Section 6.4, by introducing a drive term on the spins to target a side-band of the resonator. In that case the magnetization of the spins under cooling accumulates in the J_x basis (for a J_x drive term) rather than the J_z basis of the static field. In this situation the dephasing must also happen in the J_x basis to achieve cooling to the J_x ground state. In practice this could be engineered by using a gated protocol where a single cooling step consists of: side-band cavity cooling for a time t_{cc} , applying a collective rotation swapping the J_x and J_z eigenstates, dephasing for time t_{T_2} in the J_z basis, then applying the inverse collective rotation to rotate back to the J_x basis. This cooling step can then be repeated to form a discretized cavity cooling cycle with dephasing to reach the true ground state.

Chapter 7

Quantum Correlations in Neutron Interferometry

7.1 Introduction

A unique property of quantum theory is that when two or more quantum systems are allowed to interact they may exhibit correlations that cannot be explained classically. In the field of quantum information science protocols harnessing these correlations can exceed classical efficiencies for certain metrology applications and information processing tasks [NC00]. One of the most studied classes of correlated quantum states are entangled states as they enable extremely non-classical quantum effects such as quantum teleportation [HHHH09]. A maximally-entangled quantum state of a bipartite quantum system allows for a projective measurement of one subsystem to completely determine the outcome of the corresponding projective measurements on the other. The class of states of interest to quantum computation however is broader than purely entangled quantum states, as certain non-entangled quantum states may still possess correlations that cannot be accounted for classically. In such cases measurement on one subsystem, while not determining the state of another, may still cause a disturbance to the state of the other.

Classifying the quantum nature of correlations beyond entanglement has received much interest, with many discussions focused on quantum discord (QD) and related measures [LCS11, CMS11, MBC⁺12]. Quantum discord was proposed by Ollivier and Zurek [OZ02], and Henderson and Vedral [HV01] to characterize quantum correlations in a bipartite system. In effect, one may interpret QD as a measure of the minimum disturbance that measurement of one subsystem of a bipartite quantum system can induce on the measurement

outcomes of the other. Such classifications are of interest since certain quantum algorithms, such as DQC1, do not require entanglement to exceed classical efficiencies[KL98]. It has been shown that for the DQC1 algorithm QD is present in the output state of the computation even when entanglement is not, and hence it was suggested that QD may provide a better figure of merit of evaluating quantum resources [DSC08]. Here we investigate the quantum nature of correlations of single neutrons in a neutron interferometer (NI).

Neutron interferometry has been used for precise tests of quantum mechanical phenomena such as coherent spinor rotation [RZB⁺75] and superposition [SBR⁺83], gravitationally induced quantum interference [COW75], the Aharonov-Casher effect [COK⁺89], violation of a Bell-like inequality[HLB⁺03], generation of a single neutron entangled state[HLB⁺07], quantum contextually [BKS⁺09], and the realization of a Decoherence-Free subspace [PAC09, PHAC11]. In our case a NI provides a clean system for considering quantum correlations in a bipartite quantum systems as we are able to coherently control the spin and path-momentum degrees of freedom of a neutron beam, and manipulate the correlations between them. In addition, due to the high efficiency of single neutron detectors and the low intensity of neutrons entering the interferometer, we are able to gather statistics from performing true projective measurements on single quantum systems. In the present chapter we investigate the correlations between the spin and path degrees of freedom of the output beam from a noisy NI by observing changes in the output beam intensity as a result of a post-selected projective measurement on the neutron spin.

7.2 Theoretical Model of a Neutron Interferometer

The most common geometry for a neutron interferometer is a three-blade configuration machined from a perfect single crystal of silicon. This type of interferometer functions as a Mach–Zender interferometer on the longitudinal momentum of the neutron beam. We refer to this degree of freedom of the neutron beam as the *path* system. The neutron path can be viewed as a two level quantum system. The neutrons in the interferometer are also spin-half particles, and we can treat the spin of the neutron as a second two level quantum system that we call the *spin* system. We may couple to the neutron path and spin to form a bipartite quantum system. In this context we may view the interferometer crystal as a quantum circuit acting as illustrated in Fig. 7.1. We define the basis for the path to be the computational basis where $|0\rangle$ and $|1\rangle$ correspond to the red and blue beam paths in Fig. 7.1 respectively. For the spin-system we work in the spin-up, spin-down eigenbasis $|\uparrow\rangle, |\downarrow\rangle$ with respect to a static field in the z -direction.

The first (and third) NI blades of the interferometer act as Hadamard (H) gates on the

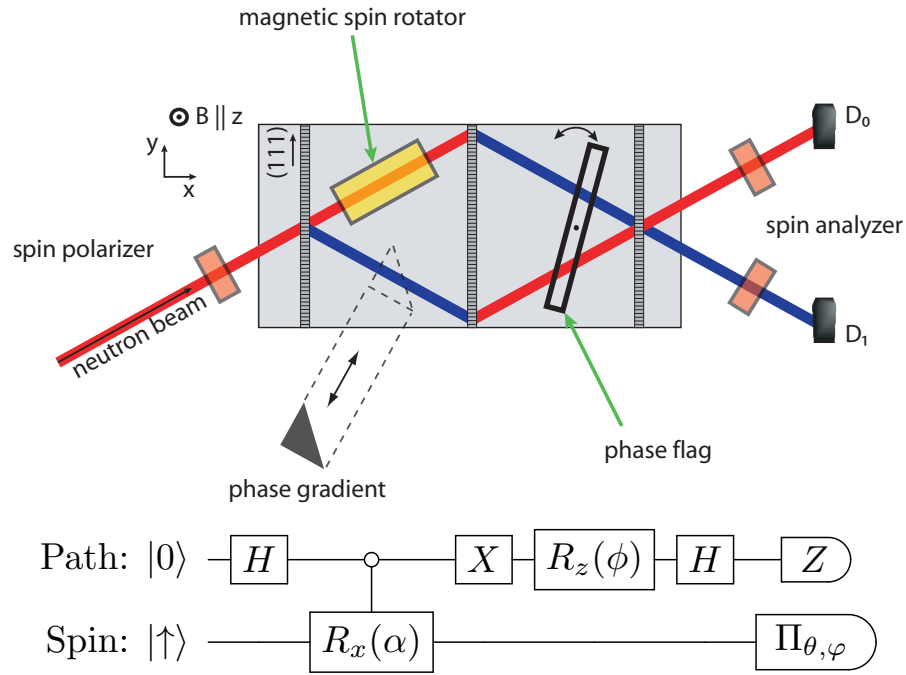


Figure 7.1: Experimental setup for the three blade neutron interferometer (top) and the corresponding quantum circuit for the ideal model (bottom). The red (blue) paths in NI schematic are defined as the $|0\rangle(|1\rangle)$ path states, H is a Hadamard gate, $R_x(\alpha)$ is a rotation of the neutron spin in the $|0\rangle$ path of α radians about x -axis, X is a bit-flip, $R_z(\phi)$ is a relative phase shift of ϕ radians between the beam paths, Π is a projective measurement performed on the spin-state (spin-analyser) in the basis $\cos(\theta)|\uparrow\rangle \pm e^{i\phi} \sin(\theta)|\downarrow\rangle$, and Z is a projective measurement of the path intensities in the $|0\rangle, |1\rangle$ basis.

neutron path by coherently splitting (and recombining) the neutron beam into two paths via Bragg scattering in the Laue geometry [Sea89]. The second blade deflects the beam by swapping the path-momentum directions, which we model as a bit-flip (X) gate. In our defined bases these are given by:

$$H = \frac{1}{\sqrt{2}} (|0\rangle\langle 0| + |0\rangle\langle 1| + |1\rangle\langle 0| - |1\rangle\langle 1|) \quad (7.1)$$

$$X = |0\rangle\langle 1| + |1\rangle\langle 0| \quad (7.2)$$

In practice the intensity of the output neutron beam is reduced compared to the input beam due to neutrons escaping the NI at the second blade. We account for this in our description of the output beam however by post-selecting on the neutrons which remain in the interferometer.

Between the first and second NI blades we couple the spin and path degrees of freedom by selectively rotating the neutron spin in the $|0\rangle$ path by an angle α . This acts a controlled- X rotation ($R_x(\alpha)$), with the spin and path as the target and control respectively:

$$C-R_x(\alpha) = |0\rangle\langle 0| \otimes R_x(\alpha) + |1\rangle\langle 1| \otimes \mathbb{1}_s \quad (7.3)$$

$$R_x(\alpha) = \exp \left[i \frac{\alpha}{2} (|\uparrow\rangle\langle \downarrow| + |\uparrow\rangle\langle \downarrow|) \right] \quad (7.4)$$

$$\mathbb{1}_s = |\uparrow\rangle\langle \uparrow| + |\downarrow\rangle\langle \downarrow| \quad (7.5)$$

We measure the intensities of the output beams using two ^3He integrating detectors called D_0 and D_1 , corresponding to projective measurements of the states $|0\rangle$ and $|1\rangle$ respectively. This performs a Z -basis measurement on the neutron path subsystem. By including spin-filters which selectively transmit neutrons with a preferred spin we may also perform post-selected spin measurements. This allows us to perform joint measurements on the spin and path of the neutron beam.

In a typical NI experiment a relative phase of ϕ is induced between the two paths by a phase flag between the second and third blades which effectively implements the Z -rotation gate:

$$R_z(\phi) = e^{-i\phi/2} |0\rangle\langle 0| + e^{i\phi/2} |1\rangle\langle 1|. \quad (7.6)$$

The relative phase ϕ parameterizes the measured beam intensity by controlling the interference between the two beam paths recombined at the third blade.

Ideally the input beam is in the spin-up polarized state $\psi_{in} = |0\rangle \otimes |\uparrow\rangle$ with respect to a uniform magnetic field in the z -direction. In practice however one is not able to perfectly

polarize the input neutron beam and in general we describe the input beam by the state

$$\rho_{in}(\epsilon) = |0\rangle\langle 0| \otimes \left(\frac{1+\epsilon}{2} |\uparrow\rangle\langle\uparrow| + \frac{1-\epsilon}{2} |\downarrow\rangle\langle\downarrow| \right) \quad (7.7)$$

where $-1 \leq \epsilon \leq 1$ parameterizes the spin-polarization of the neutron beam.

7.2.1 Output Intensities

In an ideal neutron interferometer interference effects are observed in the measured output intensity at each detector. The ideal output intensity is a function of the relative phase between interferometer paths and the angle of spin rotation in the $|0\rangle$ path. If no measurement is performed on the neutron spin subsystem, the ideal detector probabilities in the absence of noise are given by

$$\begin{aligned} D_{0,\text{Ideal}}(\phi, \alpha) &= \frac{1}{2} \left[1 + \cos\left(\frac{\alpha}{2}\right) \cos(\phi) \right] \\ D_{1,\text{Ideal}}(\phi, \alpha) &= \frac{1}{2} \left[1 - \cos\left(\frac{\alpha}{2}\right) \cos(\phi) \right]. \end{aligned} \quad (7.8)$$

which is independent of the spin-polarization of the neutron beam. In a real NI the blades do not generally have equal transmission and reflection coefficients and hence are not true 50-50 beam splitters. This doesn't effect the interference effects at detector D_0 though since both interferometer paths to this detector have the same number of transmissions and reflections.

In practice neutron interferometers cannot be machined perfectly and surface imperfections in the crystal blades lead to a distribution of phases over the cross-sectional area of the neutron beam. This results in reduced contrast of the beam intensity when averaged over the beam distribution. To include the effect of phase noise in our model we consider the output intensities with a phase shift $\phi + \phi_r$ where ϕ_r is an additional random phase shift introduced between paths by the NI blades. This random phase is assumed to be normally distributed with mean 0 and variance σ . By averaging over the distribution of ϕ_r , we may obtain the average detector intensities:

$$D_0(\phi, \alpha, \sigma) = \int_{-\infty}^{\infty} d\phi_r D_{0,\text{Ideal}}(\phi + \phi_r, \alpha) \frac{\exp\left(-\frac{\phi_r^2}{2\sigma^2}\right)}{\sqrt{2\pi\sigma^2}} \quad (7.9)$$

$$= \frac{1}{2} \left[1 + e^{-\sigma^2/2} \cos\left(\frac{\alpha}{2}\right) \cos(\phi) \right] \quad (7.10)$$

7.2.2 Contrast

The intensity curves for each detector as a function of the relative phase ϕ between interferometer paths are referred to as *contrast curves*. They are analogous to the interference pattern produced by a double slit interference experiment. The difference between the maximum and minimum intensity of the D_0 -detector as a function of a phase-flag rotation ϕ is called the *contrast* of the NI and is defined as

$$C_P = \frac{\max_{\phi}[D_0(\phi)] - \min_{\phi}[D_0(\phi)]}{\max_{\phi}[D_0(\phi)] + \min_{\phi}[D_0(\phi)]}. \quad (7.11)$$

The contrast may take values $0 \leq C_P \leq 1$ and is a measure of the strength of quantum coherence between the paths.

We also consider an alternative contrast expression where our parameter of variation in detector intensity is the angle of spin rotation α rather than the phase rotation ϕ as in Eq. (7.11). We define an alternative contrast expression called the *spin-contrast* to be given by

$$C_S = \frac{\max_{\alpha}[D_0(\alpha)] - \min_{\alpha}[D_0(\alpha)]}{\max_{\alpha}[D_0(\alpha)] + \min_{\alpha}[D_0(\alpha)]}. \quad (7.12)$$

We will refer to the standard contrast as the *path-contrast* to distinguish it from the spin-contrast.

Using the observed detector probability in Eq. (7.10) we may calculate that the path and spin contrasts of the noisy three blade NI:

$$C_P(\alpha, \sigma) = e^{-\sigma^2/2} \left| \cos\left(\frac{\alpha}{2}\right) \right| \quad (7.13)$$

$$C_S(\phi, \sigma) = e^{-\sigma^2/2} |\cos(\phi)| \quad (7.14)$$

We see here that the average contrast and spin-contrast expressions are equivalent but with the roles of α and ϕ interchanged ($C_S(\phi, \sigma) = C_P(2\phi, \sigma)$), and depend on the noise strength, and the phase of the parameter that is not optimized over for the contrast (α for path contrast and ϕ for spin-contrast).

