SCALABLE TECHNIQUES FOR QUANTUM NETWORK ENGINEERING

A DISSERTATION
SUBMITTED TO THE DEPARTMENT OF APPLIED PHYSICS
AND THE COMMITTEE ON GRADUATE STUDIES
OF STANFORD UNIVERSITY
IN PARTIAL FULFILLMENT OF THE REQUIREMENTS
FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY

Nikolas Tezak
August 2016

© 2016 by Nikolas Anton Tezak. All Rights Reserved.
Re-distributed by Stanford University under license with the author.

This work is licensed under a Creative Commons Attribution-Noncommercial 3.0 United States License.
http://creativecommons.org/licenses/by-nc/3.0/us/

This dissertation is online at: http://purl.stanford.edu/zh617nw5199

I certify that I have read this dissertation and that, in my opinion, it is fully adequate in scope and quality as a dissertation for the degree of Doctor of Philosophy.

Hideo Mabuchi, Primary Adviser

I certify that I have read this dissertation and that, in my opinion, it is fully adequate in scope and quality as a dissertation for the degree of Doctor of Philosophy.

Surya Ganguli

I certify that I have read this dissertation and that, in my opinion, it is fully adequate in scope and quality as a dissertation for the degree of Doctor of Philosophy.

Patrick Hayden

I certify that I have read this dissertation and that, in my opinion, it is fully adequate in scope and quality as a dissertation for the degree of Doctor of Philosophy.

Jelena Vuckovic

Approved for the Stanford University Committee on Graduate Studies.

Patricia J. Gumport, Vice Provost for Graduate Education

This signature page was generated electronically upon submission of this dissertation in electronic format. An original signed hard copy of the signature page is on file in University Archives.

Abstract

In the quest for creating "quantum enhanced" systems for information processing many currently pursued design strategies are difficult to scale significantly beyond a few dozen qubits. The dominant design paradigm relies on starting with near perfect quantum components and a vast overhead of classical external control. In my thesis I present tools and methods for a more integrated framework which treats quantum and hybrid quantum-classical systems on equal footing.

We have recently defined a Quantum Hardware Description Language (QHDL) capable of describing networks of interconnected open quantum systems. QHDL is compiled to symbolic and numerical system models by a custom software tool suite named QNET. This allows us to rapidly iterate over quantum network designs and derive the associated equations of motion.

Building on a recently developed model reduction technique for describing networks of nonlinear oscillators in the semi-classical regime, I present a library of nonlinear optical circuit designs useful for all-optical computation. I further present an end-to-end theoretical proposal to create all-optical neuromorphic circuits capable of supervised learning. The system is hierarchically composed of tunable linear amplifiers, analog phase memories and thresholding non-linear circuits which can be used to construct more general quantum feedback networks for nonlinear information processing.

Finally, I introduce a novel model transformation capable of dividing the description of quantum states into a low-dimensional quasi-classical part coupled to a lower complexity quantum state. This approach is exact and naturally tailored to simulating coupled quantum systems with varying degrees of dissipation.

Acknowledgements

There are a great number of people that helped me along both during my past five years of actively working towards the PhD and on the way to even getting there. First, I would like to thank my research adviser Hideo Mabuchi for giving me the opportunity to work in his group, for providing me with great impulses and ideas for research projects, for allowing me to just go play and explore with very little pressure and for keeping faith in my ability to grow into a fully fledged scientist (eventually).

Second, I would like to thank Mike Armen for always having great advice, for being an example of a true master of his craft and just for being Mike Armen.

While Joe Kerckhoff and I did not overlap for very long, it was him that first taught me the ins and outs of quantum feedback networks and quantum stochastic calculus and he has kept providing me with valuable research and career advice ever since.

I have immensely benefited from working with and learning from Dmitri Pavlichin, Gopal Sarma and Ryan Hamerly. Dmitri's expertise on all things information theoretic and his out of the box style of thinking have made him a great person to bounce ideas off of and learn from. Gopal's great enthusiasm for our research and a variety of other shared interests was contagious and he established our weekly theory meetings which provided us with some very valuable structure. Ryan – where should I begin – I am in awe! Co-evolving alongside of Ryan has shown me what growth and success is possible when great intellect meets persistence meets work ethic.

I would like to thank my group internal research collaborators, Armand Niederberger with whom I worked on QHDL and spoke German, Orion Crisafulli who first created the feedback coupled DOPO experiment, Daniel Soh who carried forward that experiment, brought it to fruition and is just generally an extremely impressive scientist, Nina Amini who has patiently worked with me on a number of projects, one of which is actually close to being completed. I have also greatly enjoyed working with Gil Tabak and Michael Goerz.

I want to thank the senior experimentalist student squad Hardeep Sanghera, Nate Bogdanowicz, Dodd Gray and Chris Rogers for being awesome guys and taking great sunset pictures. Peter McMahon has been a great recent addition to the group, his expertise and deep knowledge of the field has made him a real resource. He has also generously shared with us his amazing taste in wine. With Charles Limouse I have shared great enthusiasm for computational tools and I have greatly enjoyed our 'triathlons' (rock climbing, rowing, swimming). It has been a great pleasure witnessing Hiro Onodera and Edwin Ng starting their PhDs in our group and forming an incredible team. This alone is no easy feat and I know it will pay off. Finally, I want to thank Michael Celentano, Jeff Hill, Mike Zhang, Jie Wu, Yeon-Dae Kwon and Tony Miller for being great guys to work with and be around. I want to thank our group admin Suki Ungson for patiently walking us through all those reimbursement requests and supporting our various group activities. A giant thank you also goes out to Paula Perron and Claire Nicholas who are keeping the Department of Applied Physics up and running by their hard work and kind support.

Outside of Stanford, I would like to thank all my collaborators, Ray Beausoleil, Charles Santori, Jason Pelc, Dave Kielpinski, Thomas Van Vaerenbergh, Ranojoy Bose and Gabriel Mendoza (all at Hewlett Packard Labs) have been amazing people to work both with and for, and learn from. Rad Balu and Kurt Jacobs of the Army Research Lab have been extremely supportive and great to work for, as well. I want to thank Rad Balu especially for also being a great mentor. Through conferences I have had the great pleasure of interacting with the theory giants of our field Matt James, John Gough and Hendra Nurdin. The insight and advice they have shared with me over the years are highly appreciated.

I want to thank my thesis defense committee members and readers Surya Ganguli, Patrick Hayden and Jelena Vuckovic for invaluable discussions and interactions over the past few years as well as Tsachy Weissmann for agreeing to be the defense chair.

I want to thank Stanford University in general for providing a unique, rich and amazing environment. Sometimes I still wonder how I got here. My graduate research was supported by DARPA-MTO under award no. N66001-11-1-4106, a Stanford Graduate Fellowship and a Math+X fellowship by the Simons Foundation. I highly appreciated the academic freedom this gave me. This thesis came together over a very rapid time and I am therefore extremely grateful for proof reading on really short notice by Jenny, my dad, Paula and Michael Goerz! While at Stanford I have also had the great pleasure to serve the Stanford Optical Society in various roles, here I want to thank Marina, Cathy, Matt, Patrick, Stephen W, Alex, Andrew, Therice, Stephen H, Linda, Sage and Adam for a great time together and for putting together some awesome events and projects. I want to especially thank Marina Radulaski for being a great partner in crime in running Optical Society related things and for being a great friend.

I want to thank my family for all the love and support; my Mom and Dad for encouraging me to find my own path and having faith in me, my siblings Elena and Micha for being crazy talented and making me proud to be their brother but most of all for just being amazing people. I want to thank Paula and Uwe for their support and love.

I thank my friends, in California: Mike Tsiang, Sam Fok and the rest of our marathon club Team Awesome, Dogan, Parthi, Thomas, Le, Nora and Emma, Murad and Forest. At home in Europe, my crews in Cologne (Waldsiedlung!!), Heidelberg, Berlin, and other places.

Finally, I want to thank Jenny for being there these past years to witness it all, for love and unwavering faith in me! For being an inspiring, strong person, for being hilarious, and for having patience with me, especially during this final year. I love you!

Contents

1. Introduction 1

2. Quantum Feedback Networks 3

2.1. Quantum Stochastic Differential Calculus . . . 3
  2.1.1. The necessity of the Hudson Parthasarathy generator . . . 7
  2.1.2. The Heisenberg picture QSDEs . . . 8
  2.1.3. The output noise processes . . . 10
2.2. SLH and the Gough-James Circuit Algebra . . . 10
  2.2.1. Fundamental circuit operations . . . 11
2.3. Linear quantum feedback networks . . . 14
  2.3.1. Transferfunction and Squeezing . . . 15
  2.3.2. Coupled OPOs for spectral shaping of squeezed light . . . 15

3. Specification of photonic circuits using Quantum Hardware Description Language 19

3.1. Modeling quantum circuitry . . . 20
  3.1.1. The circuit algebra . . . 20
  3.1.2. The QHDL syntax . . . 22
  3.1.3. Parsing a network . . . 24
  3.1.4. The QHDL workflow . . . 25
3.2. An example of the QHDL workflow . . . 27
  3.2.1. The two-cavity pseudo-NAND-latch . . . 27
  3.2.2. Model reduction in the SLH context . . . 32
3.3. Conclusion . . . 36
3.4. Reduced parameters in case of signal feedback . . . 37
3.5. Latch circuit library file . . . 38

4. Ultra-Low-Power All-Optical Computation 41

4.1. Basic Single Mode Kerr Model . . . 43
  4.1.1. Semi-classical steady-state analysis . . . 44
  4.1.2. Linear Transfer Function . . . 47

4.2. Phase Sensitive Amplifier Model 48
4.2.1. Kerr Amplifier Recipe 49
4.3. Symmetric amplifier design 50
4.4. Quadrature Filter 51
4.4.1. Dynamically Resonant Quadrature Filter 53
4.4.2. Noise properties of the dynamically resonant quadrature filter 54
4.5. Two-mode Kerr-models 55
4.5.1. Cross-Kerr-effect phase modulator 57
4.5.2. An all-optical Fredkin gate 60
4.5.3. Two-mode-thresholder 60
4.6. Non-degenerate optical parametric oscillators 61
4.6.1. Eliminated pump model 63
4.6.2. Fixpoints and stability 63
4.6.3. Below threshold 64
4.6.4. Above threshold 64
4.7. A Bifurcating Kerr Amplifier 68
4.8. Final remarks on optical computing 74

5. A coherent perceptron for all-optical learning **75**
5.1. The Perceptron algorithm 77
5.1.1. The circuit modeling framework 78
5.2. The Coherent Perceptron Circuit 79
5.2.1. Tunable Gain Kerr-amplifier 81
5.2.2. Encoding and Storing the Gain 83
5.2.3. Programmable Gain Amplifier 84
5.2.4. Optical Switches 86
5.2.5. Generation of the Estimated Label 87
5.3. Results 88
5.3.1. Time scales and power budget 90
5.4. Conclusion and Outlook 92
5.5. Basic Component Models 93
5.5.1. Static, Linear Circuit Components 93
5.5.2. Resonator Models 94

6. Exact co-simulation of semi-classical and quantum dynamics **101**
6.1. Motivation 101

6.2. Quantum state compression 102
  6.2.1. The complexity functional 104
  6.2.2. Example application: The degenerate parametric oscillator 108
6.3. Dynamics in a moving basis 115
6.4. Optimal coordinate dynamics 116
  6.4.1. Gradient coupled fiducial state dynamics 117
  6.4.2. Computational results 118
6.5. Examples 118
  6.5.1. Coherent displacement 120
  6.5.2. Squeezing and displacement 121
  6.5.3. Spin coherent states 121
6.6. Conclusion and Outlook 122
6.7. Useful additional material 122
  6.7.1. General construction of the coordinate transformation 122
  6.7.2. Properties of the quantum correlation 124

7. Conclusion and Outlook **127**

A. Simplification rules for the circuit algebra **129**
A.1. Simplifying algebraic circuit expressions 129
  A.1.1. Basic algebraic properties 129
  A.1.2. Permutation objects 132
  A.1.3. Permutations and Concatenations 133
  A.1.4. Feedback of a concatenation 136
  A.1.5. Feedback of a series 137

Bibliography **141**

1. Introduction

I joined Hideo Mabuchi's group in early 2011 at a time when it appeared as if the general field of quantum information technology was having a bit of a hangover after the excitement of the 1990's and early 2000's.

Scaling up quantum systems with well-controllable qubits has proven harder than expected with theoretical advances having run greatly ahead of experimental capabilities. Recently, however, steady advances in several of the leading hardware platforms have even led to strong private sector investment into quantum technology and it is therefore a wonderful time work in this field.