7.3 Quantum Correlations in a Neutron Interferometer

We now consider quantum correlations in the output state of a three-blade NI and will follow with the mathematical model of the NI used to derive them [Section 7.2](#). In our

configuration systems A and B correspond to the path and spin degrees of freedom of a neutron respectively, which each may be modelled as a two-level quantum system (qubit). By performing a controlled spin-rotation of angle $0 \leq \alpha \leq 2\pi$ in one of the paths of the NI we may introduce entanglement between the spin and path subsystems of an initially spin-polarized neutron beam. In a realistic NI there are noise sources which introduce decoherence and reduce the effectiveness of this entangling operation. In the present paper we consider the decoherence due to surface defects of the NI blades. This noise source introduces a random phase between the two interferometer paths which degrades the coherence of the path subsystem A .

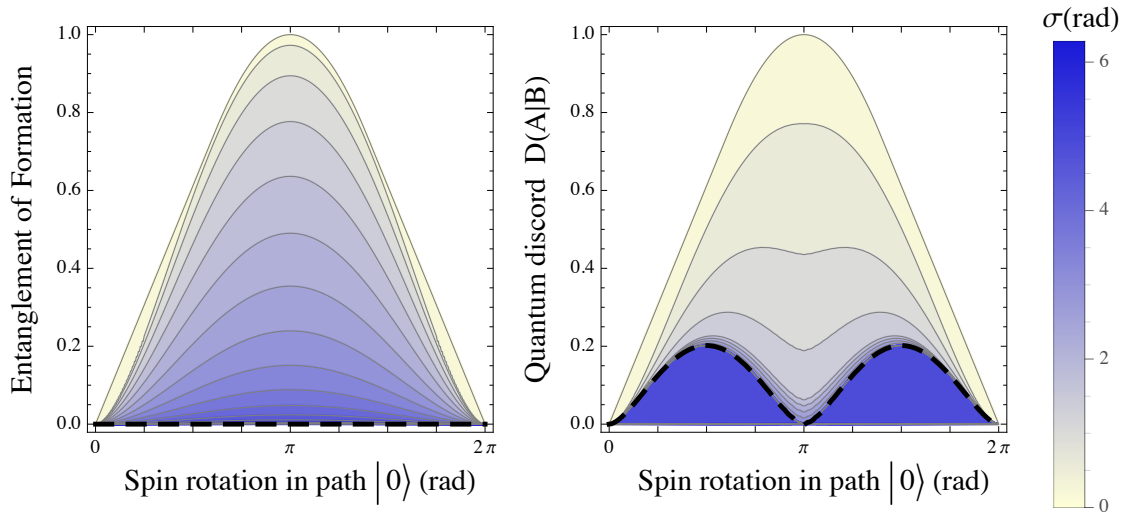


Figure 7.2: Entanglement of formation (left) and quantum discord $D(A|B)$ (right) between the spin and path degrees of freedom of neutrons exiting a three-blade NI as a function of the spin rotation angle of neutrons in the $|0\rangle$ interferometer path, and noise strength σ . The NI schematic is described in Fig. 7.1, and the noise model considered introduces a normally distributed random phase, with mean 0 and standard deviation σ , between the NI paths. The dashed line corresponds to the maximum noise case of a uniform distribution of angles. While the entanglement approaches zero for all spin rotation angles as the noise strength increases, the quantum discord remain non-zero.

Since the neutrons exiting the NI may be described by a mixed state of a 2 qubit quantum system, we use entanglement of formation (EOF) as a measure of the entanglement in the output state. Further, since the quantum state of the neutrons is rank-two we need only perform the minimization in Eq. (1.37) over PVMs on the spin subsystem to calculate the

quantum discord. We find that the EOF between the spin and path systems goes to zero asymptotically as the strength of the random phase noise increases, while the QD remains non-zero for all values of the spin rotation except $n\pi$ for integer values of n . This is illustrated in Fig. 7.2. Even though there is no entanglement between the spin and path of the neutrons in the case of strong phase noise, the non-zero quantum discord $D(A|B)$ indicates the presence of non-classical correlations. This signifies that measurements performed on the neutron spin will induce a disturbance on the path state of the output neutron beam.

The observed entanglement evolution under an increase in the strength of the phase noise can be classed as *approaching* [ZCFJ12], in contrast to entanglement sudden death [YE09]. QD has been shown to be robust to sudden death and instead asymptotically vanishes in bipartite systems subject to Markovian evolution [WSFVB09, FAC⁺10], however in our case QD remains asymptotically non-zero for most spin rotation values. Similar effects of the vanishing of entanglement but non-vanishing QD have been previously found in the theoretical analysis of the evolution of coupled quantum dots under decoherence [FCC10]. Certain initially correlated two-atom states have also been shown to have a non-vanishing quantum discord when coupled to a common dissipative cavity [ZZXG11].

7.3.1 Output Intensity with Spin-filtering

We now consider the detector intensities and contrast curves when we include a spin-filter to perform a post-selected spin measurement on the output neutron beam before detector D_0 . The spin-filter implements a post-selected projective measurement of the pure state

$$|S(\theta, \varphi)\rangle = \cos\left(\frac{\theta}{2}\right) |\uparrow\rangle + e^{i\varphi} \sin\left(\frac{\theta}{2}\right) |\downarrow\rangle \quad (7.15)$$

where θ, ϕ are the spherical coordinates parameterizing the state on the Bloch Sphere. In practice the spin-filter acts by absorbing neutrons in the orthogonal spin state before they reach the detector. After post-selection the detector intensity is proportional to

$$\begin{aligned} D_{0,S(\theta,\varphi)}(\phi, \alpha, \sigma) = \frac{1}{4} & \left(1 + \epsilon \cos^2(\alpha/2) \cos(\theta) + \frac{\epsilon}{2} \sin(\alpha) \sin(\theta) \sin(\varphi) \right. \\ & + e^{-\frac{\sigma^2}{2}} \cos(\alpha/2) (1 + \epsilon \cos(\theta)) \cos(\phi) \\ & \left. - e^{-\frac{\sigma^2}{2}} \sin(\alpha/2) \sin(\theta) \cos(\phi) [\sin(\phi) - \epsilon \sin(\varphi)] \right) \end{aligned} \quad (7.16)$$

We explicitly consider two cases, spin-filtering in the same axes as the quantizing magnetic field (Z -filter), and spin-filtering in an orthogonal basis (X -filter). These are given

by

$$\begin{aligned} Z : \quad |\uparrow\rangle &= |S(0, 0)\rangle & |\downarrow\rangle &= |S(\pi, 0)\rangle \\ X : \quad |\uparrow x\rangle &= |S(\pi/2, 0)\rangle & |\downarrow x\rangle &= |S(3\pi/2, 0)\rangle \end{aligned} \quad (7.17)$$

in terms of the (θ, ϕ) parameterization in Eq. (7.15).

In these case of the Z -filter the observed intensities at detector D_0 are proportional to

$$D_{0,\uparrow z}(\phi, \alpha, \sigma, \epsilon) = \left(\frac{1+\epsilon}{2}\right) D_0(\phi, \alpha, \sigma) + \frac{\epsilon}{8} [\cos(\alpha) - 1] \quad (7.18)$$

$$D_{0,\downarrow z}(\phi, \alpha, \sigma, \epsilon) = \left(\frac{1-\epsilon}{2}\right) D_0(\phi, \alpha, \sigma) - \frac{\epsilon}{8} [\cos(\alpha) - 1] \quad (7.19)$$

for spin-up and spin-down filtering in the z -direction respectively. Note that in this case the normalization condition for the output probabilities is that $D_{0,\downarrow} + D_{0,\uparrow} + D_{1,\downarrow} + D_{1,\uparrow} = 1$.

In these cases of the X -filter the observed intensities at detector D_0 are proportional to

$$D_{0,\uparrow x}(\phi, \alpha, \sigma) = \frac{1}{4} \left[1 + e^{-\sigma^2/2} \cos\left(\frac{\alpha}{2} + \phi\right) \right] \quad (7.20)$$

$$D_{0,\downarrow x}(\phi, \alpha, \sigma) = \frac{1}{4} \left[1 + e^{-\sigma^2/2} \cos\left(\frac{\alpha}{2} - \phi\right) \right] \quad (7.21)$$

for spin-up and spin-down filtering in the x -direction respectively. We see here that the Z -filter adds an additional term to the unfiltered contrast, while the X -filtering combines the parameters ϕ and α into a single argument of a cosine function. Further, in the weak noise case ($\sigma \approx 0$) both these expressions are observably different from the non-spin-filtered case in Eq. (7.10). However in the case of strong noise, the non-spin-filtered and X -filtered intensities approach constant values. Only the Z -filtered intensities are observably different to the non-spin-filtered intensity, and depend on the initial spin polarization ϵ , and the controlled spin-rotation angle α . We discuss the implications of these results in Section 7.3.3, but first we introduce a measure of coherence in interferometer experiments called *contrast*.

7.3.2 Contrast with spin-filtering

We now consider the theoretical path and spin-contrasts of the output beam after spin-filtering. When we post-select on the spin-up and spin-down states of the X -filter we

obtain contrast values of

$$C_{P(\uparrow x)}(\sigma) = C_{S(\uparrow x)}(\sigma) = e^{-\sigma^2/2} \quad (7.22)$$

$$C_{P(\downarrow x)}(\sigma) = C_{S(\downarrow x)}(\sigma) = e^{-\sigma^2/2}. \quad (7.23)$$

We find that the spin and path contrasts are equivalent and depend only on the strength of the phase noise. In particular the contrast decreases to zero with the increase in noise strength.

For the Z -filtered intensities we obtain post-selected path-contrasts of

$$C_{P(\uparrow z)}(\alpha, \sigma, \epsilon) = \left| \frac{(1 + \epsilon)e^{-\sigma^2/2} \cos\left(\frac{\alpha}{2}\right)}{1 + \frac{\epsilon}{2}(1 + \cos(\alpha))} \right| \quad (7.24)$$

$$C_{P(\downarrow z)}(\alpha, \sigma, \epsilon) = \left| \frac{(1 - \epsilon)e^{-\sigma^2/2} \cos\left(\frac{\alpha}{2}\right)}{1 - \frac{\epsilon}{2}(1 + \cos(\alpha))} \right|. \quad (7.25)$$

which satisfy

$$C_{P(\uparrow z)}(\alpha, \sigma, \epsilon) \geq C_P(\alpha, \sigma) \geq C_{P(\downarrow z)}(\alpha, \sigma, \epsilon) \quad (7.26)$$

for $\epsilon \geq 0$, with equality in the case of an zero spin-polarization ($\epsilon = 0$). In particular we see that $C_{path, \downarrow z}(\alpha, \sigma, 1) = 0$.

The spin-contrasts for the Z -filtered intensities are more complicated as the values of α which obtain the minimum for the detector intensities are in general functions of ϕ , ϵ and σ . For the spin-up Z -filter we have

$$C_{S(\uparrow z)}(\phi, \sigma, \epsilon) = \frac{\epsilon + (1 + \epsilon)C_S(\phi, \sigma) + C_S(\phi, \sigma)^2}{2 + \epsilon + (1 + \epsilon)C_S(\phi, \sigma) - C_S(\phi, \sigma)^2}. \quad (7.27)$$

For the spin-down Z -filter, in the range of $\frac{1}{3} \leq \epsilon \leq 1$, we have

$$C_{S(\downarrow z)}(\phi, \sigma, \epsilon) = \frac{\epsilon - (1 - \epsilon)C_S(\phi, \sigma) + \frac{(1-\epsilon)^2}{4\epsilon}C_S(\phi, \sigma)^2}{2 - \epsilon + (1 - \epsilon)C_S(\phi, \sigma) - \frac{(1-\epsilon)^2}{4\epsilon}C_S(\phi, \sigma)^2}. \quad (7.28)$$

For the specific case of unpolarized neutrons ($\epsilon = 0$) we have

$$C_{S(\uparrow z)}(\phi, \sigma, 0) = C_{S(\downarrow z)}(\phi, \sigma, 0) = C_S(\phi, \sigma) \quad (7.29)$$

and in the case of perfect polarization ($\epsilon = 1$) we find that for spin-down Z-filtering we have perfect spin-contrast:

$$C_{S(\downarrow z)}(\phi, \sigma, 1) = 1 \quad (7.30)$$

In the case of strong noise the Z-filtered spin-contrast expressions reduce to

$$C_{S(\uparrow z)}(\phi, \infty, \epsilon) = \frac{\epsilon}{2 + \epsilon} \quad (7.31)$$

$$C_{S(\downarrow z)}(\phi, \infty, \epsilon) = \frac{\epsilon}{2 - \epsilon}. \quad (7.32)$$

and depend only on the initial polarization ϵ of the neutron beam. In practice strong noise amounts to $\sigma \geq 2\pi$.

7.3.3 Interpretation of Spin-Post Selection Experiments

We now discuss the significance of previously calculated path-contrast and spin-contrast values for the noisy 3-blade neutron interferometer. In the absence of spin-filtering, while both the path-contrast and spin-contrast of the ideal 3-blade NI go to zero as the noise strength σ increases, as shown in Fig. 7.2, there is a non-zero quantum discord $D(A|B)$ between the spin and path subsystems. This implies that if we implement a measurement on the spin system the output intensities of the path system must be affected. By using a spin-filter we are able to post-select on an outcome arbitrary PVM on the spin neutron system, however to observe the influence of the spin-filter we are restricted by only being able to measure the path subsystem in the $|0\rangle, |1\rangle$ basis due to the inability to change the final blade of the NI. Hence when restricted to a single measurement basis this influence may not be observable for all spin post-selected states.

No spin post-selection

In the absence of spin-filtering we found that the path-contrast for the noisy neutron interferometer as given in Eq. (7.13), dependent only on the noise strength σ and the angle of controlled spin-rotation α . In the absence of noise, as we increase the angle of spin-rotation up to a $\alpha = \pi$ the measured contrast reduces to zero. At $\alpha = \pi$ the spin and path subsystems are maximally entangled, as shown by an EOF of 1 in Fig. 7.2. By not measuring the spin subsystem we are performing a partial trace over this subsystem which, in the case of a maximally entangled state, results in a maximally mixed reduced state of

the path subsystem, and hence zero contrast. This may be interpreted as having performed a *which-way* measurement of the path taken by the neutron through the interferometer. The neutrons passing the spin-filter are marked to have spin-down, while the neutrons which don't go through the arm with the spin-rotator will all have spin-up. By tuning $0 < \alpha < \pi$ we may control the strength of this which-way marking of the neutrons. For α close to 0 it becomes a weak which-way marking of the path taken by the neutrons through the interferometer, and hence we still retain some contrast.

In the case of spin-contrast, as given in Eq. (7.14), we have the same situation but with the roles of the rotation angle α and phase-flag ϕ reversed. In this case we are doing a spin-based magnetic interference experiment, and the relative phase between paths now performs the which-way marking of the neutron. In both cases the presence of noise reduces the value of contrast, until it is approximately zero at $\sigma = 2\pi$. This would suggest that the random phase noise destroys all relative phase information, and hence coherence, between the two paths in the interferometer. However due to the non-zero discord between the path and spin of the neutron we may attempt to recover some information by spin-measurements.

X-filter spin post-selection

In the case of X -filtering we found that both the path-contrast and spin-contrast when spin-filtering on the $|\uparrow x\rangle$ or $|\downarrow x\rangle$ spin-states depend only on the strength σ of the random-phase noise, as shown in Eq. (7.22). This is because the X -filter post-selection acts to combine the parameters α and ϕ into a single relative phase parameter $\phi + \alpha/2$ between the two NI paths which is observed at the detector. For path-contrast the spin rotation angle only shows up as a shift of the contrast curves without changing the actual contrast value. In effect the X -filter has erased the which-way marking of the neutrons in the NI due to the controlled spin-rotation angle. Similarly, for the spin-contrast the roles of ϕ and α are swapped with the spin-filter now erasing all effect of the phase-flag parameter on the output intensities. This is analogous to a *quantum eraser* in optics [SD82]. By post-selecting on the neutron spin in the x -direction we have erased the which-way measurement caused by entanglement between the spin and path neutron subsystems. However, as the noise strength increases the spin-filtered spin and path contrasts both reduce to zero and become indistinguishable from the unfiltered contrast.

Z-filter spin post-selection

When implementing a Z -filter post-selection we calculated quite different values for the spin-contrast and path-contrast. In the case where we post-select on the $|\uparrow\rangle$ spin-state,

the path-contrast in Eq. (7.24) is maximized for a perfectly polarized input ($\epsilon = 1$), as we are in effect filtering out only the portion of neutrons rotated away from $|\uparrow\rangle$ by the spin-rotator. In this sense, much like the non-post-selected case, the angle of rotation controls the strength of the which-way measurement. If instead we filter on $|\downarrow\rangle$ as given by Eq. (7.25), then we find the contrast is zero for $\epsilon = 1$. This is because in this case we are post-selecting only on the spins that were rotated, and hence we are performing a perfect which-way measurement of the path taken by the neutrons and cannot have any path based interference effects. If the incoming beam is not perfectly polarized our which-way measurement is effectively noisy and we will have a fraction of unrotated neutrons which still have spin-down polarization. In this case, as with the $|\uparrow\rangle$ filter, the angle of rotation α controls the relative strength of the which-way measurement.

For the spin-filtered spin-contrast we find that with perfect polarization the spin-filtered contrast with spin-down post selection is always 1. However with $\epsilon < 1$ the value of contrast will depend on the phase-flag ϕ , which acts as the which-way marking. As with the unfiltered case it will be maximum for $\phi = 0$, and zero at $\phi = \pi/2$.

As the noise strength increases, the dependence of ϕ is removed as we are decohering the relative phase information between paths. Hence the noise is erasing the which-way marking due to the phase-flag on the spin-contrast. In the strong noise case we find that the spin-contrast depends only on the initial polarization ϵ . If $\epsilon = 1$ we are filtering out all spins that are not rotated to $|\downarrow\rangle$, so our measured intensity is a function of the rotation angle. In the $\epsilon < 1$ case, we are effectively introducing spin noise into the system as there will now be $(1 - \epsilon)/2$ portion of neutrons with spin-down in the non-rotated path, thus reducing the spin-contrast. For the spin-up filter we have a similar situation, however the contrast is no longer unity for $\epsilon = 1$ unless $\phi = 0$ and $\sigma = 0$. In this case we are filtering out the percentage of neutrons rotated to spin-down by the spin-rotator, rather than post-selecting on them.

It has been suggested this setup might be used to demonstrate the so-called *quantum cheshire cat* paradox [APRS13]. This paradox is to weakly perform two measurements of the path a neutron takes through the interferometer simultaneously. One that is spin-based and determines that the neutron spin goes down one arm of the interferometer, and another that is not spin-based and determines that the neutron itself went down the other interferometer arm. By doing the which-way marking with a spin-filter we may measure which path the neutron spin went down by a spin based measurement. By varying α we may control the strength of this measurement. To complete the experiment would require implementing a second weak measurement simultaneously to suggest that the neutron itself was observed to go down a different path to its spin degree of freedom. This has been suggested to be implemented in a NI by using a partial absorber in the interferometer path

without the spin rotator [DGS⁺14], however the reported experiment did not implemented both the spin based, and absorption based measurements simultaneously.

7.3.4 Experimental Demonstration

Some of the theoretical results from Section 7.2.2 were demonstrated experimentally by collaborators M.O. Abutaleb, M.G. Huber and D.A. Pushin at the National Institute of Standard and Technology (NIST) [WAH⁺14]. However, temperature variations caused by the method of implementing the spin-rotator resulted in a phase drift which increases the effective strength of the phase noise in our NI. Hence while we can calculate path-contrast in the absence of spin-rotation, for the spin-contrast experiments the spin-rotator acts to increase the apparent phase noise so that $e^{-\sigma^2/2} \approx 0$. With the increase in phase noise the spin-contrast with no spin-filtering, and with X-filtering is expected to be approximately zero. We may only observe the Z-filtered contrast which in the strong noise case only depends on the neutron polarization. This Z-filter post-selection demonstrates the disturbance of the path state of the neutrons by measurement of the neutron spin in the presence of strong noise, as indicated by the non-zero QD as shown in Fig. 7.2.

We compared the contrast and spin-filtered contrasts after post-selecting on spin-down neutrons, quantized in a static magnetic field in the z -direction, for three NIs using the setup shown in Fig. 7.1. The experiment was performed at the NIST Center for Neutron Research’s Neutron Optics and Interferometer Facility, located at Gaithersburg, Maryland [PHA⁺15, NIS15]. This facility has an excellent vibration isolation and temperature stability thus allows for a good and long phase stability [PAHC08].

Our neutron beam consisted of 0.271nm wavelength neutrons, and the incident neutron beam was polarized via a transmission mode supermirror polarizer [APH⁺12] giving an initial polarization of 93% spin-up. The path-selective spin rotation was implemented using thin permalloy films deposited on Si substrate [Pyn05]. Spin-filters were implemented using either Heusler crystals or reflection-mode curved supermirrors. These were preceded by an adiabatic coil used to rotate the neutron spin so that spin-up neutrons were absorbed, and spin-down neutrons were transmitted. During this experimental work we have used two LLL type NI with different initial contrasts: “good” and “bad”, which we refer to as N_1 and N_2 respectively. To compare spin-contrast with a very low contrast NI under the same environmental conditions we used the good NI and introduced a large destructive phase gradient by adding a 45 degree fused silica wedge in one interferometer path [PCA⁺07]. We refer to the good NI with the wedge as N_3 .

The measured contrast curves in the absence of spin-filtering for the three NIs is shown

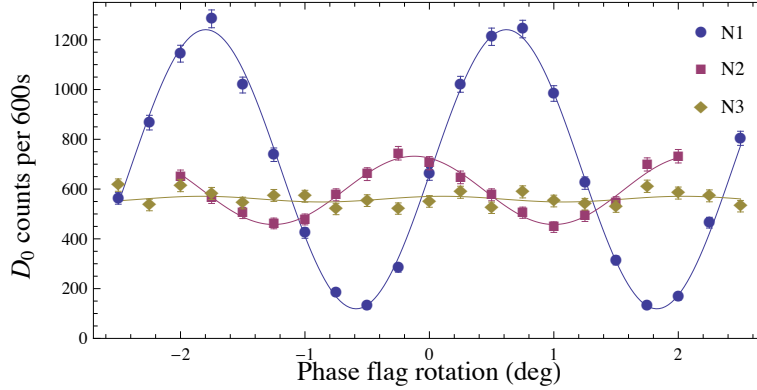


Figure 7.3: Measured intensity curves at detector D_0 as a function of phase-flag rotation for three NIs. The corresponding path-contrast values are $C_{P1} = (82.5 \pm 1.3)\%$, $C_{P2} = (23 \pm 1.5)\%$, $C_{P3} = (2 \pm 1.7)\%$ for interferometers N_1, N_2, N_3 respectively.

in Fig. 7.3, these correspond to contrast values of $C_{P1} = (82.5 \pm 1.3)\%$, $C_{P2} = (23 \pm 1.5)\%$, $C_{P3} = (2 \pm 1.7)\%$ for the interferometers N_1, N_2 and N_3 respectively. These contrast values correspond to standard deviations of $\sigma_1 = 0.62 \pm 0.03$, $\sigma_2 = 1.71 \pm 0.04$, $\sigma_3 = 2.80 \pm 0.61$ respectively in the noise model under consideration.

After applying the spin-down filter, the spin-filtered contrasts were found to be $C_{S1(\downarrow z)} = (78.0 \pm 3)\%$, $C_{S2(\downarrow z)} = (74.2 \pm 2.2)\%$, $C_{S3(\downarrow z)} = (84 \pm 4)\%$, as shown in Fig. 7.4. Our theoretical model with an initial neutron spin polarization of $(1 + \epsilon)/2 = 93\%$ predicts a spin-contrast of 75.3% for all three interferometers.

7.4 Summary

We have theoretically and experimentally investigated the role of quantum correlations in a simple bipartite quantum system in the presence of noise by using the spin and path degrees of freedom of a polarized neutron beam in a neutron interferometer. If we initially entangle the the path and spin degrees of freedom of a neutron beam by a path dependent spin-rotation, we found that that phase noise acts to reduce the amount of entanglement to zero as the noise strength increases. However a non-zero value of quantum discord $D(A|B)$ for all noise strengths indicates that there are still non-classical correlations between the neutrons spin and path degrees of freedom. The non-zero QD indicates that spin-measurements will

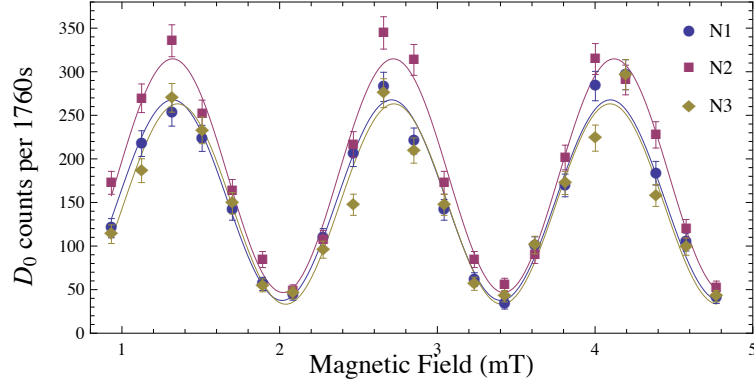


Figure 7.4: Measured intensity curves at detector D_0 as a function of spin-rotation for three NIs where we have applied a spin-filter on the output beam to select spin-down neutrons with respect to a static magnetic field in the z -direction. The corresponding spin-filtered contrast values are $C_{S1(\downarrow z)} = (78.0 \pm 3)\%$, $C_{S2(\downarrow z)} = (74.2 \pm 2.2)\%$, $C_{S3(\downarrow z)} = (84 \pm 4)\%$ for interferometers N_1, N_2, N_3 respectively.

have an influence on the quantum state of the neutron path subsystem, however due to the experimental limitations we are only able to perform measurements of the path subsystem in the basis corresponding to the beam paths as implemented by the physical neutron detectors. Restricted to this measurement basis, we are not able to see a noticeable effect for all projective measurements in the strong noise limit.

In the low noise case our analysis showed that we may think of the spin-path NI as a quantum eraser. In the absence of spin-filtering by rotating the spin state of a neutron in only one path of the interferometer we are labelling the neutrons which take this path and performing a which-way measurement of the neutron's path through the interferometer. This results in a loss of contrast proportional to the entanglement of the path and neutrons. By implementing a post-selected spin measurement in the x -direction we may erase this labelling data and restore contrast. This also held true for the spin-contrast, but with the roles of the phase flag and controlled spin rotation angle interchanged. However in the strong noise case, the X -filtered path and spin contrast both reduce to zero and so are not observably different from the non-spin-filtered contrast. Thus the effect of x -basis spin measurements on the path subsystem state are not directly observable in the NI in the presence of strong dephasing noise.

In the case of spin-contrast with post-selected spin measurement in the z -direction, the

contrast remains a function of the spin rotation angle but removes the effect of the phase noise. In the high noise case the expression for spin-contrast when we perform a Z-filter and post-select on the spin down state is a function of spin polarization only. Hence even in the high noise case we are able to experimentally observe the effect of spin-filtering on the path subsystem. Our experimental results agree with our theoretical model predicting an increase in spin-filtered contrast over phase contrast for three NIs when spin-filtering has been performed on the Spin-down state in the z -direction. The deviations between our measured spin-filtered contrast the value predicted by our theoretical model are consistent with phase variations over the acquisition time due to temperature and humidity fluctuations in the NI environment. We interpreted this non-zero quantum discord as a signature that even in the presence of strong phase noise, the NI still exhibits genuine quantum behaviour.

Chapter 8

Conclusion

In this thesis we have made several contributions to the field of open quantum systems theory. The graphical calculus and techniques we have developed provide an intuitive and pedagogical means of representing and unifying the transformations between the different representations of CP-maps, and as we demonstrated numerous times provided an elegant tool for developing new proofs, and constructions by the manipulation of the corresponding tensor networks. The culmination of this line of research was the development of the superchannel formalism which allows for the operational description of a strictly greater set of dynamics than standard quantum channels. This was demonstrated by the IC superchannel which allows for the complete characterization of an open quantum system which is initially correlated with its environment. This is an important example to consider as such a situation is one of the simplest examples of a non-Markovian environment — the initial correlations can be thought of as the result of a memory effect between the environment and system from some previous evolution or computation. Non-Markovian effects are becoming more and more important in current quantum devices which are able to achieve strong coupling between the system and control elements, however most previous work on non-Markovian quantum systems has been focused on the underlying master equations describing the interaction, and previously presented measures of non-Markovianity are difficult to measure experimentally and do not always have a clear operational interpretation. The work presented here thus provides an important first step in the operational characterization of general non-Markovian dynamics.

The work presented on the initialization of an ensemble quantum system in a high purity state via means of cavity cooling with a high-Q resonator is an important tool for state initialization in ensemble spin based quantum processors. Ensemble spin systems are a natural candidate for long life-time quantum memories in hybrid quantum architectures.