What excited me especially about the work being done in Hideo Mabuchi's group was that Hideo very explictly re-framed the problems of designing quantum information processing hardware as belonging to a genuinely new engineering discipline that required a whole new set of methods, tools and intuition. I thoroughly agree with his vision that, in particular, developing methods for coherent feedback quantum control and more generally quantum feedback networks will aid our understanding of how to incrementally synthesize complex and potentially quite heterogeneous quantum systems. Circuits and networks have been an immensely successful way of exploring the physically realizable design space given some basic building blocks. This should evenutally enable tayloring quantum systems for particular information processing tasks.

Although ultimately key enablers of 'quantum enhanced' information technology will be advances in material and individual device fabrication quality, there are some incredibly important issues that are not addressed by merely improving individual qubit devices or even by solving the problem of single qubit error correction and two-qubit fault tolerant gates: Currently proposed schemes appear to envision a quantum computer connected to an extremely sophisticated classical co-processor that processes error syndromes and other measurements, and applies optimal control pulses to achieve correcting feedback or drive specific sequences of quantum operations. However, given that the currently most promising quantum hardware platforms all require cryogenic environments, this external classical co-processor will likely have to be physically separated from its quantum counterpart necessitating a very large overhead of communication into and out from the cryogenic environment.

It is here where the framework of quantum feedback networks can be of immense value as it provides the potential of greatly reducing this ‘classical hardware’ overhead required for controlling a quantum device. By providing a unified language capable of describing heterogeneous networks of systems of varying level of dissipation and thus quantum coherence it becomes possible to design devices that autonomously self-stabilize [57, 58] and that may not require externally generated optimal control signal pulses for driving specific quantum interactions but that instead can be externally programmed in a very high level way. In the case of the above cited quantum memory proposals by Joe Kerckhoff, and also in the proposed ultra low-power all-optical decoder of Low Density Parity Check Codes [87] developed by Dmitri Pavlichin, all one must do is supply power and the hardware autonomously carries out useful operations.

The organization of this thesis roughly follows my research projects in chronological order. In Chapter 2 I introduce the theoretical framework underlying most of my work, especially that on Quantum Hardware Description Language which is presented in Chapter 3. In Chapter 4 I then turn to analyzing various building blocks featuring small networks of ultra-low power nonlinear optical resonators and propose concrete ways in which their intrinsic dynamics enables useful computation. This effort culminates in Chapter 5 in an end-to-end description of a neuromorphic, all-optical design proposal for a system that can incrementally learn through supervised training. Finally, Chapter 6 presents a novel description and simulation method for accurately representing states of open quantum systems in low dimensional bases. This method naturally yields coupled equations of motion for the reduced complexity quantum state and a set of semi-classical manifold coordinates, which will enable applying the insights gained by studying semi-classical circuit models to designing hybrid quantum-classical networks for information processing tasks.

2. Quantum Feedback Networks

In this chapter we present the general mathematical framework used to model quantum feedback networks. We first introduce a necessary and convenient quantum stochastic differential calculus and the most general Hudson Parthasarathy parametrization of Markovian open quantum systems interacting with a finite number of bosonic baths. We then proceed to summarize the extremely useful $(\mathbf{S}, \mathbf{L}, H)$ circuit algebra developed by Gough and James and finally finish with a brief description of linear quantum feedback networks and an example application.

2.1. Quantum Stochastic Differential Calculus

The framework of Quantum Stochastic Differential Equations (QSDEs) is quite formidable and can be a bit intimidating when starting out. The classic and most accessible reference for this is given by Gardiner and Collett [35] but they only derive the Heisenberg picture evolution of all operators and not a general QSDE for the unitary system+environment propagator $U(t)$.

The most general parametrization of the full QSDE for the propagator was actually worked out by Hudson and Parthasarathy in [51] and it was connected to continuous measurement processes by Barchielli and Lupieri [4], but despite some early work on cascading quantum input-output systems by Gardiner and Carmichael [33, 15] it took another 15 years before the full Hudson-Partharasathy QSDE parametrization which consists of the three objects $(\mathbf{S}, \mathbf{L}, H)$ gained a full operational interpretation and really became a practical tool in formulating a theory of Quantum Feedback Networks [40, 41].

Additionally, powerful insight was generated by the very powerful adiabatic elimination theorems proved by Bouten, Van Handel and Silberfarb [12, 13], primarily because they allowed deriving models with scattering matrices containing non-trivial operators from initial system models that are much closer to the somewhat more fundamental physical theory of cavity QED.

The QSDE for the propagator contains all the information about system and bath observables. For any specific (continuous) measurement one may then derive the corresponding *quantum filter*, i.e., a prescription of how to update the current estimate of the system state based on the measured signal. Since having the quantum state allow us to compute a prior probability distribution for the measurement signal, we can also use these known statistics to stochastically sample different measurement *trajectories* and consequently the corresponding information state[1] of the system.

Many further researchers contributed to this framework over the years and I will not be able to list all relevant resources, but some good references are [34, 38, 137, 136, 79, 15, 33, 134, 135] and more recently [10, 39, 11, 40, 12, 13, 41, 140].

I will not re-derive the Hudson-Parthasarathy QSDE from a more fundamental model here. The main difficulty encountered here lies in transforming a white Schrödinger equation from what is effectively the Stratonovich picture to the Ito formalism. Ryan Hamerly has a nice derivation in his thesis [46].

Instead, I will demonstrate the algebraic consistency of the approach and how – given the formal requirements – it really could not be formulated any different way. I will also try and provide as much physical intuition as possible to the $(\mathbf{S}, \mathbf{L}, H)$ coefficients.

We first present the result, the full QSDE for the unitary propagator of a system coupled to bosonic quantum fields

$$dU(t) = \underbrace{\left\{-\left(\frac{\mathbf{L}^\dagger \mathbf{L}}{2} + iH\right) dt - \mathbf{L}^\dagger \mathbf{S}\, \mathbf{dB}(t) + \mathbf{dB}^\dagger(t)\, \mathbf{L} + \mathrm{Tr}\left((\mathbf{S} - \mathbf{1}) \mathbf{d\Lambda}(t)^T\right)\right\}}_{dG(t)} U(t)\,. \quad (2.1)$$

This equation is in Ito form, i.e., all differentials are *forward differentials* $A(t)dQ(t) := A(t)[Q(t+dt) - Q(t)]$, which is very convenient when deriving further results from the above QSDE. The Ito form is also the primary reason why (2.1) does not look like a typical differential equation for a unitary operator: As it is, the 'generator' $dG(t)$ is not an anti-hermitian operator, i.e., it does not obviously correspond to a Hamiltonian.

If we were to work instead in the Stratonovich convention the above differential equation would be generated by a manifestly hermitian Hamiltonian. To see this I point you to Chapter 1 Section 3 of Ryan Hamerly's thesis [46] in which he explicitly converts an SDE from a central difference form that would correspond to the Stratonovich picture to the Ito form shown above. A more formal treatment but less accessible to non-mathematicians can be found in [39].

The most important assumption underlying the formalism in either convention is the white noise approximation which allows us to greatly simplify the description of the external fields and derive Markovian dynamics for a system interacting with such fields.

If we describe our external fields in terms of continuously labeled bosonic mode operators $\{b_j(t), b_j^\dagger(t), t \in [-\infty, \infty), j = 1, 2, \ldots, n\}$ then the white noise approximation essentially boils down to assuming the following commutator relationship:

---

[1] I.e., the most likely quantum state conditioned on the measurement record.

$$[b_j(t), b_k^\dagger(t')] = \delta_{jk}\delta(t-t'). \tag{2.2}$$

Due to the singular nature of the commutator relationship it is useful to introduce integrated field processes

$$B_j(t) := \int^t b_j(t')dt' \equiv \int^t dB_j(t')\,, \tag{2.3}$$

$$B_j^\dagger(t) := \int^t b_j^\dagger(t')dt' \equiv \int^t dB_j^\dagger(t')\,, \tag{2.4}$$

$$\Lambda_{jk}(t) := \int^t b_j^\dagger(t')b_k(t')dt' \equiv \int^t d\Lambda_{jk}(t')\,, \tag{2.5}$$

where we leave the lower integration bound unspecified as we will always be interested in differences between such integrals.

Here, the discrete indices j label the external field channel. For a resonator in which we neglect all but one resonant internal mode, there will be one ‘channel’ for each way in which photons can exit the resonator. Although the reality of some loss mechanism may physically be quite different from such an external bosonic bath mode, in many cases it works extremely well for the purpose of building accurate models [14].

For a system with several internal resonances, each physical environment coupler (e.g., a cavity mirror) can correspond to interaction with multiple baths if those internal resonances are spectrally very well separated. In any intermediate case some care must be taken and the formalism should not be applied blindly.

The continuous label t *does not* simply represent time in the sense of Heisenberg picture operators. Instead, it denotes the time at which this particular ‘time slice’ of the electromagnetic field interacts with the system. On the other hand, under the dynamics described by (2.1) the external field of a given channel and time is unchanged until the exact time it interacts with the system.

A final ingredient in both motivating and defining the calculus lies in the constraint that the external field is initially in a vacuum state or a coherently displaced field. Although one can still derive Heisenberg picture QSDEs for more general Gaussian field states, such as thermal states and squeezed states, there is then no general QSDE for the full unitary propagator due to additional formal problems affecting related to the *Gauge process* $\mathbf{\Lambda}(t)$.

Ultimately, we are always interested in computing certain expectation values such as correlation functions for the output fields. As we will see below, these will generally contain a contribution due to the system output but also a singular part due to the field itself. Since the output processes are effectively Heisenberg picture operators we can evaluate expectations in the initial state. For a vacuum or coherent field state, we can avoid all singularities by bringing the operators into normal order[2] since

$$
\begin{aligned}
&\langle \beta_1(t), \ldots, \beta_n(t) | \, b^\dagger_{j_1}(t) \cdots b^\dagger_{j_m}(t) b_{k_1}(t) \cdots b_{k_n}(t) \, | \beta_1(t), \ldots, \beta_n(t) \rangle & (2.6)\\
&\qquad = \beta^*_{j_1}(t) \cdots \beta^*_{j_m}(t) \beta_{k_1}(t) \cdots \beta_{k_n}(t). & (2.7)
\end{aligned}
$$

As an example let us say we are interested in carrying out a homodyne field measurement of a laser beam. This exactly corresponds to measuring the x-quadrature of a bosonic quantum noise process $X(t) - X(0) := B(t) - B(0) + B^\dagger(t) - B^\dagger(0)$ in a coherent state of constant amplitude β. This quantity appears to be of order $O(t)$ and we can easily verify that its expectation is $\langle X(t) - X(0) \rangle_\beta = [\beta + \beta^*] t$. However, if we ask for the expectation of its square we find

$$
\begin{aligned}
\langle \beta | \left(X(t) - X(0) \right)^2 | \beta \rangle &= \langle \beta | \left(\int_0^t [b(t) + b^\dagger(t)] dt \right)^2 | \beta \rangle & (2.8)\\
&= \langle \beta | \int_0^t \int_0^t \left[b(t') + b^\dagger(t') \right] \left[b(t'') + b^\dagger(t'') \right] dt' dt'' \, | \beta \rangle & (2.9)\\
&= \langle \beta | \int_0^t \int_0^t \left[b(t') b(t'') + b^\dagger(t') b(t'') \right] dt' dt'' \, | \beta \rangle & (2.10)\\
&\quad + \langle \beta | \int_0^t \int_0^t \left[\underbrace{b(t') b^\dagger(t'')}_{b^\dagger(t'') b(t') + \delta(t' - t'')} + b^\dagger(t') b^\dagger(t'') \right] dt' dt'' \, | \beta \rangle & (2.11)\\
&= (\beta + \beta^*)^2 t^2 + t. & (2.12)
\end{aligned}
$$

Thus the quantity $X(t) - X(0)$ is in this sense $O(t)$ but $[X(t) - X(0)]^2$ also has an $O(t)$ contribution. This complicates working with these quantities when taking the $t \to 0$ limit, i.e., when working with differential quantities.

To address this, one defines an extended Quantum Ito calculus that allows to still define generalized product and chain rules for such quantum stochastic differentials. In the example above, taking $t \to dt$ we then have $[X(dt) - X(0)]^2 = [dB(t) + dB(t)^\dagger]^2 = \underbrace{dB^2(t) + dB^{\dagger 2}(t) + dB^\dagger(t) dB(t)}_{0} + \underbrace{dB(t) dB^\dagger(t)}_{dt}$.