However this long-life time also prevents the rapid reset of spins via thermal relaxation processes, and thus cooling protocols that can short-circuit the thermal relaxation process become necessary. The side-band cavity cooling protocol we presented in principle allows for the gated on-demand reset of an ensemble spin system by selectively tuning the spins on resonance with the sideband of a high-Q single mode cavity. In addition as quantum systems begin to scale from 10s to 100s of qubits, the parallel removal of entropy via collective dynamics becomes critical as independently controlling and addressing each spin becomes a significant challenge. Thus the work presented here provides an important tool for state initialization in a large scale spin based quantum architecture.

Finally the work presented on quantum correlations in neutron interferometry showed that even in the presence of phase noise on the path degree of freedom, we can still observe quantum effects by coupling to the spin degree of freedom. By including a spin-filter on the output beam of the interferometer we were able to relate this to which-way interferometry measurements and to the quantum eraser. Our discussion of spin-contrast, which was found to be robust under phase noise, may be useful for developing neutron interferometer based characterization methods for magnetic materials.

References

- [AAA⁺13] J Aasi, J Abadie, BP Abbott, R Abbott, TD Abbott, MR Abernathy, C Adams, T Adams, P Addesso, RX Adhikari, et al. Enhanced sensitivity of the ligo gravitational wave detector by using squeezed states of light. *Nature Photonics*, 7(8):613–619, 2013. [doi:10.1038/nphoton.2013.177](https://doi.org/10.1038/nphoton.2013.177).
- [ABJ⁺03] J. B. Altepeter, D. Branning, E. Jeffrey, T. C. Wei, P. G. Kwiat, R.T. Thew, J. L. O’Brien, M. A. Nielsen, and A. G. White. Ancilla-assisted quantum process tomography. *Phys. Rev. Lett.*, 90:193601, 2003. [arXiv:quant-ph/0303038](https://arxiv.org/abs/quant-ph/0303038), [doi:10.1103/PhysRevLett.90.193601](https://doi.org/10.1103/PhysRevLett.90.193601).
- [Abr61] A. Abragam. *The Principles of Nuclear Magnetism*. Oxford University Press, 1961.
- [AC04] S. Abramsky and B. Coecke. A categorical semantics of quantum protocols. In *IEEE S. Log.*, 2004, pages 415–425, 2004. [arXiv:quant-ph/0402130v5](https://arxiv.org/abs/quant-ph/0402130v5), [doi:10.1109/LICS.2004.1319636](https://doi.org/10.1109/LICS.2004.1319636).
- [Aga74] G. S. Agarwal. *Quantum Optics*. Springer-Verlag, Berlin, 1974.
- [AKN⁺11] R. Amsüss, Ch. Koller, T. Nöbauer, S. Putz, S. Rotter, K. Sandner, S. Schneider, M. Schramböck, G. Steinhauser, H. Ritsch, J. Schmiedmayer, and J. Majer. Cavity qed with magnetically coupled collective spin states. *Phys. Rev. Lett.*, 107:060502, Aug 2011. [arXiv:1103.1045](https://arxiv.org/abs/1103.1045), [doi:10.1103/PhysRevLett.107.060502](https://doi.org/10.1103/PhysRevLett.107.060502).
- [APH⁺12] M. O. Abutaleb, D. A. Pushin, M. G. Huber, C. F. Majkrzak, M. Arif, and D. G. Cory. Design of remnant magnetization fecov films as compact, heatless neutron spin rotators. *Applied Physics Letters*, 101:182404, 2012. [doi:10.1063/1.4765069](https://doi.org/10.1063/1.4765069).

- [APRS13] Y. Aharonov, S. Popescu, D. Rohrlich, and P. Skrzypczyk. Quantum cheshire cats. *New Journal of Physics*, 15(11):113015, 2013. [arXiv:1202.0631](#), [doi:10.1088/1367-2630/15/11/113015](#).
- [BB11] V. Bergholm and J. D. Biamonte. Categorical quantum circuits. *J. Phys. A: Math. Theor.*, 44:245304, 2011. [arXiv:1010.4840v2](#), [doi:10.1088/1751-8113/44/24/245304](#).
- [BBGP⁺13] C. Belthangady, N. Bar-Gill, L.M. Pham, K. Arai, D. Le Sage, P. Cappellaro, and R.L. Walsworth. Dressed-state resonant coupling between bright and dark spins in diamond. *Phys. Rev. Lett.*, 110:157601, 2013. [doi:10.1103/PhysRevLett.110.157601](#).
- [BC12] T.W. Borneman and D.G. Cory. Bandwidth-limited control and ringdown suppression in high-q resonators. *J. Magn. Reson.*, 225:120–129, 2012. [arXiv:1207.1139](#), [doi:10.1016/j.jmr.2012.10.011](#).
- [BCG10] B. Q. Baragiola, B. A. Chase, and J. M. Geremia. Collective uncertainty in partially polarized and partially decohered spin- $\frac{1}{2}$ systems. *Phys. Rev. A*, 81:032104, Mar 2010. [arXiv:0910.3897](#), [doi:10.1103/PhysRevA.81.032104](#).
- [BCJ11] J. D. Biamonte, S. R. Clark, and D. Jaksch. Categorical tensor network states. *AIP Adv.*, 1:042172, 2011. [arXiv:1012.0531](#), [doi:10.1063/1.3672009](#).
- [BCM13] B. Bylicka, D. Chruscinski, and S. Maniscalco. Non-markovianity as a resource for quantum technologies. *ArXiv e-prints*, arXiv:1301.2585 [quant-ph], 2013. [arXiv:1301.2585](#).
- [BCOR09] S. Blanes, F. Casas, J. A. Oteo, and J. Ros. The Magnus expansion and some of its applications. *Phys. Rep.*, 470:151–238, 2009. [arXiv:0810.5488](#), [doi:10.1016/j.physrep.2008.11.001](#).
- [BGC12] T.W. Borneman, C.E. Granade, and D.G. Cory. Parallel information transfer in a multinode quantum information processor. *Phys. Rev. Lett.*, 108:140502, 2012. [arXiv:1107.4333](#), [doi:10.1103/PhysRevLett.108.140502](#).
- [BGTW11] A. Brodutch, A. Gilchrist, D. R. Terno, and C. J. Wood. Quantum discord in quantum computation. *Journal of Physics: Conference Series*,

- 306(1):012030, 2011. [arXiv:1012.1402](#), [doi:10.1088/1742-6596/306/1/012030](#).
- [BH12] S. Boixo and C. Heunen. Entangled and sequential quantum protocols with dephasing. *Phys. Rev. Lett.*, 102:120402, 2012. [arXiv:1108.3569](#), [doi:10.1103/PhysRevLett.108.120402](#).
- [BHW⁺04] A. Blais, R.-S. Huang, A. Wallraff, S. M. Girvin, and R. J. Schoelkopf. Cavity quantum electrodynamics for superconducting electrical circuits: An architecture for quantum computation. *Phys. Rev. A*, 69:062320, 2004. [arXiv:cond-mat/0402216](#), [doi:10.1103/PhysRevA.69.062320](#).
- [BK10a] R. Blume-Kohout. Hedged maximum likelihood quantum state estimation. *Phys. Rev. Lett.*, 105:200504, 2010. [arXiv:1001.2029](#), [doi:10.1103/PhysRevLett.105.200504](#).
- [BK10b] R. Blume-Kohout. Optimal, reliable estimation of quantum states. *New Journal of Physics*, 12(4):043034, 2010. [arXiv:quant-ph/0611080](#), [doi:10.1088/1367-2630/12/4/043034](#).
- [BK13] R. Blume-Kohout. Robust, self-consistent, closed-form tomography of quantum logic gates on a trapped ion qubit. *ArXiv e-prints*, arXiv:1310.4492 [quant-ph], 2013. [arXiv:1310.4492](#).
- [BKPV99] S. Bose, P. L. Knight, M. B. Plenio, and V. Vedral. Proposal for teleportation of an atomic state via cavity decay. *Phys. Rev. Lett.*, 83:5158–5161, 1999. [arXiv:quant-ph/9908004](#), [doi:10.1103/PhysRevLett.83.5158](#).
- [BKS⁺09] H. Bartosik, J. Klepp, C. Schmitzer, S. Sponar, A. Cabello, H. Rauch, and Y. Hasegawa. Experimental test of quantum contextuality in neutron interferometry. *Phys. Rev. Lett.*, 103:040403, 2009. [arXiv:0904.4576](#), [doi:10.1103/PhysRevLett.103.040403](#).
- [BL11] J. C. Baez and A. Lauda. A prehistory of n-categorical physics. In *Deep Beauty*, pages 13–128. Cambridge University Press, 2011. [arXiv:0908.2469](#), [doi:10.1017/CB09780511976971.003](#).
- [BMR⁺05] J. Baugh, O. Moussa, C. A. Ryan, A. Nayak, and R. Laflamme. Experimental implementation of heat-bath algorithmic cooling using solid-state nuclear magnetic resonance. *Nature*, 438(7067):470–473, 11 2005. [doi:10.1038/nature04272](#).

- [BMT⁺13] O.W.B. Benningshof, H.R. Mohebbi, I.A.J. Taminiau, G.X. Miao, and D.G. Cory. Superconducting microstrip resonator for pulsed esr of thin films. *J. Magn. Reson.*, 230:84 – 87, 2013. doi:[10.1016/j.jmr.2013.01.010](https://doi.org/10.1016/j.jmr.2013.01.010).
- [BP70] R. Bonifacio and G. Preparata. Coherent spontaneous emission. *Phys. Rev. A*, 2:336 – 347, 1970. doi:[10.1103/PhysRevA.2.336](https://doi.org/10.1103/PhysRevA.2.336).
- [BS58] N. Bloembergen and P.P. Sorokin. Nuclear magnetic resonance in the cesium halides. *Phys. Rev.*, 110:865–875, 1958. doi:[10.1103/PhysRev.110.865](https://doi.org/10.1103/PhysRev.110.865).
- [BSS⁺12] G. Bachar, O. Suchoi, O. Shtempluck, A. Blank, and E. Buks. Nonlinear induction detection of electron spin resonance. *Appl. Phys. Lett.*, 101:022602, 2012. arXiv:[1204.3588](https://arxiv.org/abs/1204.3588), doi:[10.1063/1.473450](https://doi.org/10.1063/1.473450).
- [Bul87] R. K. Bullough. Photon, Quantum and Collective, effects from Rydberg Atoms in Cavities. *Hyperfine Interactions*, 37:71–108, 1987. doi:[10.1007/BF02395705](https://doi.org/10.1007/BF02395705).
- [BW11] M. C. Butler and D. P. Weitekamp. Polarization of nuclear spins by a cold nanoscale resonator. *Phys. Rev. A*, 84:063407, 2011. doi:[10.1103/PhysRevA.84.063407](https://doi.org/10.1103/PhysRevA.84.063407).
- [BZ06] I. Bengtsson and K. Życzkowski. *Geometry of Quantum States*. Cambridge University Press, Cambridge, 2006.
- [CBHP13] C. Cormick, A. Bermudez, S. F. Huelga, and M. B. Plenio. Dissipative ground-state preparation of a spin chain by a structured environment. *New Journal of Physics*, 15(7):073027, 2013. arXiv:[1304.2201](https://arxiv.org/abs/1304.2201), doi:[10.1088/1367-2630/15/7/073027](https://doi.org/10.1088/1367-2630/15/7/073027).
- [CDP08] G. Chiribella, G. M. D’Ariano, and P. Perinotti. Transforming quantum operations: Quantum supermaps. *EPL (Europhysics Letters)*, 83:30004, 2008. arXiv:[0804.0180](https://arxiv.org/abs/0804.0180), doi:[10.1209/0295-5075/83/30004](https://doi.org/10.1209/0295-5075/83/30004).
- [CDP11] G. Chiribella, Giacomo M. D’Ariano, G. M., and P. Perinotti. Informational derivation of quantum theory. *Phys. Rev. A*, 84:012311, 2011. arXiv:[1011.6451](https://arxiv.org/abs/1011.6451), doi:[10.1103/PhysRevA.84.012311](https://doi.org/10.1103/PhysRevA.84.012311).

- [CGB⁺10] I. Chiorescu, N. Groll, S. Bertaina, T. Mori, and S. Miyashita. Magnetic strong coupling in a spin-photon system and transition to classical regime. *Phys. Rev. B*, 82:024413, 2010. [arXiv:1004.3605](#), [doi:10.1103/PhysRevB.82.024413](#).
- [CGM⁺12] A. Chiuri, C. Greganti, L. Mazzola, M. Paternostro, and P. Mataloni. Linear optics simulation of quantum non-markovian dynamics. *Sci. Rep.*, 2:12, 2012. [arXiv:1208.1630](#), [doi:10.1038/srep00968](#).
- [CH11] B. Coecke and C. Heunen. Pictures of complete positivity in arbitrary dimension. *ArXiv e-prints*, arXiv:1110.3055v1, 2011. [arXiv:1110.3055](#).
- [Cha05] H. F. Chau. Unconditionally secure key distribution in higher dimensions by depolarization. *Information Theory, IEEE Transactions on*, 51:1451–1468, 2005. [arXiv:quant-ph/0405016](#), [doi:10.1109/TIT.2005.844076](#).
- [CHHC06] P. Cappellaro, J. S. Hodges, T. F. Havel, and D. G. Cory. Principles of control for decoherence-free subsystems. *The Journal of Chemical Physics*, 125(4):044514, 2006. [arXiv:quant-ph/0604203](#), [doi:10.1063/1.2216702](#).
- [Cho75] M. D. Choi. Completely positive linear maps on complex matrices. *Linear Algebra Appl.*, 10:285, 1975. [doi:10.1016/0024-3795\(75\)90075-0](#).
- [CJHL09] P. Cappellaro, L. Jiang, J.S. Hodges, and M.D. Lukin. Coherence and control of quantum registers based on electronic spin in a nuclear spin bath. *Phys. Rev. Lett.*, 102:210502, 2009. [arXiv:0901.0444](#), [doi:10.1103/PhysRevLett.102.210502](#).
- [CLLW15] R. Cleve, D. Leung, L. Liu, and C. Wang. Near-linear constructions of exact unitary 2-designs. *ArXiv e-prints*, arXiv:1501.04592 [quant-ph], 2015. [arXiv:1501.04592](#).
- [CMS11] L. C. Céleri, J. Maziero, and R. M. Serra. Theoretical and experimental aspects of quantum discord and related measures. *Int. J. Quantum Inf.*, 09:1837, 2011. [arXiv:1107.3428](#), [doi:10.1142/S0219749911008374](#).
- [CN10] B. Collins and I. Nechita. Random quantum channels I: Graphical calculus and the Bell state phenomenon. *Commun. Math. Phys.*, 297:345, 2010. [arXiv:0905.2313](#), [doi:10.1007/s00220-010-1012-0](#).

- [CN11] B. Collins and I. Nechita. Random quantum channels II: Entanglement of random subspaces, Rényi entropy estimates and additivity problems. *Adv. Math.*, 226:1181, 2011. [arXiv:0906.1877](#), [doi:10.1016/j.aim.2010.08.002](#).
- [COK+89] A. Cimmino, G. I. Opat, A. G. Klein, H. Kaiser, S. A. Werner, M. Arif, and R. Clothier. Observation of the topological aharonov-casher phase shift by neutron interferometry. *Phys. Rev. Lett.*, 63, 1989. [doi:10.1103/PhysRevLett.63.380](#).
- [COW75] R. Colella, A. W. Overhauser, and S. A. Werner. Observation of gravitationally induced quantum interference. *Phys. Rev. Lett.*, 34, 1975. [doi:10.1103/PhysRevLett.34.1472](#).
- [CPP09] B. Coecke, E. O. Paquette, and D. Pavlovic. Classical and quantum structuralism. *ArXiv e-prints*, 0904.1997v2 [quant-ph], 2009. [arXiv:0904.1997](#).
- [CS06] B. Collins and P. Śniady. Integration with respect to the haar measure on unitary, orthogonal and symplectic group. *Communications in Mathematical Physics*, 264(3):773–795, 2006. [arXiv:math/0511253](#), [doi:10.1007/s00220-006-1554-3](#).
- [CS12] B. Coecke and R. W. Spekkens. Picturing classical and quantum Bayesian inference. *Synthese*, 186:651, 2012. [arXiv:1102.2368](#), [doi:10.1007/s11229-011-9917-5](#).
- [CTZ08] H. A. Carteret, D. R. Terno, and K. Życzkowski. Dynamics beyond completely positive maps: Some properties and applications. *Phys. Rev. A*, 77:042113, 2008. [arXiv:quant-ph/0512167](#), [doi:10.1103/PhysRevA.77.042113](#).
- [DCEL09] C. Dankert, R. Cleve, J. Emerson, and E. Livine. Exact and approximate unitary 2-designs and their application to fidelity estimation. *Phys. Rev. A*, 80:012304, 2009. [doi:10.1103/PhysRevA.80.012304](#).
- [DGS+14] T. Denkmayr, H. Geppert, S. Sponar, H. Lemmel, A. Matzkin, J. Tollaksen, and Y. Hasegawa. Observation of a quantum cheshire cat in a matter-wave interferometer experiment. *Nature communications*, 5, 2014. [arXiv:1312.3775](#), [doi:10.1038/ncomms5492](#).

- [Dic54] R. H. Dicke. Coherence in spontaneous radiation processes. *Phys. Rev.*, 93:99–110, Jan 1954. doi:[10.1103/PhysRev.93.99](https://doi.org/10.1103/PhysRev.93.99).
- [DL15] J. M. Dominy and D. A. Lidar. Beyond complete positivity. *ArXiv e-prints*, arXiv:1503.05342 [quant-ph], 2015. arXiv:[1503.05342](https://arxiv.org/abs/1503.05342).
- [DLP01] G. M. D’Ariano and P. Lo Presti. Quantum tomography for measuring experimentally the matrix elements of an arbitrary quantum operation. *Phys. Rev. Lett.*, 86:4195, 2001. arXiv:[quant-ph/0012071](https://arxiv.org/abs/quant-ph/0012071), doi:[10.1103/PhysRevLett.86.4195](https://doi.org/10.1103/PhysRevLett.86.4195).
- [DLP03] G. M. D’Ariano and P. Lo Presti. Imprinting complete information about a quantum channel on its output state. *Phys. Rev. Lett.*, 91:047902, 2003. arXiv:[quant-ph/0211133](https://arxiv.org/abs/quant-ph/0211133), doi:[10.1103/PhysRevLett.91.047902](https://doi.org/10.1103/PhysRevLett.91.047902).
- [DMP00] G.M. D’Ariano, L. Maccone, and M.G.A. Paris. Orthogonality relations in quantum tomography. *Physics Letters A*, 276(1-4):25–30, 2000. arXiv:[quant-ph/0005111](https://arxiv.org/abs/quant-ph/0005111), doi:[10.1016/S0375-9601\(00\)00660-5](https://doi.org/10.1016/S0375-9601(00)00660-5).
- [DN02] J. L. Dodd and M. A. Nielsen. Simple operational interpretation of the fidelity of mixed states. *Phys. Rev. A*, 66:044301, 2002. doi:[10.1103/PhysRevA.66.044301](https://doi.org/10.1103/PhysRevA.66.044301).
- [DPF⁺11] I. Diniz, S. Portolan, R. Ferreira, J. M. Gérard, P. Bertet, and A. Auffèves. Strongly coupling a cavity to inhomogeneous ensembles of emitters: Potential for long-lived solid-state quantum memories. *Phys. Rev. A*, 84:063810, Dec 2011. arXiv:[1101.1842](https://arxiv.org/abs/1101.1842), doi:[10.1103/PhysRevA.84.063810](https://doi.org/10.1103/PhysRevA.84.063810).
- [DSC08] A. Datta, A. Shaji, and C. M. Caves. Quantum discord and the power of one qubit. *Phys. Rev. Lett.*, 100:050502, 2008. arXiv:[0709.0548](https://arxiv.org/abs/0709.0548), doi:[10.1103/PhysRevLett.100.050502](https://doi.org/10.1103/PhysRevLett.100.050502).
- [EAŻ05] J. Emerson, R. Alicki, and K. Życzkowski. Scalable noise estimation with random unitary operators. *J. Opt. B: Quantum Semiclass. Opt.*, 7:S347, 2005. arXiv:[quant-ph/0503243](https://arxiv.org/abs/quant-ph/0503243), doi:[10.1088/1464-4266/7/10/021](https://doi.org/10.1088/1464-4266/7/10/021).
- [EBW⁺87] R. R. Ernst, G. Bodenhausen, A. Wokaun, et al. *Principles of nuclear magnetic resonance in one and two dimensions*, volume 14. Clarendon Press, Oxford, 1987.

- [EEQ⁺10] G.R. Eaton, S.S. Eaton, R.W. Quine, D. Mitchell, V. Kathirvelu, and R.T. Weber. A signal-to-noise standard for pulsed EPR. *J. Magn. Reson.*, 205(1):109–113, 2010. doi:[10.1016/j.jmr.2010.04.006](https://doi.org/10.1016/j.jmr.2010.04.006).
- [ESM⁺07] J. Emerson, M. Silva, O. Moussa, C. Ryan, M. Laforest, J. Baugh, D. G. Cory, and R. Laflamme. Symmetrized characterization of noisy quantum processes. *Science*, 317(5846):1893–1896, 2007. arXiv:[0707.0685](https://arxiv.org/abs/0707.0685), doi:[10.1126/science.1145699](https://doi.org/10.1126/science.1145699).
- [EV09] G. Evenbly and G. Vidal. Algorithms for entanglement renormalization. *Phys. Rev. B*, 79:144108, 2009. arXiv:[0707.1454](https://arxiv.org/abs/0707.1454), doi:[10.1103/PhysRevB.79.144108](https://doi.org/10.1103/PhysRevB.79.144108).
- [FAC⁺10] A. Ferraro, L. Aolita, D. Cavalcanti, F. M. Cucchietti, and A. Acín. Almost all quantum states have nonclassical correlations. *Phys. Rev. A*, 81:052318, 2010. arXiv:[0908.3157](https://arxiv.org/abs/0908.3157), doi:[10.1103/PhysRevA.81.052318](https://doi.org/10.1103/PhysRevA.81.052318).
- [FBK12] C. Ferrie and R. Blume-Kohout. Estimating the bias of a noisy coin. *AIP Conference Proceedings*, 1443:14–21, 2012. arXiv:[1201.1493](https://arxiv.org/abs/1201.1493), doi:[10.1063/1.3703615](https://doi.org/10.1063/1.3703615).
- [FCC10] F. F. Fanchini, L. K. Castelano, and A. O. Caldeira. Entanglement versus quantum discord in two coupled double quantum dots. *New Journal of Physics*, 12(7):073009, 2010. arXiv:[0912.1468](https://arxiv.org/abs/0912.1468), doi:[10.1088/1367-2630/12/7/073009](https://doi.org/10.1088/1367-2630/12/7/073009).
- [FGLE12] S. T. Flammia, D. Gross, Y.-K. Liu, and J. Eisert. Quantum tomography via compressed sensing: error bounds, sample complexity and efficient estimators. *New Journal of Physics*, 14(9):095022, 2012. arXiv:[1205.2300](https://arxiv.org/abs/1205.2300), doi:[10.1088/1367-2630/14/9/095022](https://doi.org/10.1088/1367-2630/14/9/095022).
- [FSW07] A. S. Fletcher, P. W. Shor, and M. Z. Win. Optimum quantum error recovery using semidefinite programming. *Phys. Rev. A*, 75:012338, 2007. arXiv:[quant-ph/0606035](https://arxiv.org/abs/quant-ph/0606035), doi:[10.1103/PhysRevA.75.012338](https://doi.org/10.1103/PhysRevA.75.012338).
- [FTP⁺07] A. Fedrizzi, H. Thomas, A. Poppe, T. Jennewein, and A. Zeilinger. A wavelength-tunable fiber-coupled source of narrowband entangled photons. *Opt. Express*, 15(23):15377–15386, 2007. arXiv:[0706.2877](https://arxiv.org/abs/0706.2877), doi:[10.1364/OE.15.015377](https://doi.org/10.1364/OE.15.015377).

- [Ful97] W. Fulton. *Young tableaux: with applications to representation theory and geometry*, volume 35. Cambridge University Press, 1997.
- [FvdG99] C.A. Fuchs and J. van de Graaf. Cryptographic distinguishability measures for quantum-mechanical states. *Information Theory, IEEE Transactions on*, 45(4):1216–1227, 1999. [arXiv:quant-ph/9712042](https://arxiv.org/abs/quant-ph/9712042), [doi:10.1109/18.761271](https://doi.org/10.1109/18.761271).
- [GAE07] D. Gross, K. Audenaert, and J. Eisert. Evenly distributed unitaries: On the structure of unitary designs. *Journal of Mathematical Physics*, 48(5):052104, 2007. [arXiv:quant-ph/0611002](https://arxiv.org/abs/quant-ph/0611002), [doi:10.1063/1.2716992](https://doi.org/10.1063/1.2716992).
- [Gar11] B. M. Garraway. The dicke model in quantum optics: Dicke model revisited. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 369(1939):1137–1155, 2011. [doi:10.1098/rsta.2010.0333](https://doi.org/10.1098/rsta.2010.0333).
- [GB08] M. Grant and S. Boyd. Graph implementations for nonsmooth convex programs. In V. Blondel, S. Boyd, and H. Kimura, editors, *Recent Advances in Learning and Control*, Lecture Notes in Control and Information Sciences, pages 95–110. Springer-Verlag Limited, 2008. http://stanford.edu/~boyd/graph_dcp.html.
- [GB11] M. Gessner and H.-P. Breuer. Detecting Nonclassical System-Environment Correlations by Local Operations. *Phys. Rev. Lett.*, 107(18):180402, 2011. [arXiv:1108.2699](https://arxiv.org/abs/1108.2699), [doi:10.1103/PhysRevLett.107.180402](https://doi.org/10.1103/PhysRevLett.107.180402).
- [GB14] M. Grant and S. Boyd. CVX: Matlab software for disciplined convex programming, version 2.1. <http://cvxr.com/cvx>, March 2014.
- [GCM⁺12] J. M. Gambetta, A. D. Córcoles, S. T. Merkel, B. R. Johnson, J. A. Smolin, J. M. Chow, C. A. Ryan, C. Rigetti, S. Poletto, T. A. Ohki, M. B. Ketchen, and M. Steffen. Characterization of addressability by simultaneous randomized benchmarking. *Phys. Rev. Lett.*, 109:240504, 2012. [arXiv:1204.6308](https://arxiv.org/abs/1204.6308), [doi:10.1103/PhysRevLett.109.240504](https://doi.org/10.1103/PhysRevLett.109.240504).
- [GEB99] R. Ghose, T. R. Eykyn, and G. Bodenhausen. Average liouvillian theory revisited: cross-correlated relaxation between chemical shift anisotropy and dipolar couplings in the rotating frame in nuclear magnetic resonance. *Molecular Physics*, 96(8):1281–1288, 1999. [doi:10.1080/00268979909483072](https://doi.org/10.1080/00268979909483072).

- [GFC15] C. Granade, C. Ferrie, and D. G. Cory. Accelerated randomized benchmarking. *New Journal of Physics*, 17(1):013042, 2015. [arXiv:1404.5275](#), [doi:10.1088/1367-2630/17/1/013042](#).
- [GFWC12] C. E. Granade, C. Ferrie, N. Wiebe, and D. G. Cory. Robust on-line hamiltonian learning. *New Journal of Physics*, 14(10):103013, 2012. [arXiv:1207.1655](#), [doi:10.1088/1367-2630/14/10/103013](#).
- [GGZ11] F. Galve, G. L. Giorgi, and R. Zambrini. Orthogonal measurements are almost sufficient for quantum discord of two qubits. *EPL*, 96:40005, 2011. [arXiv:1107.2005](#), [doi:10.1209/0295-5075/96/40005](#).
- [Gho00] R. Ghose. Average liouvillian theory in nuclear magnetic resonance—principles, properties, and applications. *Concepts in Magnetic Resonance*, 12(3):152–172, 2000. [doi:10.1002/\(SICI\)1099-0534\(2000\)12:3<152::AID-CMR4>3.0.CO;2-P](#).
- [GLN05] A. Gilchrist, N. K. Langford, and M. A. Nielsen. Distance measures to compare real and ideal quantum processes. *Phys. Rev. A*, 71:062310, 2005. [arXiv:quant-ph/0408063](#), [doi:10.1103/PhysRevA.71.062310](#).
- [GLP⁺13] K. Geerlings, Z. Leghtas, I.M. Pop, S. Shankar, L. Frunzio, R.J. Schoelkopf, M. Mirrahimi, and M.H. Devoret. Demonstrating a driven reset protocol for a superconducting qubit. *Phys. Rev. Lett.*, 110:120501, 2013. [arXiv:1211.0491](#), [doi:10.1103/PhysRevLett.110.120501](#).
- [GPR⁺14] P. Gumann, O. Patange, C. Ramanathan, H. Haas, O. Moussa, M. L. W. Thewalt, H. Riemann, N. V. Abrosimov, P. Becker, H-J. Pohl, K. M. Itoh, and D. G. Cory. Inductive measurement of optically hyperpolarized phosphorous donor nuclei in an isotopically enriched silicon-28 crystal. *Phys. Rev. Lett.*, 113:267604, 2014. [arXiv:1407.5352](#), [doi:10.1103/PhysRevLett.113.267604](#).
- [GRP⁺13] M. Gessner, M. Ramm, T. Pruttivarasin, A. Buchleitner, H-P. Breuer, and H. Häffner. Local detection of quantum correlations with a single trapped ion. *Nat. Phys.*, 10(2):105–109, 2013. [arXiv:1311.4489](#), [doi:10.1038/nphys2829](#).
- [GTW09] A. Gilchrist, D. R. Terno, and C. J. Wood. Vectorization of quantum operations and its use. *ArXiv e-prints*, arXiv:0911.2539 [quant-ph], 2009. [arXiv:0911.2539](#).