More generally we can derive a full Ito table (Table 2.1) based on bringing all differentials into normal order and substituting "$\delta(0) dt \approx 1$". Then, taking the product of any two Quantum stochstic processes $Q = AB$, the differential is given by

$$dQ = dAB + AdB + dAdB \tag{2.13}$$

with the Ito rules applied to the last term. A generalization of the chain rule is then given by

$$F = f(A) \Rightarrow dF = f'(A) dA + \frac{1}{2} f''(A) dA dA. \tag{2.14}$$

---

[2] I.e., creation operators are commuted all the way to the left of annihilation operators.

$dX \setminus dY$	dt	dB_j	$dB_j^\dagger$	$d\Lambda_{jk}$
dB_m	0	0	$\delta_{mj}dt$	$\delta_{mj}dB_k$
$dB_m^\dagger$	0	0	0	0
$d\Lambda_{mn}$	0	0	$\delta_{nj}dB_m^\dagger$	$\delta_{nj}d\Lambda_{mk}$

Table 2.1.: The Quantum Ito table. Each entry gives the result of the product $dXdY$ where dX is enumerated downwards and dY is enumerated to the right.

We now proceed to demonstrate that the parametrization of the Hudson-Parthasarathy QSDE is in some sense the most general.

2.1.1. The necessity of the Hudson Parthasarathy generator

We now prove that any Ito QSDE that generates a unitary quantum stochastic process and features the quantum noise processes introduced above, is necessarily of Hudson Parthasarathy form, i.e., the coefficients appearing in the QSDE will necessarily obey the same constraints. To see this, assume first that we wished to define an alternate QSDE that generates a unitary quantum stochastic process $V(t)$. We make the following Ansatz:

$$dV = \left[K dt + \mathbf{M}^\dagger \mathbf{dB} + \mathbf{dB}^\dagger \mathbf{N} + \mathrm{Tr}\left(\mathbf{Q}\mathbf{d\Lambda}^T\right)\right] V(t) \tag{2.15}$$

where $\mathbf{N}, \mathbf{M}$ are n dimensional vectors of system operators and we generally take $\mathbf{X}^\dagger$ to be the transpose *and* element-wise adjoint of a matrix or vector of operators. $\mathbf{Q}$ is an $n \times n$ matrix of system operators. We assume that the noise increments $\mathbf{dB}, \mathbf{dB}^\dagger, \mathbf{d\Lambda}$ commute with any system operator and they also commute with the unitary $V(t)$ itself since they are forward differentials (this is where using the Ito formalism pays off). Demanding that $V(t)$ be unitary at all times is equivalent to demanding that $V(t)^\dagger V(t) \equiv \mathbf{1}$. Applying our Ito-product rule yields

$$\begin{aligned}
d[V^\dagger V] &= dV^\dagger V + V^\dagger dV + dV^\dagger dV & (2.16)\\
&= V^\dagger \left[K^\dagger dt + \mathbf{dB}^\dagger \mathbf{M} + \mathbf{N}^\dagger \mathbf{dB} + \mathrm{Tr}\left(\mathbf{Q}^\dagger \mathbf{d\Lambda}^T\right)\right] & (2.17)\\
&\quad + V^\dagger \left[K dt + \mathbf{M}^\dagger \mathbf{dB} + \mathbf{dB}^\dagger \mathbf{N} + \mathrm{Tr}\left(\mathbf{Q}\mathbf{d\Lambda}^T\right)\right] V & (2.18)\\
&\quad + V^\dagger \left[K^\dagger dt + \mathbf{dB}^\dagger \mathbf{M} + \mathbf{N}^\dagger \mathbf{dB} + \mathrm{Tr}\left(\mathbf{Q}^\dagger \mathbf{d\Lambda}^T\right)\right] & (2.19)\\
&\quad\quad \times \left[K dt + \mathbf{M}^\dagger \mathbf{dB} + \mathbf{dB}^\dagger \mathbf{N} + \mathrm{Tr}\left(\mathbf{Q}\mathbf{d\Lambda}^T\right)\right] V & (2.20)
\end{aligned}$$

We can now apply the Ito table to the last two rows and collect contributions to different differentials:

$$\begin{aligned} d[V^\dagger V] &= V^\dagger \left[K^\dagger + K + \mathbf{N}^\dagger \mathbf{N}\right] V dt \quad &(2.21)\\ &+ V^\dagger \mathbf{dB}^\dagger \left[\mathbf{M} + \mathbf{N} + \mathbf{Q}^\dagger \mathbf{N}\right] V \quad &(2.22)\\ &+ V^\dagger \left[\mathbf{M} + \mathbf{N} + \mathbf{Q}^\dagger \mathbf{N}\right]^\dagger \mathbf{dB} V \quad &(2.23)\\ &+ V^\dagger \mathrm{Tr}\left(\left[\mathbf{Q} + \mathbf{Q}^\dagger + \mathbf{Q}^\dagger \mathbf{Q}\right] \mathbf{d\Lambda}^T\right) V \quad &(2.24) \end{aligned}$$

We must require all the square bracketed expressions to vanish independently. The first constraint is most generally satisfied by

$$K^\dagger + K + \mathbf{N}^\dagger \mathbf{N} = 0 \Leftrightarrow K = -iH - \frac{1}{2}\mathbf{N}^\dagger \mathbf{N}, \tag{2.25}$$

where $H = H^\dagger$ is an arbitrary Hermitian system operator that we will come to identify as the system Hamiltonian. The last constraint can be equivalently written as

$$[\mathbf{1} + \mathbf{Q}]^\dagger [\mathbf{1} + \mathbf{Q}] = \mathbf{1}, \tag{2.26}$$

which suggests that $\mathbf{S} := \mathbf{1} + \mathbf{Q}$ is a unitary matrix of system operators. With this we can rewrite the second (and the equivalent third) constraint as

$$\mathbf{M} + \mathbf{S}^\dagger \mathbf{N} = 0 \Leftrightarrow \mathbf{M} = -\mathbf{S}^\dagger \mathbf{N}. \tag{2.27}$$

If we now make the final relabeling $\mathbf{N} \to \mathbf{L}$ and reinsert all these into the original QSDE for V we find

$$dV = \left[-\left(iH + \frac{1}{2}\mathbf{N}^\dagger\right) dt - \mathbf{L}^\dagger \mathbf{S} \mathbf{dB} + \mathbf{dB}^\dagger \mathbf{L} + \mathrm{Tr}\left([\mathbf{S} - \mathbf{1}]\mathbf{d\Lambda}^T\right)\right] V(t) \tag{2.28}$$

which is fully identical to (2.1)! We can thus see that any unitary operator process generated by the linear and quadratic noise operators $\mathbf{dB}, \mathbf{dB}^\dagger$ and $\mathbf{d\Lambda}$ must have the Hudson-Parthasarathy form. We now proceed to give operational meaning to the $(\mathbf{S}, \mathbf{L}, H)$ parameters by deriving the Heisenberg equations of motion as well as the input-output relationship.

2.1.2. The Heisenberg picture QSDEs

Defining the Heisenberg picture operators as $j_t(X) := U^\dagger(t) X U(t)$ we derive its SDE via the Ito rules:

$$\begin{aligned} dj_t(X) &= dU^\dagger X U + U^\dagger X dU + dU^\dagger X dU \quad &(2.29)\\ &= U^\dagger \left(i[H, X] + \frac{1}{2}\mathbf{L}^\dagger [X, \mathbf{L}] + \frac{1}{2}[\mathbf{L}^\dagger, X]\mathbf{L}\right) U dt \quad &(2.30)\\ &\quad + \mathbf{dB}^\dagger U^\dagger \mathbf{S}^\dagger [X, \mathbf{L}] U + U^\dagger [\mathbf{L}^\dagger, X] \mathbf{S} U \mathbf{dB} \quad &(2.31)\\ &\quad + \mathrm{Tr}\left(U^\dagger \left[\mathbf{S}^\dagger X \mathbf{S} - X\right] U \mathbf{d\Lambda}^T\right). \quad &(2.32) \end{aligned}$$

We can rewrite this as

$$\begin{aligned} dj_t(X) = j_t &\left(\underbrace{i[H,X] + \frac{1}{2}\mathbf{L}^\dagger[X,\mathbf{L}] + \frac{1}{2}[\mathbf{L}^\dagger, X]\mathbf{L}}_{\mathcal{L}^* X} \right) dt \\ &+ \mathbf{dB}^\dagger j_t\left(\mathbf{S}^\dagger[X,\mathbf{L}]\right) + j_t\left([\mathbf{L}^\dagger, X]\mathbf{S}\right)\mathbf{dB} \\ &+ \operatorname{Tr}\left(j_t\left(\mathbf{S}^\dagger X \mathbf{S} - X\right) \mathbf{d\Lambda}^T\right) , \end{aligned} \tag{2.33}$$

where we have implicitly defined the adjoint Liouville super operator $\mathcal{L}^*X := i[H,X] + \frac{1}{2}\mathbf{L}^\dagger[X,\mathbf{L}] + \frac{1}{2}[\mathbf{L}^\dagger, X]\mathbf{L}$. This Heisenberg picture QSDE is very useful as we can compute expectation values in an initially factorizable state $\rho = \rho_S \otimes \Omega_R$ where Ω_R denotes the collective vacuum state for the external fields. All the quantum noise increments have zero expectation in the vacuum state and therefore the expectation of X evolves as

$$d\,\langle X \rangle_{\rho_S\otimes\Omega_R} = \left\langle j_t \underbrace{\left(i[H,X] + \frac{1}{2}\mathbf{L}^\dagger[X,\mathbf{L}] + \frac{1}{2}[\mathbf{L}^\dagger, X]\mathbf{L} \right)}_{\mathcal{L}^* X} \right\rangle_{\rho_S} dt \tag{2.34}$$

$$= \operatorname{Tr}\left(\rho_S U^\dagger(t)[\mathcal{L}^* X] U(t)\right) dt \tag{2.35}$$

$$= \operatorname{Tr}\left(\underbrace{U(t)\rho_S U^\dagger(t)}_{\rho_S(t)} \mathcal{L}^* X \right) dt \tag{2.36}$$

$$= \operatorname{Tr}\left([\mathcal{L}\rho_S(t)]X\right) dt \tag{2.37}$$

$$\overset{!}{=} \operatorname{Tr}\left(d\rho_S(t) X\right) \tag{2.38}$$

where we have taken the super-adjoint of $\mathcal{L}^*$ which is given by

$$\mathcal{L}\rho = -i[H,\rho] + \sum_k \left[L_k \rho L_k^\dagger - \frac{1}{2}\{L_k^\dagger L_k, \rho\} \right] . \tag{2.39}$$

Thus, we have derived the standard Lindblad master equation directly from the Heisenberg picture QSDE:

$$\dot{\rho}(t) = \mathcal{L}\rho = -i[H,\rho] + \sum_k \left[L_k \rho L_k^\dagger - \frac{1}{2}\{L_k^\dagger L_k, \rho\} \right] . \tag{2.40}$$

The formalism is quite powerful and can also be used to derive dynamics conditioned on a continuous measurement (quantum filtering) but a more general treatment would exceed the scope of this introduction. A thorough overview of this is given in [11].

2.1.3. The output noise processes

Defining the output field processes as

$$\mathbf{B}'(t) := U^\dagger(t)\mathbf{B}(t)U(t) \tag{2.41}$$

$$\mathbf{\Lambda}'(t) := U^\dagger(t)\mathbf{\Lambda}(t)U(t) \tag{2.42}$$

we can again use the Ito rules to derive their associated SDEs

$$\mathbf{dB}'(t) = j_t(\mathbf{S})\mathbf{dB}(t) + j_t(\mathbf{L})dt \tag{2.43}$$

$$\mathbf{d\Lambda}'(t) = j_t(\mathbf{S}^\sharp)\mathbf{d\Lambda}(t)j_t(\mathbf{S}^T) + j_t(\mathbf{S}^\sharp)\mathbf{dB}^\sharp(t)j_t(\mathbf{L}^T) \tag{2.44}$$

$$+ j_t(\mathbf{L}^\sharp)\mathbf{dB}^T(t)j_t(\mathbf{S}^T) + j_t(\mathbf{L}^\sharp\mathbf{L}^T)dt \tag{2.45}$$

where we use the same $\mathbf{X}^\sharp$ notation as Gough and James [41] for the elementwise adjoint of an operator matrix $\mathbf{X}$. Equation (2.43) is especially enlightening as it demonstrates that the scattering matrix literally 'scatters' input fields to outputs and that the $\mathbf{L}$ operator appears as a linear contribution to the output field and as a quadratic contribution to the output Gauge process. In the next section I will introduce the Gough-James circuit algebra which will provide further intuition for the $(\mathbf{S}, \mathbf{L}, H)$ parametrization.