- [Hae76] U. Haeberlen. *High resolution NMR in Solids*. Academic, New York, 1976.
- [Har11] L. Hardy. Reformulating and reconstructing quantum theory. *ArXiv e-prints*, arXiv:1104.2066v3 [quant-ph], 2011. [arXiv:1104.2066v3](#).
- [Har12] S. Hartmann. Generalized Dicke states. *ArXiv e-prints*, arXiv:1201.1732 [quant-ph], 2012. [arXiv:1201.1732](#).
- [HH62] S.R. Hartmann and E.L. Hahn. Nuclear double resonance in the rotating frame. *Phys. Rev.*, 128:2042–2053, 1962. [doi:10.1103/PhysRev.128.2042](#).
- [HH12] F. Huszár and N. M. T. Houlby. Adaptive bayesian quantum tomography. *Phys. Rev. A*, 85:052120, 2012. [arXiv:1107.0895](#), [doi:10.1103/PhysRevA.85.052120](#).
- [HHH99] M. Horodecki, P. Horodecki, and R. Horodecki. General teleportation channel, singlet fraction, and quasidistillation. *Phys. Rev. A*, 60:1888, 1999. [arXiv:quant-ph/9807091](#), [doi:10.1103/PhysRevA.60.1888](#).
- [HHHH09] R. Horodecki, P. Horodecki, M. Horodecki, and K. Horodecki. Quantum entanglement. *Rev. Mod. Phys.*, 81:865, 2009. [arXiv:quant-ph/0702225](#), [doi:10.1103/RevModPhys.81.865](#).
- [HJ85] R. A. Horn and C. R. Johnson. *Matrix Analysis*. Cambridge University Press, Cambridge, 1985.
- [HLB⁺03] Y. Hasegawa, R. Loidl, G. Badurek, M. Baron, and H. Rauch. Violation of a bell-like inequality in single-neutron interferometry. *Nature*, 425:45–48, 2003. [doi:10.1038/nature01881](#).
- [HLB⁺07] Y. Hasegawa, R. Loidl, G. Badurek, S. Filipp, J. Klepp, and H. Rauch. Evidence for entanglement and full tomographic analysis of bell states in a single-neutron system. *Phys. Rev. A*, 76:052108, 2007. [doi:10.1103/PhysRevA.76.052108](#).
- [Hra97] Z. Hradil. Quantum-state estimation. *Phys. Rev. A*, 55(3):R1561–R1564, 1997. [doi:10.1103/PhysRevA.55.R1561](#).
- [HV01] L. Henderson and V. Vedral. Classical, quantum and total correlations. *J. Phys. A: Math. Gen.*, 34:6899, 2001. [arXiv:quant-ph/0105028](#), [doi:10.1088/0305-4470/34/35/315](#).

- [HW68] U. Haeberlen and J. S. Waugh. Coherent averaging effects in magnetic resonance. *Phys. Rev.*, 175:453–467, 1968. doi:[10.1103/PhysRev.175.453](https://doi.org/10.1103/PhysRev.175.453).
- [Jam72] A. Jamiołkowski. Linear transformations which preserve trace and positive semidefiniteness of operators. *Rep. Math. Phys.*, 3:275, 1972. doi:[10.1016/0034-4877\(72\)90011-0](https://doi.org/10.1016/0034-4877(72)90011-0).
- [JK11] N. Johnston and D. W. Kribs. Quantum gate fidelity in terms of Choi matrices. *J. Phys. A: Math. Theor.*, 44:495303, 2011. arXiv:[1102.0948](https://arxiv.org/abs/1102.0948), doi:[10.1088/1751-8113/44/49/495303](https://doi.org/10.1088/1751-8113/44/49/495303).
- [JKMW01] Daniel F. V. James, Paul G. Kwiat, William J. Munro, and Andrew G. White. Measurement of qubits. *Phys. Rev. A*, 64(5):052312, 2001. arXiv:[quant-ph/0103121](https://arxiv.org/abs/quant-ph/0103121), doi:[10.1103/PhysRevA.64.052312](https://doi.org/10.1103/PhysRevA.64.052312).
- [JKP09] N. Johnston, D. W. Kribs, and V. I. Paulsen. Computing stabilized norms for quantum operations via the theory of completely bounded maps. *Quant. Info. Comp.*, 9(1):16–35, 2009. arXiv:[0711.3636](https://arxiv.org/abs/0711.3636).
- [JS61] T. F. Jordan and E. C. G. Sudarshan. Dynamical mappings of density operators in quantum mechanics. *J. Math. Phys.*, 2:772, 1961. doi:[10.1063/1.1724221](https://doi.org/10.1063/1.1724221).
- [JS91] R. Joyal and A. Street. The geometry of tensor calculus I. *Adv. Math.*, 88:55, 1991.
- [JSS04] T. F. Jordan, A. Shaji, and E. C. G. Sudarshan. Dynamics of initially entangled open quantum systems. *Phys. Rev. A*, 70:052110, Nov 2004. URL: <http://link.aps.org/doi/10.1103/PhysRevA.70.052110>, doi:[10.1103/PhysRevA.70.052110](https://doi.org/10.1103/PhysRevA.70.052110).
- [KdSR⁺14] S. Kimmel, M. P. da Silva, C. A. Ryan, B. R. Johnson, and T. Ohki. Robust extraction of tomographic information via randomized benchmarking. *Phys. Rev. X*, 4:011050, 2014. arXiv:[1306.2348](https://arxiv.org/abs/1306.2348), doi:[10.1103/PhysRevX.4.011050](https://doi.org/10.1103/PhysRevX.4.011050).
- [KGD⁺11] Y. Kubo, C. Grezes, A. Dewes, T. Umeda, J. Isoya, H. Sumiya, N. Morishita, H. Abe, S. Onoda, T. Ohshima, V. Jacques, A. Dréau, J.-F. Roch, I. Diniz, A. Auffeves, D. Vion, D. Esteve, and P. Bertet. Hybrid quantum circuit with a superconducting qubit coupled to a spin

- ensemble. *Phys. Rev. Lett.*, 107:220501, Nov 2011. [arXiv:1110.2978](#), [doi:10.1103/PhysRevLett.107.220501](#).
- [Kit97] A. Y. Kitaev. Quantum computations: algorithms and error correction. *Russian Mathematical Surveys*, 52(6):1191–1249, 1997.
- [KL80] G. M. Kelly and M. L. Laplaza. Coherence for compact closed categories. *J. Pure Appl. Algebra*, 19:193, 1980.
- [KL98] E. Knill and R. Laflamme. Power of one bit of quantum information. *Phys. Rev. Lett.*, 81:5672, 1998. [arXiv:quant-ph/9802037](#), [doi:10.1103/PhysRevLett.81.5672](#).
- [KLM01] E. Knill, R. Laflamme, and G. J. Milburn. A scheme for efficient quantum computation with linear optics. *Nature*, 409(6816):46–52, 2001. [doi:10.1038/35051009](#).
- [KMRRS07] A.-M. Kuah, K. Modi, C. A. Rodríguez-Rosario, and E. C. G. Sudarshan. How state preparation can affect a quantum experiment: Quantum process tomography for open systems. *Phys. Rev. A*, 76(4):042113, 2007. [arXiv:0706.0394](#), [doi:10.1103/PhysRevA.76.042113](#).
- [KMW02] David Kielpinski, Chris Monroe, and David J Wineland. Architecture for a large-scale ion-trap quantum computer. *Nature*, 417(6890):709–711, 2002. [doi:10.1038/nature00784](#).
- [KOB⁺10] Y. Kubo, F. R. Ong, P. Bertet, D. Vion, V. Jacques, D. Zheng, A. Dreau, J. F. Roch, A. Auffeves, F. Jelezko, J. Wrachtrup, M. F. Barthe, P. Bergonzo, and D. Esteve. Strong coupling of a spin ensemble to a superconducting resonator. *Phys. Rev. Lett.*, 105:140502, 2010. [arXiv:1006.0251](#), [doi:10.1103/PhysRevLett.105.140502](#).
- [Kra83] K. Kraus. *States, Effects and Operations: Fundamental Notions of Quantum Theory*, volume 190 of *Lecture Notes in Physics*. Springer, Berlin, 1983. [doi:10.1007/3-540-12732-1](#).
- [KS14] R. Koenig and J. A. Smolin. How to efficiently select an arbitrary clifford group element. *Journal of Mathematical Physics*, 55(12):122202, 2014. [arXiv:1406.2170](#), [doi:http://dx.doi.org/10.1063/1.4903507](#).

- [Kub62] R. Kubo. Generalized cumulant expansion method. *Journal of the Physical Society of Japan*, 17(7):1100–1120, 1962. doi:[10.1143/JPSJ.17.1100](https://doi.org/10.1143/JPSJ.17.1100).
- [KWM11] Z. Kurucz, J. H. Wesenberg, and K. Mølmer. Spectroscopic properties of inhomogeneously broadened spin ensembles in a cavity. *Phys. Rev. A*, 83:053852, May 2011. arXiv:[1101.4828](https://arxiv.org/abs/1101.4828), doi:[10.1103/PhysRevA.83.053852](https://doi.org/10.1103/PhysRevA.83.053852).
- [LCS11] M. D. Lang, C. M. Caves, and A. Shaji. Entropic measures of non-classical correlations. *Int. J. Quantum Inf.*, 09:1553, 2011. arXiv:[1105.4920](https://arxiv.org/abs/1105.4920), doi:[10.1142/S021974991100826X](https://doi.org/10.1142/S021974991100826X).
- [LD98] Daniel Loss and David P. DiVincenzo. Quantum computation with quantum dots. *Phys. Rev. A*, 57:120–126, 1998. doi:[10.1103/PhysRevA.57.120](https://doi.org/10.1103/PhysRevA.57.120).
- [LDB92] M. H. Levitt and L. Di Bari. Steady state in magnetic resonance pulse experiments. *Phys. Rev. Lett.*, 69:3124–3127, Nov 1992. doi:[10.1103/PhysRevLett.69.3124](https://doi.org/10.1103/PhysRevLett.69.3124).
- [Leu03] D. W. Leung. Choi’s proof as a recipe for quantum process tomography. *Journal of Mathematical Physics*, 44(2):528–533, 2003. arXiv:[quant-ph/0201119](https://arxiv.org/abs/quant-ph/0201119), doi:[10.1063/1.1518554](https://doi.org/10.1063/1.1518554).
- [LGR+13] Y. Lin, J. P. Gaebler, F. Reiter, T. R. Tan, R. Bowler, A. S. Sørensen, D. Leibfried, and D. J. Wineland. Dissipative production of a maximally entangled steady state of two quantum bits. *Nature*, 504(7480):415–418, 2013. arXiv:[1307.4443](https://arxiv.org/abs/1307.4443), doi:[10.1038/nature12801](https://doi.org/10.1038/nature12801).
- [LPB10] E.-M. Laine, J. Piilo, and H.-P. Breuer. Witness for initial system-environment correlations in open-system dynamics. *Europhys. Lett.*, 92(6):60010, 2010. arXiv:[1004.2184](https://arxiv.org/abs/1004.2184), doi:[10.1209/0295-5075/92/60010](https://doi.org/10.1209/0295-5075/92/60010).
- [LTLG11] C.-F. Li, J.-S. Tang, Y.-L. Li, and G.-C. Guo. Experimentally witnessing the initial correlation between an open quantum system and its environment. *Phys. Rev. A*, 83(6):064102, 2011. doi:[10.1103/PhysRevA.83.064102](https://doi.org/10.1103/PhysRevA.83.064102).
- [LWP+05] N. Langford, T. Weinhold, R. Prevedel, K. Resch, A. Gilchrist, J. O’Brien, G. Pryde, and A. White. Demonstration of a Simple Entangling Optical

- Gate and Its Use in Bell-State Analysis. *Phys. Rev. Lett.*, 95(21):210504, 2005. [arXiv:quant-ph/0506262](#), [doi:10.1103/PhysRevLett.95.210504](#).
- [Mag54] W. Magnus. On the exponential solution of differential equations for a linear operator. *Comm. Pure Appl. Math.*, 7:649–673, 1954. [doi:10.1002/cpa.3160070404](#).
- [MBC⁺12] K. Modi, A. Brodutch, H. Cable, T. Paterek, and V. Vedral. The classical-quantum boundary for correlations: Discord and related measures. *Rev. Mod. Phys.*, 84:1655, 2012. [arXiv:1112.6238](#), [doi:10.1103/RevModPhys.84.1655](#).
- [MBKE11] E. Magesan, R. Blume-Kohout, and J. Emerson. Gate fidelity fluctuations and quantum process invariants. *Phys. Rev. A*, 84:012309, 2011. [arXiv:0910.1315](#), [doi:10.1103/PhysRevA.84.012309](#).
- [MGE11] E. Magesan, J. M. Gambetta, and J. Emerson. Scalable and robust randomized benchmarking of quantum processes. *Phys. Rev. Lett.*, 106:180504, 2011. [arXiv:1009.3639](#), [doi:10.1103/PhysRevLett.106.180504](#).
- [MGE12] E. Magesan, J. M. Gambetta, and J. Emerson. Characterizing quantum gates via randomized benchmarking. *Phys. Rev. A*, 85:042311, 2012. [arXiv:1109.6887](#), [doi:10.1103/PhysRevA.85.042311](#).
- [MGJ⁺12] E. Magesan, J. M. Gambetta, B. R. Johnson, C. A. Ryan, J. M. Chow, S. T. Merkel, M. P. da Silva, G. A. Keefe, M. B. Rothwell, T. A. Ohki, M. B. Ketchen, and M. Steffen. Efficient measurement of quantum gate error by interleaved randomized benchmarking. *Phys. Rev. Lett.*, 109:080505, 2012. [arXiv:1203.4550](#), [doi:10.1103/PhysRevLett.109.080505](#).
- [MGS⁺13] S. T. Merkel, J. M. Gambetta, J. A. Smolin, S. Poletto, A. D. Córcoles, B. R. Johnson, C. A. Ryan, and M. Steffen. Self-consistent quantum process tomography. *Phys. Rev. A*, 87:062119, 2013. [arXiv:1211.0322](#), [doi:10.1103/PhysRevA.87.062119](#).
- [MMK⁺95] C. Monroe, D.M. Meekhof, B.E. King, S.R. Jefferts, W.M. Itano, D.J. Wineland, and P. Gould. Resolved-sideband raman cooling of a bound atom to the 3d zero-point energy. *Phys. Rev. Lett.*, 75:4011 – 4014, 1995. [doi:10.1103/PhysRevLett.75.4011](#).