2.2. SLH and the Gough-James Circuit Algebra

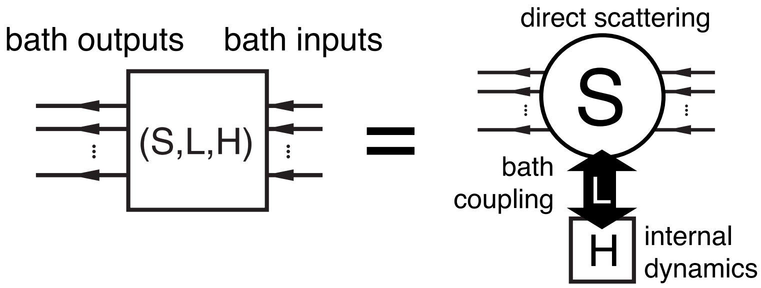


Figure 2.1.: A quantum network component model interacting with n external quantum fields is parametrized by three objects: a scattering matrix S mediating direct scattering of inputs to outputs, a coupling vector L describing the coupling of each external field to the internal degrees of freedom and a Hamiltonian that induces the internal dynamics.

In [40, 41], Gough and James have introduced an algebraic method to derive the QSDE $(\mathbf{S}, \mathbf{L}, H)$ parameters for a full network of cascaded quantum systems from the individual

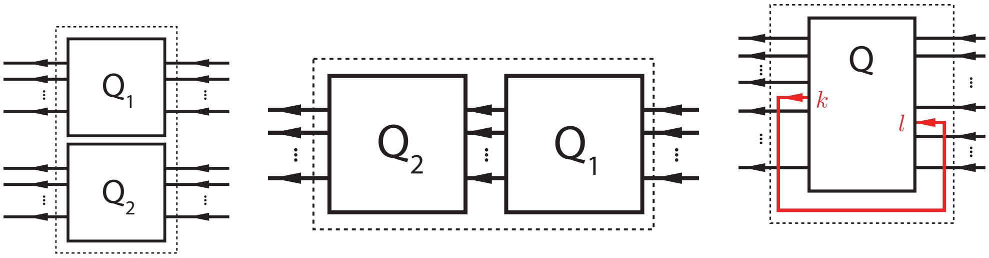
$(\mathbf{S}, \mathbf{L}, H)$ parameters of its constituents. A general system with an equal number n of input and output channels is described by the parameter triplet $(\mathbf{S}, \mathbf{L}, H)$, where H is the effective internal *Hamiltonian* for the system, $\mathbf{L} = (L_1, L_2, \ldots, L_n)^T$ the *coupling vector* and $\mathbf{S} = (S_{jk})_{j,k=1}^n$ is the *scattering matrix* (whose elements are themselves operators). An element L_k of the coupling vector is given by a system operator that describes the system's coupling to the k-th output channel. Similarly, the elements S_{jk} of the scattering matrix are in general given by system operators describing the scattering between different field channels j and k. We have visualized the role of the individual operators in Figure 2.1.

As we have explicitly verified in Section 2.1.1, the only conditions on the parameters are that the Hamiltonian is self-adjoint and the scattering matrix is unitary:

$$H^* = H \text{ and } \mathbf{S}^\dagger \mathbf{S} = \mathbf{S}\mathbf{S}^\dagger = \mathbb{1}_n. \tag{2.46}$$

We adhere to the conventions used by Gough and James, i.e., the imaginary unit is given by $i := \sqrt{-1}$, and the adjoint of an operator A is given by A^*, the element-wise adjoint of an operator matrix $\mathbf{M}$ is given by $\mathbf{M}^\sharp$. Its transpose is given by $\mathbf{M}^T$ and the combination of these two operations, i.e. the adjoint operator matrix is given by $\mathbf{M}^\dagger = (\mathbf{M}^T)^\sharp = (\mathbf{M}^\sharp)^T$.

2.2.1. Fundamental circuit operations



(a) $Q_1 \boxplus Q_2$ (b) $Q_2 \lhd Q_1$ (c) $[Q]_{1\to 4}$

Figure 2.2.: Basic operations of the Gough-James circuit algebra.

In [41], Gough and James have introduced two operations that allow the construction of quantum optical 'feedforward' networks:

1. The *concatenation* product (Figure 2.2(a)) describes the situation where two arbitrary systems are attached to each other without optical scattering between the two systems' in- and output channels:

$$(\mathbf{S}_1, \mathbf{L}_1, H_1) \boxplus (\mathbf{S}_2, \mathbf{L}_2, H_2) = \left(\begin{pmatrix} \mathbf{S}_1 & 0 \\ 0 & \mathbf{S}_2 \end{pmatrix}, \begin{pmatrix} \mathbf{L}_1 \\ \mathbf{L}_1 \end{pmatrix}, H_1 + H_2 \right). \tag{2.47}$$

Note however, that even without optical scattering, the two subsystems may interact directly via shared quantum degrees of freedom.

2. The *series* product (Figure 2.2(b)) is to be used for two systems $Q_j = (\mathbf{S}_j, \mathbf{L}_j, H_j)$, $j = 1, 2$ of equal channel number n where all output channels of Q_1 are fed into the corresponding input channels of Q_2

$$(\mathbf{S}_2, \mathbf{L}_2, H_2) \lhd (\mathbf{S}_1, \mathbf{L}_1, H_1) = \left(\mathbf{S}_2\mathbf{S}_1, \mathbf{L}_2 + \mathbf{S}_2\mathbf{L}_1, H_1 + H_2 + \operatorname{Im}\left[\left\{\mathbf{L}_2^\dagger \mathbf{S}_2 \mathbf{L}_1\right\}\right]\right) \tag{2.48}$$

From their definition it can be seen that both the series product and the concatenation product not only yield valid circuit component triplets that obey the constraints (2.46), but they are also associative operations.[3] To make the network operations complete in the sense that it can also be applied to situations with optical feedback, an additional rule is required: The *feedback* operation (Figure 2.2(c)) describes the case where the k-th output channel of a system with $n \geq 2$ is fed back into the l-th input channel. The result is a component with $n - 1$ channels:

$$[\ (\mathbf{S}, \mathbf{L}, H)\ ]_{k \to l} = \left(\tilde{\mathbf{S}}, \tilde{\mathbf{L}}, \tilde{H}\right), \tag{2.49}$$

---

[3] For the concatenation product this is immediately clear, for the series product in can be quickly verified by computing $(Q_1 \lhd Q_2) \lhd Q_3$ and $Q_1 \lhd (Q_2 \lhd Q_3)$.

where the effective parameters are given by [40]

$$\tilde{\mathbf{S}} = \mathbf{S}_{[\cancel{k},\cancel{l}]} + \begin{pmatrix} S_{1l} \\ S_{2l} \\ \vdots \\ S_{k-1\,l} \\ S_{k+1\,l} \\ \vdots \\ S_{nl} \end{pmatrix} (1 - S_{kl})^{-1} \begin{pmatrix} S_{k1} & S_{k2} & \cdots & S_{kl-1} & S_{kl+1} & \cdots & S_{kn} \end{pmatrix}, \tag{2.50}$$

$$\tilde{\mathbf{L}} = \mathbf{L}_{[\cancel{k}]} + \begin{pmatrix} S_{1l} \\ S_{2l} \\ \vdots \\ S_{k-1\,l} \\ S_{k+1\,l} \\ \vdots \\ S_{nl} \end{pmatrix} (1 - S_{kl})^{-1} L_k, \tag{2.51}$$

$$\tilde{H} = H + \Im \left\{ \left[\sum_{j=1}^{n} L_j^* S_{jl} \right] (1 - S_{kl})^{-1} L_k \right\}. \tag{2.52}$$

Here we have written $\mathbf{S}_{[\cancel{k},\cancel{l}]}$ as a shorthand notation for the matrix $\mathbf{S}$ with the k-th row and l-th column removed and similarly $\mathbf{L}_{[\cancel{k}]}$ is the vector $\mathbf{L}$ with its k-th entry removed. These resulting parameters fulfill the conditions[4] for circuit components. Moreover, it can be shown that in the case of multiple feedback loops, the result is independent of the order in which the feedback operation is applied[5]. The possibility of treating the quantum circuits algebraically offers some valuable insights: A given full-system triplet $(\mathbf{S}, \mathbf{L}, H)$ may very well allow for different ways of decomposing it algebraically into networks of physically realistic subsystems. The algebraic treatment thus establishes a notion of dynamic equivalence between potentially very different physical setups. Given a certain number of fundamental building blocks such as beamsplitters, phases and cavities, from which we construct complex networks, we can investigate what kinds of composite systems can be realized. If we also take into account the adiabatic limit theorems for QSDEs [12, 13] the set of physically realizable systems is further expanded. Hence, the algebraic methods not only facilitate the analysis of quantum circuits, but ultimately may very well lead to an understanding of how to construct a general system $(\mathbf{S}, \mathbf{L}, H)$ from some set of elementary systems. There already exist some investigations along these lines for the particular subclass of *linear* systems [85] which can be thought of as a networked collection of quantum harmonic and parametric oscillators. Additional useful references for quantum feedback networks can be found in [140].

---

[4] This is obvious for $\tilde{\mathbf{L}}$ and $\tilde{H}$, for a proof that $\tilde{\mathbf{S}}$ is indeed unitary see Gough and James's original paper [40].

[5] Note however that some care has to be taken with the indices of the feedback channels when permuting the feedback operation.

2.3. Linear quantum feedback networks

There exists a special class of $(\mathbf{S}, \mathbf{L}, H)$ models consisting only of harmonic and/or parametric oscillator degrees of freedom represented by mode operators $\{a_k, a_k^\dagger, k = 1, 2, \dots, m\}$ interacting with n external fields. Among such systems, we may further constrain the $(\mathbf{S}, \mathbf{L}, H)$ parameters such that

1. $\mathbf{S} \in \mathbb{C}^{n\times n} \cap U(n)$, i.e., $\mathbf{S}$ is a purely number valued unitary matrix,
2. the coupling vector is linear in the mode operators which we assemble into a vector $\mathbf{a} = (a_1, \dots, a_m)^T$ such that $\mathbf{L} = \mathbf{C}_-\mathbf{a} + \mathbf{C}_+\mathbf{a}^\sharp$ with $\mathbf{C}_\pm \in \mathbb{C}^{n\times m}$
3. the Hamiltonian is quadratic[6] $H = \mathbf{a}^\dagger\Omega_-\mathbf{a} + \frac{1}{2}\left[\mathbf{a}^\dagger\Omega_+\mathbf{a}^\sharp + \mathbf{a}^T\Omega_+^\dagger\mathbf{a}\right]$ with $\Omega_-^\dagger = \Omega_-$ and $\Omega_\pm \in \mathbb{C}^{m\times m}$.

We call such systems linear quantum systems because the Heisenberg equations of motion (cf Equation (2.33)) for the mode operators are linear as well as the input-output relationship (cf Equation (2.43)). Specifically they can be cast into the form

$$\begin{aligned} \mathbf{d}\breve{\mathbf{a}}(t) &= \mathbf{A}\,\breve{\mathbf{a}}(t)\mathrm{d}t + \mathbf{B}\,\mathbf{d}\breve{\mathbf{B}}(t), && (2.53)\\ \mathbf{d}\breve{\mathbf{B}}'(t) &= \mathbf{C}\,\breve{\mathbf{a}}(t)\mathrm{d}t + \mathbf{D}\,\mathbf{d}\breve{\mathbf{B}}(t), && (2.54) \end{aligned}$$

where we use the doubled-up mode vectors of [43, 140] taken to be in the Heisenberg picture, i.e., $\mathbf{a}(t) = j_t(\mathbf{a})$:

$$\breve{\mathbf{a}}(t) = \begin{pmatrix} \mathbf{a}(t) \\ \mathbf{a}^\sharp(t) \end{pmatrix}, \quad \mathbf{d}\breve{\mathbf{B}} = \begin{pmatrix} \mathbf{dB}(t) \\ \mathbf{dB}^\sharp(t) \end{pmatrix}, \quad \mathbf{d}\breve{\mathbf{B}}' = \begin{pmatrix} \mathbf{dB}' \\ \mathbf{dB}'^* \end{pmatrix}. \tag{2.55}$$

The matrices are again defined through 'double-up notation' $\breve{\Delta}(\mathbf{X}, \mathbf{Y}) \equiv \begin{pmatrix} \mathbf{X} & \mathbf{Y} \\ \mathbf{Y}^* & \mathbf{X}^* \end{pmatrix}$, as

$$\mathbf{A} = \breve{\Delta}(\mathbf{A}_-, \mathbf{A}_+), \quad \mathbf{B} \equiv -\breve{\Delta}(\mathbf{C}_-^\dagger, -\mathbf{C}_+^T)\breve{\Delta}(\mathbf{S}, \mathbf{0}), \quad \mathbf{C} = \breve{\Delta}(\mathbf{C}_-, \mathbf{C}_+), \quad \mathbf{D} \equiv \breve{\Delta}(\mathbf{S}, \mathbf{0}), \tag{2.56}$$

and where

$$\mathbf{A}_\pm = -i\Omega_\pm - \frac{1}{2}\left(\mathbf{C}_-^\dagger\mathbf{C}_\pm - \mathbf{C}_+^T\mathbf{C}_\pm^\sharp\right). \tag{2.57}$$

---

[6] Linear drive terms are permitted in H as well as constant terms in $\mathbf{L}$ but they can easily be added later.