- [Mod11] K. Modi. Preparation of states in open quantum mechanics. *Open Systems & Information Dynamics*, 18(03):253–260, 2011. [arXiv:0903.2027](#), [doi:10.1142/S1230161211000170](#).
- [Mod12] K. Modi. Operational approach to open dynamics and quantifying initial correlations. *Sci. Rep.*, 2:581, 2012. [arXiv:1011.6138](#), [doi:10.1038/srep00581](#).
- [Mol69] B.R. Mollow. Power spectrum of light scattered by two-level systems. *Phys. Rev.*, 188:1969 – 1975, 1969. [doi:10.1103/PhysRev.188.1969](#).
- [Mos09] Michele Mosca. Quantum algorithms. In Robert A. Meyers, editor, *Encyclopedia of Complexity and Systems Science*, pages 7088–7118. Springer New York, 2009. [arXiv:0808.0369](#), [doi:10.1007/978-0-387-30440-3_423](#).
- [MRD⁺13] D. H. Mahler, L. A. Rozema, A. Darabi, C. Ferrie, R. Blume-Kohout, and A. M. Steinberg. Adaptive quantum state tomography improves accuracy quadratically. *Phys. Rev. Lett.*, 111:183601, 2013. [arXiv:1303.0436](#), [doi:10.1103/PhysRevLett.111.183601](#).
- [MS10] Kavan Modi and E. C. G. Sudarshan. Role of preparation in quantum process tomography. *Phys. Rev. A*, 81(5):052119, 2010. [arXiv:0904.4663](#), [doi:10.1103/PhysRevA.81.052119](#).
- [MSH⁺08] JR Maze, PL Stanwix, JS Hodges, S Hong, JM Taylor, P Cappellaro, L Jiang, MV Gurudev Dutt, E Togan, AS Zibrov, et al. Nanoscale magnetic sensing with an individual electronic spin in diamond. *Nature*, 455(7213):644–647, 2008. [doi:10.1038/nature07279](#).
- [MSM⁺12] S. Miyashita, T. Shirai, T. Mori, H. De Raedt, S. Bertaina, and I. Chiorescu. Photon and spin dependence of the resonance line shape in the strong coupling regime. *Journal of Physics B: Atomic, Molecular and Optical Physics*, 45(12):124010, 2012. [arXiv:1206.2394](#), [doi:10.1088/0953-4075/45/12/124010](#).
- [MTB⁺08] J.J.L. Morton, A.M. Tyryshkin, R.M. Brown, S. Shankar, B.W. Lovett, A. Ardavan, T. Schenkel, E.E. Haller, J.W. Ager, and S.A. Lyon. Solid-state quantum memory using the ³¹p nuclear spin. *Nature*, 455:1085–1088, 2008. [doi:10.1038/nature07295](#).

- [MVZ⁺12] K.W. Murch, U. Vool, D. Zhou, S.J. Weber, S.M. Girvin, and I. Siddiqi. Cavity-assisted quantum bath engineering. *Phys. Rev. Lett.*, 109:183602, 2012. [arXiv:1207.0053](https://arxiv.org/abs/1207.0053), [doi:http://dx.doi.org/10.1103/PhysRevLett.109.183602](https://doi.org/10.1103/PhysRevLett.109.183602).
- [Nar07] G. Narang. Simulating a single-qubit channel using a mixed-state environment. *Phys. Rev. A*, 75(3):032305, March 2007. [arXiv:quant-ph/0611058](https://arxiv.org/abs/quant-ph/0611058), [doi:10.1103/PhysRevA.75.032305](https://doi.org/10.1103/PhysRevA.75.032305).
- [NC00] M. A. Nielsen and I. L. Chuang. *Quantum Computation and Quantum Information*. Cambridge University Press, 2000.
- [Nie02] M. A. Nielsen. A simple formula for the average gate fidelity of a quantum dynamical operation. *Phys. Lett. A*, 303:249, 2002. [arXiv:quant-ph/0205035](https://arxiv.org/abs/quant-ph/0205035), [doi:10.1016/S0375-9601\(02\)01272-0](https://doi.org/10.1016/S0375-9601(02)01272-0).
- [NIS] NIST. Quantum algorithm zoo. URL: <http://math.nist.gov/quantum/zoo/>.
- [NIS15] NIST. Neutron interferometry and optics facility, 2015. URL: <http://physics.nist.gov/MajResFac/InterFer/text.html>.
- [OZ02] H. Ollivier and W. H. Zurek. Quantum discord: A measure of the quantumness of correlations. *Phys. Rev. Lett.*, 88:017901, 2002. [doi:10.1103/PhysRevLett.88.017901](https://doi.org/10.1103/PhysRevLett.88.017901).
- [PAC09] D. A. Pushin, M. Arif, and D. G. Cory. Decoherence-free neutron interferometry. *Physical Review A*, 79:053635, 2009. [doi:10.1103/PhysRevA.79.053635](https://doi.org/10.1103/PhysRevA.79.053635).
- [PAHC08] D. A. Pushin, M. Arif, M. Huber, and D. G. Cory. Measurements of the vertical coherence length in neutron interferometry. *Physical Review Letters*, 100:250404, 2008. [doi:10.1103/PhysRevLett.100.250404](https://doi.org/10.1103/PhysRevLett.100.250404).
- [Pau02] Vern Paulsen. *Completely bounded maps and operator algebras*, volume 78. Cambridge University Press, 2002.
- [PB02] F. Petruccione and H.-P. Breuer. *The theory of open quantum systems*. Oxford Univ. Press, 2002.

- [PCA⁺07] D. A. Pushin, D. G. Cory, M. Arif, D. L. Jacobson, and M. G. Huber. Reciprocal space approaches to neutron imaging. *Applied Physics Letters*, 90:224104, 2007. doi:[10.1063/1.2737390](https://doi.org/10.1063/1.2737390).
- [Pec94] P. Pechukas. Reduced dynamics need not be completely positive. *Phys. Rev. Lett.*, 73:1060–1062, 1994. doi:[10.1103/PhysRevLett.73.1060](https://doi.org/10.1103/PhysRevLett.73.1060).
- [Pen56] R. Penrose. On best approximate solutions of linear matrix equations. *Mathematical Proceedings of the Cambridge Philosophical Society*, 52:17–19, 1956. doi:[10.1017/S0305004100030929](https://doi.org/10.1017/S0305004100030929).
- [Pen71] R. Penrose. Applications of negative dimensional tensors. In *Combinatorial Mathematics and its Applications*. Academic Press, 1971.
- [PHA⁺15] D. A. Pushin, M. G. Huber, M. Arif, C. B. Shahi, J. Nsofini, C. J. Wood, D. Sarenac, and D. G. Cory. Neutron interferometry at the national institute of standards and technology. *Advances in High Energy Physics*, 2015:687480, 2015. doi:[10.1155/2015/687480](https://doi.org/10.1155/2015/687480).
- [PHAC11] D. A. Pushin, M. G. Huber, M. Arif, and D. G. Cory. Experimental realization of decoherence-free subspace in neutron interferometry. *Phys. Rev. Lett.*, 107:150401, 2011. doi:[10.1103/PhysRevLett.107.150401](https://doi.org/10.1103/PhysRevLett.107.150401).
- [PHBK99] M. Plenio, S. Huelga, A. Beige, and P. Knight. Cavity-loss-induced generation of entangled atoms. *Phys. Rev. A*, 59(3):2468–2475, March 1999. [arXiv:quant-ph/9811003](https://arxiv.org/abs/quant-ph/9811003), doi:[10.1103/PhysRevA.59.2468](https://doi.org/10.1103/PhysRevA.59.2468).
- [Pur46] E. M. Purcell. Spontaneous emission probabilities at radio frequencies. In *Proceedings of the American Physical Society*, 1946.
- [Pyn05] Roger Pynn. Broadband spin flippers constructed from thin magnetic films. *Physica B: Condensed Matter*, 356:178–181, 2005. doi:[10.1016/j.physb.2004.10.072](https://doi.org/10.1016/j.physb.2004.10.072).
- [Ram08] C. Ramanathan. Dynamic nuclear polarization and spin diffusion in non-conducting solids. *Appl. Magn. Reson.*, 34:409, 2008. [arXiv:0801.2170](https://arxiv.org/abs/0801.2170), doi:[10.1007/s00723-008-0123-7](https://doi.org/10.1007/s00723-008-0123-7).
- [RHP10] Ángel Rivas, Susana F. Huelga, and Martin B. Plenio. Entanglement and non-markovianity of quantum evolutions. *Phys. Rev. Lett.*, 105:050403, 2010. [arXiv:0911.4270](https://arxiv.org/abs/0911.4270), doi:[10.1103/PhysRevLett.105.050403](https://doi.org/10.1103/PhysRevLett.105.050403).

- [RHP14] Ángel Rivas, Susana F Huelga, and Martin B Plenio. Quantum non-markovianity: characterization, quantification and detection. *Reports on Progress in Physics*, 77(9):094001, 2014. [arXiv:1405.0303](#), [doi:10.1088/0034-4885/77/9/094001](#).
- [RMBL08] C.A. Ryan, O. Moussa, J. Baugh, and R. Laflamme. Spin based heat engine: Demonstration of multiple rounds of algorithmic cooling. *Phys. Rev. Lett.*, 100:140501, 2008. [arXiv:0706.2853](#), [doi:10.1103/PhysRevLett.100.140501](#).
- [RRMAG10] César A. Rodríguez-Rosario, Kavan Modi, and Alán Aspuru-Guzik. Linear assignment maps for correlated system-environment states. *Phys. Rev. A*, 81:012313, 2010. URL: <http://link.aps.org/doi/10.1103/PhysRevA.81.012313>, [arXiv:0910.5568](#), [doi:10.1103/PhysRevA.81.012313](#).
- [RRMK⁺08] C. A. Rodríguez-Rosario, K. Modi, A.-M. Kuah, A. Shaji, and E. C. G. Sudarshan. Completely positive maps and classical correlations. *J. Phys. A–Math. Theor.*, 41(20):205301, 2008. [arXiv:quant-ph/0703022](#), [doi:10.1088/1751-8113/41/20/205301](#).
- [RRMMAG12] C. A. Rodríguez-Rosario, K. Modi, L. Mazzola, and A. Aspuru-Guzik. Unification of witnessing initial system-environment correlations and witnessing non-Markovianity. *Europhys. Lett.*, 99(2):20010, 2012. [arXiv:1204.2197](#), [doi:10.1209/0295-5075/99/20010](#).
- [RS09] A. Roy and A.J. Scott. Unitary designs and codes. *Designs, Codes and Cryptography*, 53:13, 2009. [arXiv:0809.3813](#), [doi:10.1007/s10623-009-9290-2](#).
- [RW05] B. Rosgen and J. Watrous. On the Hardness of Distinguishing Mixed-State Quantum Computations. In *20th Annual IEEE Conference on Computational Complexity (CCC'05)*, pages 344–354. IEEE, 2005. [arXiv:cs/0407056](#), [doi:10.1109/CCC.2005.21](#).
- [RWC⁺11] D Z. Rossatto, T. Werlang, L K. Castelano, C J. Villas-Boas, and F F. Fanchini. Purity as a witness for initial system-environment correlations in open-system dynamics. *Phys. Rev. A*, 84(4):042113, 2011. [arXiv:1108.2370](#), [doi:10.1103/PhysRevA.84.042113](#).

- [RWM⁺15] M. Ringbauer, C. J. Wood, K. Modi, A. Gilchrist, A. G. White, and A. Fedrizzi. Characterizing quantum dynamics with initial system-environment correlations. *Phys. Rev. Lett.*, 114:090402, 2015. [arXiv:1410.5826](#), [doi:10.1103/PhysRevLett.114.090402](#).
- [RZB⁺75] H. Rauch, A. Zeilinger, G. Badurek, A. Wilfing, W. Bauspiess, and U. Bonse. Verification of coherent spinor rotation of fermions. *Physics Letters A*, 54, 1975. [doi:10.1016/0375-9601\(75\)90798-7](#).
- [ŠB01] P Štelmachovič and V Bužek. Dynamics of open quantum systems initially entangled with environment: Beyond the kraus representation. *Phys. Rev. A*, 64:062106, 2001. [arXiv:quant-ph/0108136](#), [doi:10.1103/PhysRevA.64.062106](#).
- [SBC⁺11] A. Smirne, D. Brivio, S. Cialdi, B. Vacchini, and M. G. A. Paris. Experimental investigation of initial system-environment correlations via trace-distance evolution. *Phys. Rev. A*, 84(3):032112, 2011. [arXiv:1105.0174](#), [doi:10.1103/PhysRevA.84.032112](#).
- [SBR⁺83] J. Summhammer, G. Badurek, H. Rauch, U. Kischko, and A. Zeilinger. Direct observation of fermion spin superposition by neutron interferometry. *Phys. Rev. A*, 27, 1983. [doi:10.1103/PhysRevA.27.2523](#).
- [Sch96] B. Schumacher. Sending entanglement through noisy quantum channels. *Phys. Rev. A*, 54:2614, 1996. [arXiv:quant-ph/9604023](#), [doi:10.1103/PhysRevA.54.2614](#).
- [SD82] Marlan O. Scully and Kai Drühl. Quantum eraser: A proposed photon correlation experiment concerning observation and "delayed choice" in quantum mechanics. *Phys. Rev. A*, 25:2208–2213, 1982. [doi:10.1103/PhysRevA.25.2208](#).
- [Sea89] V. F. Sears. *Neutron Optics*. Oxford University Press, 1989.
- [Sel07] P. Selinger. Dagger compact closed categories and completely positive maps: (extended abstract). *Electron. Notes Theor. Comput. Sci.*, 170:139, 2007. [doi:10.1016/j.entcs.2006.12.018](#).
- [Sel11a] P. Selinger. Finite dimensional hilbert spaces are complete for dagger compact closed categories (extended abstract). *Electron. Notes Theor. Comput. Sci.*, 270:113, 2011. [doi:10.1016/j.entcs.2011.01.010](#).