In addition to the linear systems described here, one may also define more general linear quantum feedback networks by directly specifying the $(\mathbf{A}, \mathbf{B}, \mathbf{C}, \mathbf{D})$ matrices and by relaxing the conditions on the input field states to general Gaussian states. The $(\mathbf{A}, \mathbf{B}, \mathbf{C}, \mathbf{D})$ matrices are always subject to some physical realizability constraints [84] but within these constraint one may construct systems that do not admit any $(\mathbf{S}, \mathbf{L}, H)$ representation, although in some case one may construct $(\mathbf{S}, \mathbf{L}, H)$ models that will approximate the linear system in some parameter limit [43].

2.3.1. Transferfunction and Squeezing

By Fourier transforming equations (2.53) and (2.54), we can find the system transfer function Ξ that directly links the Fourier transformed input and output fields $\tilde{\breve{\mathbf{b}}}^{(\prime)}(\omega) := \frac{1}{2\pi}\int_{-\infty}^{\infty} e^{i\omega t}\mathbf{d}\breve{\mathbf{B}}^{(\prime)}(t)$ as

$$\tilde{\breve{\mathbf{b}}}^{\prime}(\omega) = \Xi(\omega)\tilde{\breve{\mathbf{b}}}(\omega), \tag{2.58}$$

where

$$\Xi(\omega) \equiv \mathbf{D} + \mathbf{C}(-i\omega\mathbf{1} - \mathbf{A})^{-1}\mathbf{B}. \tag{2.59}$$

Due to the inherent redundancy of the doubled up notation, the transfer function can be decomposed as

$$\Xi(\omega) = \begin{pmatrix} \mathcal{S}^{-}(\omega) & \mathcal{S}^{+}(\omega) \\ \mathcal{S}^{+*}(-\omega) & \mathcal{S}^{-*}(-\omega) \end{pmatrix}. \tag{2.60}$$

From these matrices one can calculate the quadrature dynamics and eventually calculate the power spectral density of a given quadrature, which is known as squeezing spectrum [42]

$$\mathcal{P}_j^{\theta}(\omega) = 1 + \mathcal{N}_j(\omega) + \mathcal{N}_j(-\omega) + e^{2i\theta}\mathcal{M}_j(\omega) + e^{-2i\theta}\mathcal{M}_j(-\omega), \tag{2.61}$$

for $j = 1, 2, ...$ representing each output port and θ as the quadrature angle. In our case, the feedback system output is $j = 1$. Here, the parameters $\mathcal{N}_j(\omega)$ and $\mathcal{M}_j(\omega)$ are defined through

$$\mathcal{N}_j(\omega) \equiv \sum_k |\mathcal{S}_{jk}^{+}(\omega)|^2, \quad \mathcal{M}_j(\omega) \equiv \sum_k S_{jk}^{-}(\omega)\mathcal{S}_{jk}^{+}(-\omega), \tag{2.62}$$

where $\mathcal{S}_{jk}^{-}$ and $\mathcal{S}_{jk}^{+}$ are the (j,k) entries of the matrices $\mathcal{S}^{-}(\omega)$ and $\mathcal{S}^{+}(\omega)$, respectively. There are more general expressions for computing the multi-mode squeezing. We refer to [42] for more details.

2.3.2. Coupled OPOs for spectral shaping of squeezed light

In this section I will briefly summarize the theory and experimental results of an experiment carried out in our group first by Orion Crisafulli and then by Daniel Soh. These results were also published in [23]. The physical setup consisted of two degenerate parametric oscillators (DOPOs) in a feedback configuration. DOPOs are in some sense the ideal systems for realizing quantum limited, phase sensitive gain and producing squeezed light [66].

Squeezed light is characterized by an asymmetry in the variance of the field quadratures $\Delta X^2, \Delta P^2$ [130]. Like vacuum states or coherent states, squeezed states of light are minimum uncertainty states, i.e., they saturate Heisenberg's uncertainty principle $\Delta X \Delta P = \frac{\hbar}{2}$ but in contrast to vacuum states and coherent states the variances are different $\Delta X \neq \Delta P$. Squeezed states can be used for enhancing the sensitivity of any measurement in which the measured signal is encoded in a particular quadrature of some light signal. This has been exploited in increasing the sensitivity of interferometers such as those appearing in gravitational wave detectors [18].

Squeezed light can be produced in an optical parametric oscillator when the parametric gain becomes nearly as large as the linear losses. When the parametric gain approaches the linear loss, the linear model becomes unstable and the linear description breaks down. From [42] it is known that static self feedback of a DOPO can increase the amount of squeezing a system is capable of by effectively modulating the total linear loss of the system.

In our experiment, we investigated what happens when the feedback signal itself was acted on by a second DOPO. Specifically, a probe signal reflects off one mirror of a first DOPO which we will call the plant, then reflects off a mirror of a second DOPO which we call the controller and is then fed back to another mirror of the plant where it can interfere with that mirror's output field.

A full experimental schematic is shown in Figure 2.3.

We found that the feedback allows for substantially tuning the resulting squeezing spectrum by modifying the frequency at which maximum squeezing and and anti-squeezing occurs. The results are presented in Figure 2.4. This effect is strongly sensitive to the optical path lengths in the feedback loop as these influence the effectively formed supermode spectrum of the total network. A system such as ours could be very useful as a source of tunable squeezed light in sensing applications where the frequency of the signal to be measured can shift or is initially only known to be fairly close to the frequency of our squeezed light.

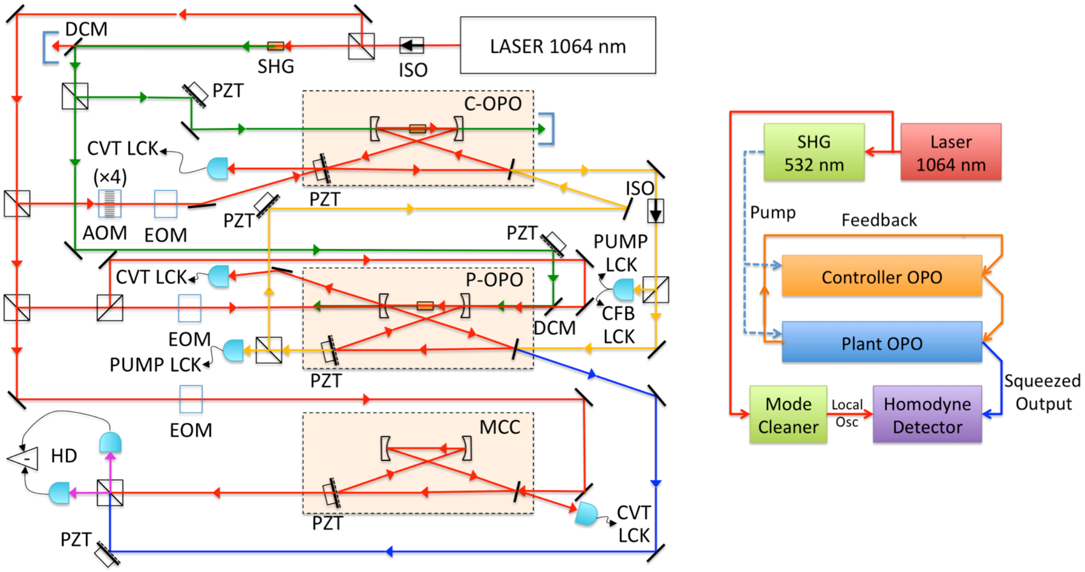
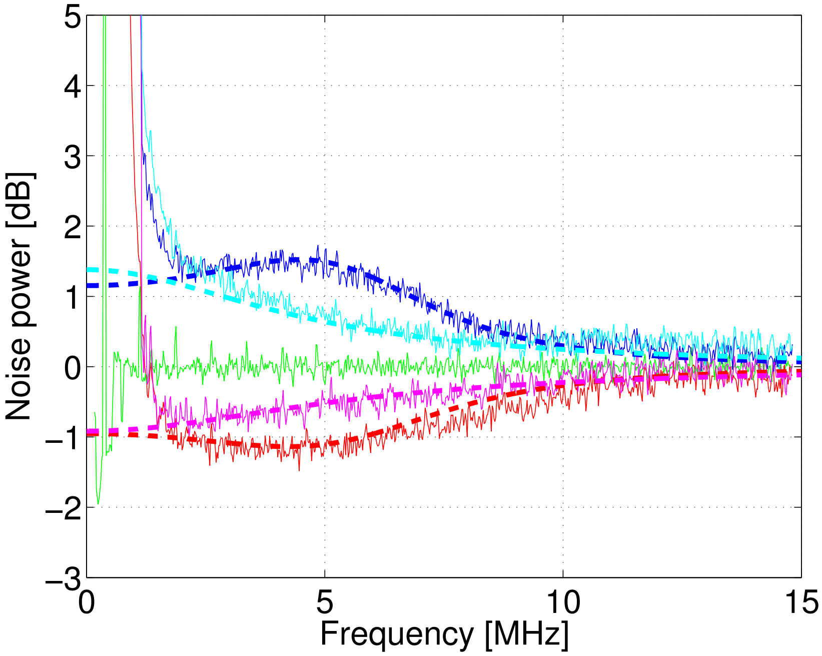
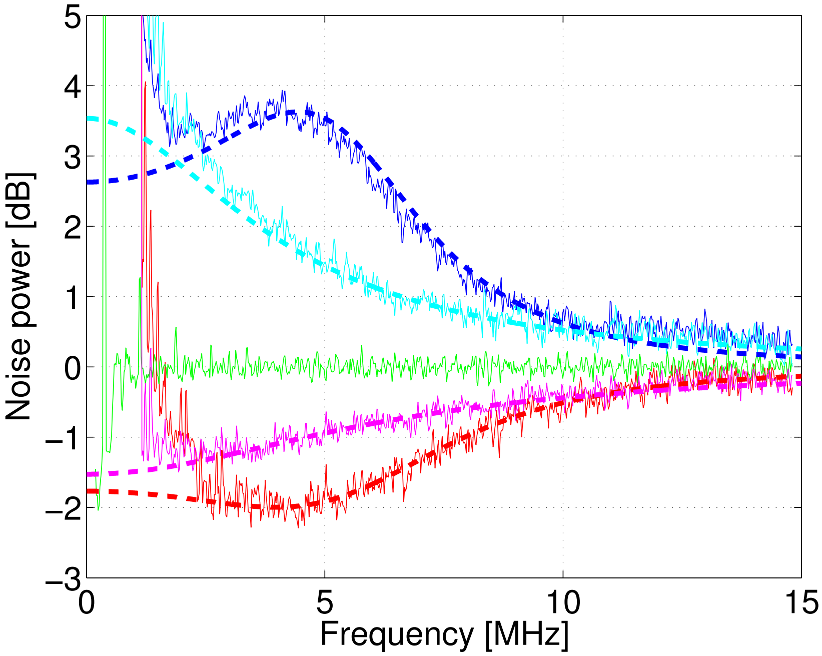

DCM
SHG
ISO
LASER 1064 nm
PZT
C-OPO
CVT LCK
(×4)
AOM
EOM
PZT
PZT
ISO
PZT
PUMP LCK
CVT LCK
P-OPO
CFB LCK
DCM
EOM
PUMP LCK
PZT
EOM
MCC
HD
PZT
CVT LCK
PZT
SHG 532 nm
Laser 1064 nm
Pump
Feedback
Controller OPO
Plant OPO
Squeezed Output
Mode Cleaner
Local Osc
Homodyne Detector

Figure 2.3.: The full experimental setup for two OPOs in a mutual feedback loop.



(a) $x = 0.17$



(b) $x = 0.33$

Figure 2.4.: Open-loop and closed loop squeezing spectra of the maximally amplified and squeezed quadratures. The green traces give the vacuum level which is identical for all quadratures. The pink and cyan traces give the open loop squeezing spectra of the maximally squeezed and anti-squeezed quadratures, respectively, while the red and dark blue traces were obtained in closed loop (though without pumping the controller). Figure 2.4 (a) is at relative plant pump amplitude of $x = 0.17$ and (b) at approximately twice the pump amplitude $x = 0.33$ where the actual pump power is then given by $x = \sqrt{\frac{P_p}{P_{\text{th}}}}$.

3. Specification of photonic circuits using Quantum Hardware Description Language

This chapter was also published as [117].

In this work we propose and demonstrate a modeling and simulation workflow based on schematic capture using a Quantum Hardware Description Language (QHDL) for nanophotonic circuits, which we will define as a proper subset of the standard VHSIC Hardware Description Language (VHDL). Our approach utilizes a mixture of common open-source software packages and custom processing scripts to provide a high-level, modular interface to the quantum circuit 'algebra' of Gough and James [40, 41] (which generalizes earlier work on cascaded open quantum systems by Carmichael [15] and by Gardiner [33]). The natural hierarchical organization of VHDL and the schematic capture workflow should facilitate future work on model reduction and design abstractions for nanophotonic circuits, which seems essential given the extremely high dimension (variable count) associated with many-component quantum models.

In the following sections we first review the formal setting of $(\mathbf{S}, \mathbf{L}, H)$ component models and the concatenation and series and products as introduced by Gough and James, which have recently been used to derive quantum nonlinear photonic circuit models by hand or using custom-coded computer algebra scripts [57, 58, 69, 68]. While we will restrict our attention here to linear and cavity nonlinear optics, it should be noted that the $(\mathbf{S}, \mathbf{L}, H)$ formalism can in principle be used to describe hybrid circuits incorporating suitable spintronic, nanomechanical [55], and/or quantum-electronic components. Likewise, the approach we describe here could be extended straightforwardly to admit static Bogoliubov components as described in [43]. We then review the proposed syntax of QHDL and illustrate its use in the specification of a simple interferometer as a network of elementary optical components. After describing methods that can be used to parse QHDL circuit descriptions to derive quantum equations of motion for analysis and/or numerical simulation, we illustrate the full schematic capture workflow using an example of constructing a bistable latch from cavity nonlinear optical components. The paper closes with a brief consideration of model reduction in the $(\mathbf{S}, \mathbf{L}, H)$ context.

3.1. Modeling quantum circuitry

Within this section we use $\{Q_j,\, j = 1, 2, 3, \ldots, N\}$ to denote individual quantum input-output components. We clearly distinguish between input and output ports and do not consider bi-directional ports, although for physical reasons every input port is assumed to have an associated output port and *v*ice versa.

3.1.1. The circuit algebra

Our modeling workflow is based on the Gough-James synthesis results for open quantum systems [40, 41], which provide a purely algebraic method to derive quantum Markov models for a network of interconnected quantum components.

A component with an equal number n of input and output channels is described by the parameter triplet $(\mathbf{S}, \mathbf{L}, H)$, where H is the effective internal *Hamilton operator* for the system, $\mathbf{L} = (L_1, L_2, \ldots, L_n)^T$ the *coupling vector* and $\mathbf{S} = (S_{jk})_{j,k=1}^n$ is the *scattering matrix*, whose elements are themselves operators.

Each element L_k of the coupling vector is given by an operator that describes the system's coupling to the k-th input channel. Similarly, the elements S_{jk} of the scattering matrix are given by system operators describing the scattering between different field channels j and k. The only mathematical conditions on the parameters are that the Hamiltonian operator be self-adjoint and the scattering matrix be unitary:

$$H^\dagger = H \text{ and } \mathbf{S}^\dagger \mathbf{S} = \mathbf{S}\mathbf{S}^\dagger = \mathbb{1}_n.$$

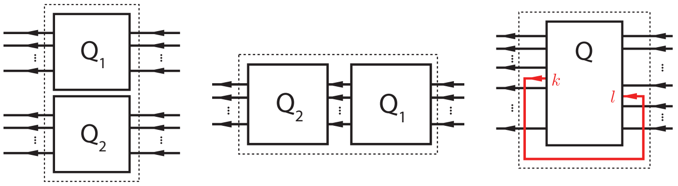
The master equation [36] corresponding to a given $(\mathbf{S}, \mathbf{L}, H)$ model is

$$\frac{d\rho_t}{dt} = -i[H, \rho_t] + \sum_{j=1}^n \left(L_j \rho_t L_j^\dagger - \frac{1}{2} \left\{ L_j^\dagger L_j, \rho_t \right\} \right) \tag{3.1}$$

Here $[A, B] \equiv AB - BA$ and $\{A, B\} \equiv AB + BA$, while ρ_t is a density matrix describing the evolving state of the internal degrees of freedom. It is also straightforward to obtain the quantum filtering equations [11, 138] for stochastic simulation of a given $(\mathbf{S}, \mathbf{L}, H)$ model.

While the scattering matrix elements S_{jk} do not appear in Eq. (3.1) they are required for the composition rules described below, which can be used to derive the overall parameter triplet for a network of interconnected quantum input-output components. The $(\mathbf{S}, \mathbf{L}, H)$ circuit algebra plus simple correspondences such as Eq. (3.1) provide all that is needed to obtain overall equations of motion for complex photonic circuits.

In [41], Gough and James have introduced two operations that allow for the construction of arbitrary networks of optical *feedforward* circuits:



(a) $Q_1 \boxplus Q_2$ (b) $Q_2 \lhd Q_1$ (c) $[Q]_{k\to l}$

Figure 3.1.: Basic operations of the Gough-James circuit algebra.

1. The *concatenation* product (cf. Figure 3.1(a)) describes a formal adjoining of two systems in which there is no optical scattering between the systems:

$$(\mathbf{S}_1, \mathbf{L}_1, H_1) \boxplus (\mathbf{S}_2, \mathbf{L}_2, H_2) \equiv \left(\begin{pmatrix} \mathbf{S}_1 & 0 \\ 0 & \mathbf{S}_2 \end{pmatrix}, \begin{pmatrix} \mathbf{L}_1 \\ \mathbf{L}_2 \end{pmatrix}, H_1 + H_2 \right) \tag{3.2}$$

   Note however, that even without optical scattering, the two subsystems may interact via shared quantum degrees of freedom. A simple example of this scenario is given by a two-mode resonator (such as a ring-resonator) with an atom that interacts with both optical modes, but in which there is no direct scattering between the modes.

2. The *series* product (cf. Figure 3.1(b)) describes a configuration in which two systems $Q_j = (\mathbf{S}_j, \mathbf{L}_j, H_j)$, $j = 1, 2$ possessing an equal number of channels n are connected in such a manner that all output channels of Q_1 are fed into the corresponding input channels of Q_2. The resulting system is then given by

$$(\mathbf{S}_2, \mathbf{L}_2, H_2) \lhd (\mathbf{S}_1, \mathbf{L}_1, H_1) \equiv \left(\mathbf{S}_2\mathbf{S}_1, \mathbf{L}_2 + \mathbf{S}_2\mathbf{L}_1, H_1 + H_2 + \Im\left\{ \mathbf{L}_2^\dagger \mathbf{S}_2 \mathbf{L}_1 \right\} \right), \tag{3.3}$$

   where we define the imaginary part of an operator as $\Im\{A\} \equiv \frac{A - A^\dagger}{2i}$.

To make the network operations complete, one additional rule is required: The *feedback* operation (cf. Figure 3.1(c)) describes the case where the k-th output channel of a system with $n \geq 2$ channels is fed back into the l-th input channel. The result is a system with $n - 1$ channels:

$$[(\mathbf{S}, \mathbf{L}, H)]_{k\to l} \equiv \left(\tilde{\mathbf{S}}, \tilde{\mathbf{L}}, \tilde{H} \right) \tag{3.4}$$

Formulae for the resulting parameter triplet are provided in 3.4.

Note that the series product can be expressed in terms of the concatenation and feedback operations (*e*.g., for two components with $n = 1$ we have $Q_2 \lhd Q_1 = [Q_1 \boxplus Q_2]_{1\to 2}$), and consequently, the latter two operations are sufficient to perform all network calculations. However, the series product is a useful shorthand and allows for a more intuitive network expression.

For use in the following we define the identity system with n channels

$$\mathbb{1}_n \equiv (\mathbf{1}_n, \mathbf{0}, 0)\,, \tag{3.5}$$

where $\mathbf{1}_n = (\delta_{kl})_{k,l=1}^n$ is the identity matrix in n dimensions, as well as the channel permuting system

$$P_\sigma \equiv (\mathbf{P}_\sigma, \mathbf{0}, 0)\,, \tag{3.6}$$

where the permutation matrix is defined by $\mathbf{P}_\sigma \equiv \left(\delta_{k,\sigma(l)}\right)_{k,l=1}^n$. This definition ensures that $P_{\sigma_2} \triangleleft P_{\sigma_1} = P_{\sigma_2 \circ \sigma_1}$.

3.1.2. The QHDL syntax

QHDL is a subset of structural VHDL [89], which we will use as a formal syntax for specifying photonic circuits in terms of interconnections among referenced quantum input-output components. These components can themselves represent composite networks of subcomponents, facilitating hierarchical approaches to photonic circuit design. It is useful to start with a set of basic components such as beamsplitters and phase-shifts, as well as linear and non-linear cavity models with one or more coupling mirrors[1], which can be collected in a shared library file. The set of such *primitive* components within a QHDL software environment can of course be extended at any time.

Within the context of a single QHDL file, the exact physical model (parameter triplet) of any referenced component is left unspecified except for its external ports and parametric dependencies. This approach allows the circuit designer to operate at a high level of abstraction, facilitating last-minute substitution of alternative physical component models (including effective models with reduced simulation complexity) into a given interconnection topology.

In the following section, we will introduce the QHDL syntax by means of a very simple circuit that realizes a Mach-Zehnder interferometer.

A QHDL file begins with the *entity* declaration, which defines the abstract interface of the circuit being specified: it specifies a list of named input and output ports (of the overall circuit), which are required in order for the circuit itself to be callable as a composite QHDL component, as well as any numeric parameters required for physical modeling. Note that we require that all input ports appear before all output ports.

---

[1] It is important to note that the Gough-James circuit algebra cannot be used to build dynamical systems from static components, *e.g.*, it cannot create the Fock space and operator algebra for an optical resonator mode as an automatic result of cascading beamsplitters and phase-shifts in the configuration of a ring cavity. All such dynamic components therefore must be implemented as primitive $(\mathbf{S}, \mathbf{L}, H)$ models.

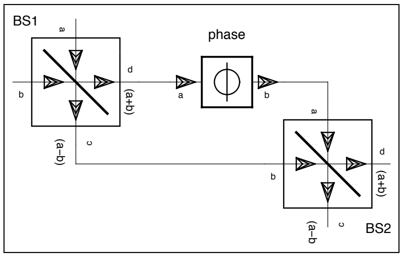


Figure 3.2.: A basic Mach-Zehnder setup.

Listing 3.1: Entity declaration

```
entity Mach_Zehnder is
   generic (phi_mz: real := 0);
   port (In1, VacIn: in fieldmode; Out1, Out2: out fieldmode);
end Mach_Zehnder;
```

For this entity we must then have one or more *architecture* declarations in the same QHDL file. These provide alternative ways of realizing the internal structure of the circuit. The architecture declaration consists of a head which specifies the interfaces of all *components* used in the architecture body and all internal *signals*. The component declarations are very similar to the entity declaration- they serve to establish an interface for each subcomponent.

Listing 3.2: Architecture head

```
architecture structure_MZ of Mach_Zehnder is
   component beamsplitter
      port (a, b: in fieldmode; c, d: out fieldmode);
   end component beamsplitter;

   component phase
      generic (phi: real);
      port (a: in fieldmode; b: out fieldmode);
   end component phase;

   signal bs1_phase, bs1_bs2, phase_bs2: fieldmode;
```

The architecture body then consists of a series of *instance* assignments for each occurrence of any of the previously specified component types. Each instance assignment specifies the relationship between the component-instance parameters and the entity parameters. In addition, it specifies a port map detailing how the component-instance is connected to the internal signals or the external ports.