- [Sel11b] P. Selinger. A survey of graphical languages for monoidal categories. In *New Structures for Physics*, volume 813, pages 289–355. Springer Berlin, 2011. [arXiv:0908.3347](#), [doi:10.1007/978-3-642-12821-9_4](#).
- [SKM⁺11] A. Shabani, R. L. Kosut, M. Mohseni, H. Rabitz, M. A. Broome, M. P. Almeida, A. Fedrizzi, and A. G. White. Efficient Measurement of Quantum Dynamics via Compressive Sensing. *Phys. Rev. Lett.*, 106(10):100401, 2011. [arXiv:0910.5498](#), [doi:10.1103/PhysRevLett.106.100401](#).
- [SL09] A. Shabani and D. A. Lidar. Maps for general open quantum systems and a theory of linear quantum error correction. *Phys. Rev. A*, 80:012309, 2009. [arXiv:0902.2478](#), [doi:10.1103/PhysRevA.80.012309](#).
- [SMW05] L. J. Schulman, T. Mor, and Y. Weinstein. Physical limits of heat-bath algorithmic cooling. *Phys. Rev. Lett.*, 94:120501, 2005. [doi:10.1103/PhysRevLett.94.120501](#).
- [SMWF⁺07] T. Schmitt-Manderbach, H. Weier, M. Fürst, R. Ursin, F. Tiefenbacher, T. Scheidl, J. Perdigues, Z. Sodnik, C. Kurtsiefer, J. G. Rarity, A. Zeilinger, and H. Weinfurter. Experimental demonstration of free-space decoy-state quantum key distribution over 144 km. *Phys. Rev. Lett.*, 98:010504, 2007. [doi:10.1103/PhysRevLett.98.010504](#).
- [SS05] A. Shaji and E.C.G. Sudarshan. Who’s afraid of not completely positive maps? *Physics Letters A*, 341(1–4):48 – 54, 2005. [doi:10.1016/j.physleta.2005.04.029](#).
- [SSGF04] D. Salgado, J. L. Sánchez-Gómez, and M. Ferrero. Evolution of any finite open quantum system always admits a kraus-type representation, although it is not always completely positive. *Phys. Rev. A*, 70:054102, 2004. URL: <http://link.aps.org/doi/10.1103/PhysRevA.70.054102>, [doi:10.1103/PhysRevA.70.054102](#).
- [SST⁺12] M. Steger, K. Saeedi, M. L. W. Thewalt, J. J. L. Morton, H. Riemann, N. V. Abrosimov, P. Becker, and H.-J. Pohl. Quantum information storage for over 180 s using donor spins in a 28si “semiconductor vacuum”. *Science*, 336(6086):1280–1283, 2012. [doi:10.1126/science.1217635](#).
- [Sti55] W. F. Stinespring. Positive functions on C*-algebras. *Proc. Am. Math. Soc.*, 6:211, 1955.

- [SV99] L. J. Schulman and U. V. Vazirani. Molecular scale heat engines and scalable quantum computation. In *Proceedings of the thirty-first annual ACM symposium on Theory of computing*, pages 322–329. ACM, 1999. [doi:10.1145/301250.301332](https://doi.org/10.1145/301250.301332).
- [TC68] M. Tavis and F. W. Cummings. Exact solution for an n -molecule-radiation-field hamiltonian. *Phys. Rev.*, 170:379–384, 1968. [doi:10.1103/PhysRev.170.379](https://doi.org/10.1103/PhysRev.170.379).
- [TC69] M. Tavis and F. W. Cummings. Approximate solutions for an n -molecule-radiation-field hamiltonian. *Phys. Rev.*, 188:692–695, Dec 1969. [doi:10.1103/PhysRev.188.692](https://doi.org/10.1103/PhysRev.188.692).
- [TDL01] B. M. Terhal, D. P. DiVincenzo, and D. W. Leung. Hiding bits in bell states. *Phys. Rev. Lett.*, 86:5807–5810, 2001. [arXiv:quant-ph/0011042](https://arxiv.org/abs/quant-ph/0011042), [doi:10.1103/PhysRevLett.86.5807](https://doi.org/10.1103/PhysRevLett.86.5807).
- [Ter15] B.M. Terhal. Quantum error correction for quantum memories. *Rev. Mod. Phys.*, 87(2):307–346, 2015. [arXiv:1302.3428](https://arxiv.org/abs/1302.3428), [doi:10.1103/RevModPhys.87.307](https://doi.org/10.1103/RevModPhys.87.307).
- [TW05] V. V. Temnov and U. Woggon. Superradiance and subradiance in an inhomogeneously broadened ensemble of two-level systems coupled to a low- q cavity. *Phys. Rev. Lett.*, 95:243602, 2005. [doi:10.1103/PhysRevLett.95.243602](https://doi.org/10.1103/PhysRevLett.95.243602).
- [VB96] L. Vandenberghe and S. Boyd. Semidefinite programming. *SIAM Review*, 38(1):49–95, 1996. [doi:10.1137/1038003](https://doi.org/10.1137/1038003).
- [VC00] V. Vuletic and S. Chu. Laser cooling of atoms, ions, or molecules by coherent scattering. *Phys. Rev. Lett.*, 84:3787 – 3790, 2000. [doi:10.1103/PhysRevLett.84.3787](https://doi.org/10.1103/PhysRevLett.84.3787).
- [Vid08] G. Vidal. Class of quantum many-body states that can be efficiently simulated. *Phys. Rev. Lett.*, 101:110501, 2008. [arXiv:quant-ph/0610099](https://arxiv.org/abs/quant-ph/0610099), [doi:10.1103/PhysRevLett.101.110501](https://doi.org/10.1103/PhysRevLett.101.110501).
- [VM14] S. Vinjanampathy and K. Modi. Second law for quantum operations. *ArXiv e-prints*, arXiv:1405.6140 [quant-ph], 2014. [arXiv:1405.6140](https://arxiv.org/abs/1405.6140).

- [VMC08] F. Verstraete, V. Murg, and J. I. Cirac. Matrix product states, projected entangled pair states, and variational renormalization group methods for quantum spin systems. *Adv. Phys.*, 57:143, 2008. [arXiv:0907.2796](#), [doi:10.1080/14789940801912366](#).
- [VOB⁺06] S.O. Valenzuela, W.D. Oliver, D.M. Berns, K.K. Berggren, L.S. Levitov, and T.P. Orlando. Microwave-induced cooling of a superconducting qubit. *Science*, 314:1589, 2006. [arXiv:cond-mat/0702190](#), [doi:10.1126/science.1134008](#).
- [VWC09] F. Verstraete, M. M. Wolf, and J. I. Cirac. Quantum computation and quantum-state engineering driven by dissipation. *Nat Phys*, 5(9):633–636, 09 2009. [doi:10.1038/nphys1342](#).
- [WAH⁺14] C. J. Wood, M. O. Abutaleb, M. G. Huber, M. Arif, D. G. Cory, and D. A. Pushin. Quantum correlations in a noisy neutron interferometer. *Phys. Rev. A*, 90:032315, 2014. [arXiv:1304.6935](#), [doi:10.1103/PhysRevA.90.032315](#).
- [Wat08] J. Watrous. Theory of quantum information. *Lecture notes, University of Waterloo*, 2008. URL: <https://cs.uwaterloo.ca/~watrous/CS766/>.
- [Wat09] J. Watrous. Semidefinite programs for completely bounded norms. *ArXiv e-prints*, arXiv:0901.4709 [quant-ph], 2009. [arXiv:0901.4709](#).
- [WBC14] C. J. Wood, T. W. Borneman, and D. G. Cory. Cavity cooling of an ensemble spin system. *Phys. Rev. Lett.*, 112:050501, 2014. [doi:10.1103/PhysRevLett.112.050501](#).
- [WBC15] C. J. Wood, J. D. Biamonte, and D. G. Cory. Tensor networks and graphical calculus for open quantum systems. *Quant. Inf. Comp.*, 15:0579–0811, 2015. [arXiv:1111.6950](#).
- [WC15] C. J. Wood and D. G. Cory. Cavity cooling to the ground state of an ensemble quantum system. *ArXiv e-prints*, arXiv:1506.03007 [quant-ph], 2015. [arXiv:1506.03007](#).
- [WECC08] M. M. Wolf, J. Eisert, T. S. Cubitt, and J. I. Cirac. Assessing non-markovian quantum dynamics. *Phys. Rev. Lett.*, 101:150402, 2008. [arXiv:0711.3172](#), [doi:10.1103/PhysRevLett.101.150402](#).

- [WF14] J. J. Wallman and S. T. Flammia. Randomized benchmarking with confidence. *New Journal of Physics*, 16(10):103032, 2014. [arXiv:1404.6025](#), [doi:10.1088/1367-2630/16/10/103032](#).
- [WGFC14] N. Wiebe, C. Granade, C. Ferrie, and D.G. Cory. Hamiltonian learning and certification using quantum resources. *Phys. Rev. Lett.*, 112:190501, 2014. [arXiv:1309.0876](#), [doi:10.1103/PhysRevLett.112.190501](#).
- [WHE⁺04] Y. S. Weinstein, T. F. Havel, J. Emerson, N. Boulant, M. Saraceno, S. Lloyd, and D. G. Cory. Quantum process tomography of the quantum fourier transform. *J. Chem. Phys.*, 121:6117, 2004. [doi:10.1063/1.1785151](#).
- [WHG15] C. J. Wood, I. N. Hincks, and C. E. Granade. Quantum utils for mathematics [online]. 2015. URL: <http://quantumutils.org>.
- [WI79] D.J. Wineland and W.M. Itano. Laser cooling of atoms. *Phys. Rev. A*, 20:1521 – 1540, 1979. [doi:10.1103/PhysRevA.20.1521](#).
- [Wic54] G. C. Wick. Properties of Bethe-Salpeter wave functions. *Phys. Rev.*, 96:1124, 1954. [doi:10.1103/PhysRev.96.1124](#).
- [WJ06] J. Wrachtrup and F. Jelezko. Processing quantum information in diamond. *Journal of Physics: Condensed Matter*, 18(21):S807, 2006. [doi:10.1088/0953-8984/18/21/S08](#).
- [WKV⁺10] J.R. Weber, W.F. Koehl, J.B. Varley, A. Janotti, B.B. Buckley, C.G. Van de Walle, and D.D. Awschalom. Quantum computing with defects. *Proc. Natl. Acad. Sci. U.S.A.*, 107:8513 – 8518, 2010. [doi:10.1073/pnas.1003052107](#).
- [WLB13] S. Wißmann, B. Leggio, and H.-P. Breuer. Detecting initial system-environment correlations: Performance of various distance measures for quantum states. *Phys. Rev. A*, 88:022108, 2013. [arXiv:1306.3248](#), [doi:10.1103/PhysRevA.88.022108](#).
- [WM02] J. Wesenberg and K. Mølmer. Mixed collective states of many spins. *Phys. Rev. A*, 65:062304, May 2002. [arXiv:quant-ph/0201043](#), [doi:10.1103/PhysRevA.65.062304](#).

- [Woo98] W. K. Wootters. Entanglement of formation of an arbitrary state of two qubits. *Phys. Rev. Lett.*, 80:2245, 1998. [arXiv:quant-ph/9709029](#), [doi:10.1103/PhysRevLett.80.2245](#).
- [Woo08] C. J. Wood. Non-completely positive maps: properties and applications. Master’s thesis, Macquarie University, 2008. [arXiv:0911.3199](#).
- [WRLZ08] M. Wallquist, P. Rabl, M.D. Lukin, and P. Zoller. Theory of cavity-assisted microwave cooling of polar molecules. *New J. Phys.*, 10:063005, 2008. [arXiv:0803.2666](#), [doi:10.1088/1367-2630/10/6/063005](#).
- [WSFVB09] T. Werlang, S. Souza, F. F. Fanchini, and C. J. Villas Boas. Robustness of quantum discord to sudden death. *Phys. Rev. A*, 80:024103, 2009. [arXiv:0905.3376](#), [doi:10.1103/PhysRevA.80.024103](#).
- [XTH13] M. Xu, D. A. Tieri, and M. J. Holland. Simulating open quantum systems by applying $su(4)$ to quantum master equations. *Phys. Rev. A*, 87:062101, 2013. [arXiv:1302.6284](#), [doi:10.1103/PhysRevA.87.062101](#).
- [YE09] T. Yu and J. H. Eberly. Sudden death of entanglement. *Science*, 323(5914):598–601, 2009. [arXiv:0910.1396](#), [doi:10.1126/science.1167343](#).
- [YI99] Y. Yamamoto and A. Imamoglu. *Mesoscopic Quantum Optics*, chapter 6.4. John Wiley & Sons, Inc., 1999.
- [ZCFJ12] D. Zhou, G-W. Chern, J. Fei, and R. Joynt. Topology of entanglement evolution of two qubits. *International Journal of Modern Physics B*, 26, 2012. [arXiv:1007.1749](#), [doi:10.1142/S0217979212500543](#).
- [Zim06] M. Ziman. Quantum process tomography: the role of initial correlations. *ArXiv e-prints*, arXiv:quant-ph/0603166, 2006. [arXiv:quant-ph/0603166](#).
- [ZZXG11] Ying-Jie Zhang, Xu-Bo Zou, Yun-Jie Xia, and Guang-Can Guo. Quantum discord dynamics in the presence of initial system–cavity correlations. *Journal of Physics B: Atomic, Molecular and Optical Physics*, 44(3):035503, 2011. [doi:10.1088/0953-4075/44/3/035503](#).