Listing 3.3: Architecture body

```
begin
   BS1: beamsplitter
```

```
        port map (a => In1, b => VacIn, c => bs1_bs2, d => bs1_phase);
    phase: phase
        generic map (phi => phi_mz);
        port map (a => bs1_phase, b => phase_bs2);
    BS2: beamsplitter
        port map (a => phase_bs2, b => bs1_bs2, c => Out1, d => Out2);
end structure_MZ;
```

In the port map, each internal component port is assigned to either an entity port or a signal. Any instance *in* (*out*) port must be connected either to an entity *in* (*out*) port or to a signal that is connected to another instance's *out* (*in*) port.

Listing 3.4: Port map statement

```
        port map (a => phase_bs2, b => bs1_bs2, c => Out1, d => Out2);
```

Each signal therefore connects exactly two ports: one instance input and one instance output *or* one instance input (output) and an entity input (output).

3.1.3. Parsing a network

Here we present a simple algorithm to parse a general network into a circuit expression. We assume that the QHDL file has been preprocessed such that we have the lists of ports, components, instances, signals and port mappings in native data structures accessible to our algorithm.

1. We denote the list of internal signals by S. For each instance assignment $j = 1, 2 \dots N$ in the architecture body:
    - Generate the network triplet $Q_j = (\mathbf{S}_j, \mathbf{L}_j, H_j)$ with the correct parametrization as specified in the `generic map` statement.
    - Generate the correctly ordered[2] list of input port names I_j and the correctly ordered list of output port names O_j where each portname is entry is of the form *instance-name:port-name*.
2. Concatenate all triplets $Q = Q_1 \boxplus Q_2 \boxplus \cdots \boxplus Q_N$ and similarly concatenate the input and output port lists $I = I_1 + I_2 + \cdots + I_N$ and $O = O_1 + O_2 + \cdots + O_N$
3. For each internal signal $s \in S$ concatenate the full circuit triplet Q with a single channel identity system $\mathbb{1}_1$ resulting in $Q_f^{(0)} = Q \boxplus \mathbb{1}_{|S|}$,

---

[2] As defined via the the component declaration in the architecture head.

4. Now, each element in the full list of output ports O corresponds to an entry of the form *instance-name:port-name*. Make copies of $O' = O$ and $S' = S$ and iterate over all output ports in the following fashion:

   If the output port is connected to a global output (i.e. an entity output port), continue to the next entry.

   If the output port is connected to the j-th signal in the *current* signal list S', let k be the index of the output port in the *current* output port O' list and update the model triplet $Q_f^{(n)} \to Q_f^{(n+1)} = [Q_f^{(n)}]_{k\to M+j}$, where $M = |O'|$ is the length of the current output port list. Then, remove the k-th entry from O', and the j-th entry of S'.

5. Now, let $M_f = |S'|$ and iterate over a copy of the input port list $I' = I$ and a new copy of the signal list $S'' = S$:

   If the input port is connected to a global input (i.e. an entity input port), continue to the next entry.

   If the input port is connected to the j-th entry of S'' update $Q_f^{(n)} \to Q_f^{(n+1)} = [Q_f^{(n)}]_{M_f+j\to k}$ where k is the index of the current port in I'. Then, remove the k-th entry of I' and the j-th entry of S''.

6. By construction, the only remaining ports of our resulting triplet $Q_f^{\mathrm{r}es}$ lead to global/entity ports. Iterating over O' and the list of entity output ports O_E, construct a suitable permutation σ_{out} that maps every output port index from O' to the correct index of the entity output port. In a similar fashion, iterate over I' and the list of entity input ports I_E to generate a permutation $\sigma_{\mathrm{in}}^{-1}$, mapping the indices from I' to the correct indices of the entity input ports within I_E. Then, invert this permutation $\sigma_{\mathrm{in}} = (\sigma_{\mathrm{in}}^{-1})^{-1}$ to obtain a mapping from I_E to I'. Finally, the model triplet for the circuit is given by

$$Q_{\mathrm{f}inal} = P_{\sigma_{out}} \lhd Q_f^{\mathrm{r}es} \lhd P_{\sigma_{\mathrm{in}}}.$$

If one is interested in working with the actual network expressions as opposed to the more concrete level of the actual Hilbert space operators, there exist other, more complex, approaches to parsing a network, which directly yield simpler overall network expression. Combined with a sufficiently sophisticated set of circuit expression simplification rules, the above algorithm works just as well.

3.1.4. The QHDL workflow

The circuit design workflow relies heavily on symbolic computer algebra methods. Using symbolic algebra, rather than working with numerical matrix representations of all the operators appearing in the component parameter triplets, makes it possible to view the overall circuit $(\mathbf{S}, \mathbf{L}, H)$ in analytic form. It also allows the designer to defer choosing the values of numerical parameters, which could be convenient for optimization scenarios, as well as details such as the upper photon-number limits to use for truncated Fock spaces in numerical simulations.

In fact we can define our own algebraic types, operations and simplification rules not just for Hilbert space operators and scalar coefficients, but also for circuit algebra components. This approach enables us to extend the hierarchical design principle even to our compiled QHDL component library, as will become clear in the following outline of the modeling workflow:

1. **Circuit design** In step 1, we visually compose the circuit using a schematic capture tool and then export to QHDL[3] or directly describe the circuit in text-based QHDL. Since QHDL describes the connections between *functional* entities, it is not necessary at this stage to specify how referenced components are implemented.

2. **Component model specification** The QHDL file is then parsed to generate the circuit expression in which referenced components appear as symbols. This expression is stored in a library file along with information about model parameters and the component names of the referenced subcomponents. Note that a library file can be treated as a standalone entity for future circuit designs. When this file is imported at runtime, the referenced subcomponent models are dynamically loaded from their respective library files. Now, the full $(\mathbf{S}, \mathbf{L}, H)$ parameter triplet can be generated by explicitly evaluating the circuit algebra operations. By means of the symbolic operator algebra, the final operator matrices and the Hamiltonian are still in fully symbolic form, which can be used to generate the quantum master equation or an appropriate stochastic differential equation in symbolic form. This allows for the application of analytical model reduction techniques before turning to purely numerical methods.

3. **Numerical simulation** Define all scalar model parameters and (truncated) Hilbert space dimensions, and compute the behavior of the circuit.

In Table 3.1 we list the necessary software tools to implement the QHDL circuit design workflow. We plan to publicly release our custom tools in the near future.

---

[3] In our case we have modified the VHDL exporting functionality of the gEDA toolsuite to generate QHDL.

3.2. An example of the QHDL workflow

In this section, we present a detailed example applying the QHDL workflow to the design, analysis, and simulation of an all-optical $\overline{SR}$-latch as recently proposed in [69, 68]. The elementary component models $\{(\mathbf{S}_j, \mathbf{L}_j, H_j)\}$ required for this circuit are the following:

1. **Beamsplitters**
$$Q_{BS}(\theta) \equiv \left(\begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}, \mathbf{0}, 0 \right),$$
2. **P**hase-delays $U(\phi) \equiv (e^{i\phi}, 0, 0)$
3. **C**oherent displacements $W(\alpha) \equiv (1, \alpha, 0)$ which models a laser source outputting a coherent field with amplitude $\alpha \in \mathbb{C}$,
4. **K**err-nonlinear cavity (here a unidirectional ring cavity with two input/output ports)
$$Q_K(\Delta, \chi, \{\kappa_j\}) = \left(\mathbb{1}_2, \begin{pmatrix} \sqrt{\kappa_1} a \\ \sqrt{\kappa_2} a \end{pmatrix}, \Delta a^\dagger a + \chi a^\dagger a^\dagger a a \right)$$

As the circuits we discuss here are meant to be used as logical gates in larger circuits, we need not include the laser sources in our circuit schematics. Instead, they can easily be added at the level of the circuit algebra by feeding a concatenated block of laser displacements (sources) into the full network $Q_{\text{with input}} = Q \lhd (W_{\alpha 1} \boxplus W_{\alpha 2} \boxplus \cdots \boxplus W_{\alpha n})$.

3.2.1. The two-cavity pseudo-NAND-latch

We have recently proposed [69] several different optical circuits to realize three classical logic gates: an AND-gate, a NOT-gate with integrated fanout of two and a combined (but imperfect) NAND-gate (Figure 3.3(a)), which in the following we will call pseudo-NAND gate as it only works properly when at any given time at least one input is in the 'on' state. The first two of these gates used in sequence also realizes a NAND gate, but the advantage of the pseudo-NAND is that it requires only a single Kerr-nonlinear cavity component. The QHDL workflow can be readily applied to design the pseudo-NAND circuit and automatically generate the circuit expression in terms of its components[4]:

$$\begin{aligned} &\big(\mathbb{1}_1 \boxplus \big((\mathbb{1}_1 \boxplus ((\Phi \boxplus \mathbb{1}_1) \lhd B_2)) \lhd P_{(1\,3\,2)} \lhd (K \boxplus \mathbb{1}_1)\big)\big) \\ &\qquad \lhd \big(B_1 \boxplus \big(P_{(2\,1)} \lhd (W \boxplus \mathbb{1}_1)\big)\big) \end{aligned} \tag{3.7}$$

---

[4] Here we represent a permutation $\sigma = \begin{pmatrix} 1 & 2 & \dots & n \\ \sigma(1) & \sigma(2) & \dots & \sigma(n) \end{pmatrix}$ just by its image tuple $(\sigma(1)\, \sigma(2)\, \dots\, \sigma(n))$.

Table 3.1.: List of software components necessary to realize our QHDL-workflow.

Requirement	**Our solution**	**Alternatives**
Graphical **schematic capture** tool with VHDL/QHDL export capabilities	gschem and gnetlist from the gEDA suite [119]	Graphical design tool from system modeling environments / modeling languages, such as Modelica [120]
QHDL-Parser that computes the circuit expression	A custom parser written in Python using the open source PLY [5] package	A parser for computing a circuit expression from the Modelica specification, written in, e.g. *Mathematica*
Symbolic computer algebra system with support for: non-commutative operator algebra, commutative scalar algebra coefficient, operator-valued matrix algebra and the Gough-James circuit algebra	A custom computer algebra system written in Python [122] and interfacing with SymPy [107] for the scalar coefficient algebra	*Mathematica* [139] plus an implementation of the Gough-James circuit algebra
Numerical backend to convert symbolic operator expressions into matrices and simulate the system	Custom algorithms for solving the Master equation as well as quantum stochastic differential equations implemented in Python and C using optimized numerical libraries for linear algebra [3, 54].	The Quantum Optics Toolbox [110] for MATLAB [72] or similar library for modeling the dynamics of open quantum systems, such as QuTIP [56]

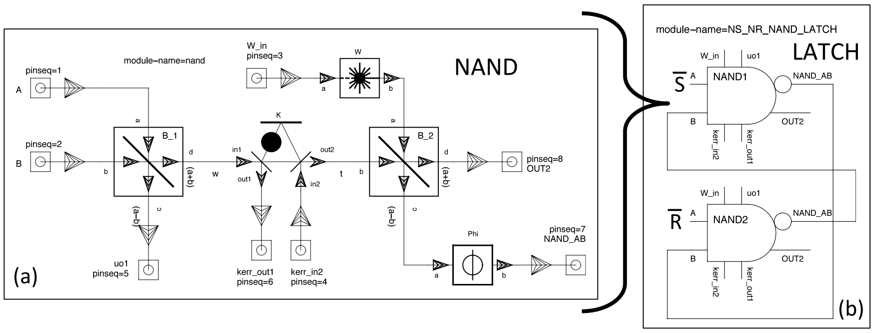

module–name=nand
NAND
pinseq=1
A
pinseq=2
B
W_in
pinseq=3
W
a
b
K
B_1
d
b
c
(a+b)
(a−b)
in1
out1
out2
in2
w
t
B_2
pinseq=8
OUT2
Phi
pinseq=7
NAND_AB
uo1
pinseq=5
kerr_out1
pinseq=6
kerr_in2
pinseq=4
(a)
module–name=NS_NR_NAND_LATCH
LATCH
W_in
uo1
$\overline{\mathrm{S}}$
A
NAND1
NAND_AB
B
OUT2
kerr_in2
kerr_out1
$\overline{\mathrm{R}}$
NAND2
(b)

Figure 3.3.: Pseudo-NAND circuit schematic (a) as created with gschem and its device symbol embedded as a component in a SR-NAND-latch circuit (b).

where the beamsplitter symbols are defined by $B_1 \equiv Q_{BS}(\frac{\pi}{4})$, $B_2 \equiv Q_{BS}(\theta)$, the output correction phase is $\Phi \equiv U(\phi)$, the constant coherent displacement component is $W \equiv W(\beta)$ and the Kerr cavity is given by $K \equiv Q_K(\Delta, \chi, \kappa_1 = \kappa_2 = \kappa)$. The network expression (3.7) looks complicated but can be verified easily by comparing its visual representation (Fig. 3.4(a))[5] with the original circuit schematic. Moreover, since the scattering matrix of K is in block-diagonal form, it is possible to decompose the cavity component $K = K_1 \boxplus K_2$, where the Hamiltonian of K can be assigned to either of the two blocks. Upon substituting this decomposable form into the expression, the automatic expression simplification built into our circuit algebra implementation yields the following form:

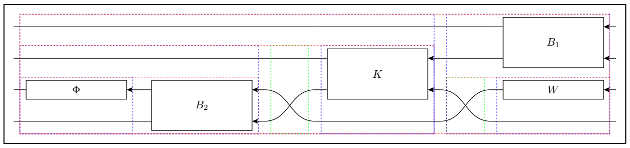
$$\Big\{ (\mathbb{1}_1 \boxplus K_1) \lhd B_1 \Big\} \boxplus \Big\{ (\Phi \boxplus \mathbb{1}_1) \lhd B_2 \lhd (W \boxplus K_2) \Big\}, \tag{3.8}$$

which is visually represented in Figure 3.4(b). The numerical model parameters as given in [69] are $\theta = 0.891$, $\chi = -5/6$, $\Delta = 50$, $\kappa = 25$, $\phi = 2.546$ and the auxiliary constant input amplitude is given by $\beta = -34.289 - 11.909i$. The coherent input amplitudes corresponding to the logical signals 'on' and 'off' are then given by $\alpha = 22.6274$ and 0, respectively.

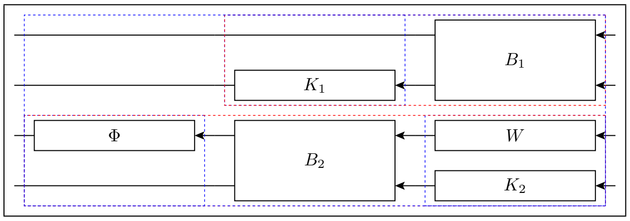
As in classical circuit theory, two NAND-gates in a mutual feedback configuration, as shown in Figure 3.3(b), can be used to realize a latch with inverted inputs $\overline{\mathrm{S}}$ and $\overline{\mathrm{R}}$. A latch features controllable bistable behavior and thus realizes a single-bit memory unit. It has two inputs: S(ET) and R(ESET), which can be activated individually to control the internal logical state to 'on' or 'off', respectively. Ideally, when both S and R are 'off' (HOLD-condition), the internal state remains stable. In practice, quantum fluctuations and noisy inputs lead to spontaneous switching between the two internal states. One of the design goals is thus to decrease the rate at which this spontaneous switching occurs. The QHDL code as produced by gnetlist [119] (slightly edited to be more concise) can

---

[5] These visualizations were automatically generated using another software tool we have implemented.



(a) Circuit expression (3.7) as generated by the QHDL-Parser



(b) Simplified expression (3.8)

Figure 3.4.: Pseudo-NAND circuit expression visualizations. As can easily be verified visually, the simplified expression follows from decomposing $K = K_1 \boxplus K_2$ and 'pulling' K_2 down into the fourth row. These expression simplifications are automatically performed by our symbolic circuit algebra software.

be found in Listing 3.5 and the circuit component library file generated by the QHDL-parser is presented in Listing 3.6 in 3.5. Substituting the individual component models into the circuit expression yields the full triplet $(\mathbf{S}_0, \mathbf{L}_0, H_0)$ for the latch. Finally, after feeding in the coherent input signals $\overline{\mathrm{S}}$ and $\overline{\mathrm{R}}$ into their respective ports: $(\mathbf{S}, \mathbf{L}, H) = (\mathbf{S}_0, \mathbf{L}_0, H_0) \lhd (W(\overline{\mathrm{S}}) \boxplus \mathbb{1}_2 \boxplus W(\overline{\mathrm{R}}) \boxplus \mathbb{1}_2)$ the parameters assume the following form

$$\mathbf{S} = \left(\begin{array}{c|c} \mathbf{S}_1 & \mathbf{0} \\ \hline \mathbf{0} & \mathbf{S}_2 \end{array}\right), \text{ where } \mathbf{S}_1 = \mathbf{S}_2 = \begin{pmatrix} \frac{1}{\sqrt{2}} & -\frac{\cos\theta e^{i\phi}}{\sqrt{2}} & \frac{\sin\theta e^{i\phi}}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & \frac{\cos\theta e^{i\phi}}{\sqrt{2}} & -\frac{\sin\theta e^{i\phi}}{\sqrt{2}} \\ 0 & \sin\theta & \cos\theta \end{pmatrix}, \tag{3.9}$$

$$\mathbf{L} = \begin{pmatrix} \sqrt{\frac{\kappa}{2}} \sin\theta e^{i\phi}\; b + \frac{\overline{\mathrm{S}}}{\sqrt{2}} - \frac{\beta}{\sqrt{2}} \cos\theta e^{i\phi} \\ \sqrt{\kappa}\; a - \sqrt{\frac{\kappa}{2}} \sin\theta e^{i\phi}\; b + \frac{\overline{\mathrm{S}}}{\sqrt{2}} + \frac{\beta}{\sqrt{2}} \cos\theta e^{i\phi} \\ \sqrt{\kappa} \cos\theta\; b + \beta \sin\theta \\ \sqrt{\frac{\kappa}{2}} \sin\theta e^{i\phi}\; a + \frac{\overline{\mathrm{R}}}{\sqrt{2}} - \frac{\beta}{\sqrt{2}} \cos\theta e^{i\phi} \\ \sqrt{\kappa}\; b - \sqrt{\frac{\kappa 2}{\sin}} \theta e^{i\phi}\; a + \frac{\overline{\mathrm{R}}}{\sqrt{2}} + \frac{\beta}{\sqrt{2}} \cos\theta e^{i\phi} \\ \sqrt{\kappa} \cos\theta\; a + \beta \sin\theta \end{pmatrix}, \tag{3.10}$$

$$\begin{aligned} H = {} & \Delta \left(a^\dagger a + b^\dagger b\right) + \chi \left(a^\dagger a^\dagger a a + b^\dagger b^\dagger b b\right) - \frac{\kappa}{\sqrt{2}} \sin\theta \sin\phi \left(a b^\dagger + a^\dagger b\right) \\ & + \frac{\sqrt{2\kappa}}{4} i \left[\left(\overline{\mathrm{S}}^* + \beta^* \cos\theta e^{-i\phi}\right) a - \text{h.c.}\right] + \frac{\sqrt{2\kappa}}{4} i \left[\left(\overline{\mathrm{R}}^* + \beta^* \cos\theta e^{-i\phi}\right) b - \text{h.c.}\right]. \end{aligned} \tag{3.11}$$

Listing 3.5: QHDL source for the pseudo-NAND latch.

```
--pseudo-NAND latch
ENTITY NS_NR_NAND_LATCH IS
   PORT (NS, W1, kerr2_extra, NR, W2, kerr1_extra : in fieldmode;
        BS1_1_out, kerr1_out2, OUT2_2, BS1_2_out, kerr2_out2, OUT2_1 : out fieldmode);
END NS_NR_NAND_LATCH;

ARCHITECTURE netlist OF NS_NR_NAND_LATCH IS
   COMPONENT nand
      PORT (A, B, W_in, kerr_in2 : in fieldmode;
           uo1, kerr_out1, NAND_AB, OUT2 : out fieldmode);
   END COMPONENT;

   SIGNAL FB12, FB21 : fieldmode; -- feedback signals

BEGIN
   NAND2 : nand
   PORT MAP (
      A => NR, B => FB12, W_in => W2, kerr_in2 => kerr2_extra,
      uo1 => BS1_2_out, kerr_out1 => kerr2_out2, NAND_AB => FB21, OUT2 => OUT2_2);

   NAND1 : nand
   PORT MAP (
      A => NS, B => FB21, W_in => W1, kerr_in2 => kerr1_extra,
      uo1 => BS1_2_out, kerr_out1 => kerr1_out2, NAND_AB => FB12, OUT2 => OUT2_1);
END netlist;
```

Due to the symmetry of the underlying circuit model, the model parameters are invariant under exchange of the two pseudo-NAND gates, which corresponds to simultaneously exchanging $\overline{\mathrm{S}} \leftrightarrow \overline{\mathrm{R}}$, $(a, a^\dagger) \leftrightarrow (b, b^\dagger)$, $(L_1, L_2, L_3) \leftrightarrow (L_4, L_5, L_6)$ and $\mathbf{S}_1 \leftrightarrow \mathbf{S}_2$. This symmetry suggests that the most likely candidates for the internal logical states 'on' and 'off' correspond to the case where one internal cavity mode is in a high power state and the other one in a low power state and the opposite case, obtained by exchanging the cavities states. This is indeed the case, and in fact it follows from the basic way in which we have designed our pseudo-NAND gate; 'on' $\Leftrightarrow$ {NAND1 cavity power is low, NAND2 cavity power is high} and 'off' $\Leftrightarrow$ (NAND1 cavity power is high, NAND2 cavity power is low).

To understand our model's dynamic behavior we turn to numerical methods. The simulation of this model is carried out by representing the operators as numerical matrices in a truncated product basis of Fock-states of total dimension $N^2 = 75^2 = 5625$[6]. We carried out a large number of quantum jump trajectory simulations [110, 50] with the following sequence of alternating input conditions: 0.5 time units of SET, 5 units of HOLD, 0.5 units of RESET, 5 units of HOLD (repeated twice). The upper plot in Figure 3.5 presents a typical simulated trace where the system is subjected to this sequence of input conditions. We generally find that the SET and RESET input conditions successfully drive the system into the desired cavity states, while the cavities remain in their states during the HOLD condition. Although a simulation of the full master equation is feasible using current HPC hardware and sparse matrix storage [69], quantum jump simulations exhibit the inherently bistable nature of our synthesized latch more clearly.

---

[6] *I.e.*, each individual cavity basis is given by $\{|0\rangle, |1\rangle, \dots |N-1\rangle\}$.

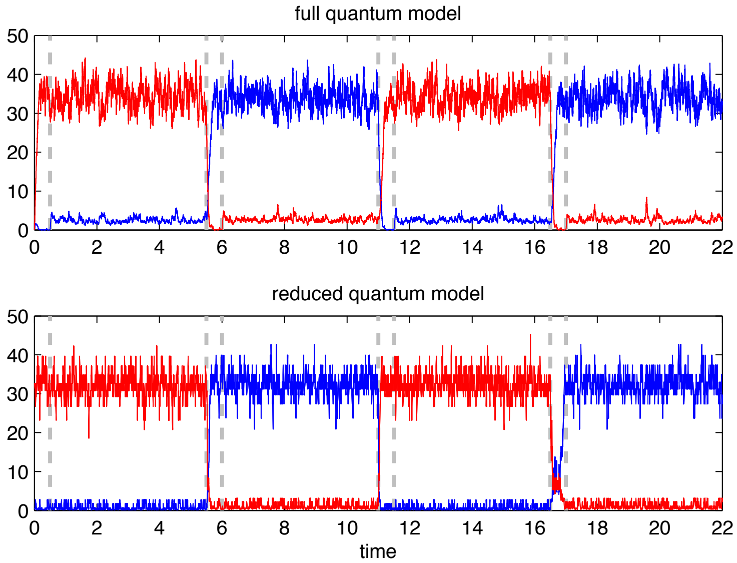


Figure 3.5.: Simulated input sequence for the full pseudo-NAND latch model (upper) and our reduced model (lower). The red trace (lighter, in grayscale) is given by $\langle a^\dagger a \rangle$ and the blue trace presents $\langle b^\dagger b \rangle$. The SET and RESET input conditions, marked by the short intervals between dashed vertical lines, induce transitions to their respective target latch states: 'on' corresponds to the a-mode being in a high photon number state, while 'off' corresponds to a high photon number in mode b.

3.2.2. Model reduction in the SLH context

As the latch could readily be used as a component in more complex circuits, such as flip-flops or even quantum memories [57, 58], it would be highly desirable to reduce the Hilbert space dimension N^2 required to represent it. Since we are working with quantum circuit models, we are ultimately limited by exponential scaling of the state space with the number of components (although it may be possible to develop efficient simulation procedures when components are only weakly entangled, as should be the case in ultra-low power classical signal processing). However, there is clearly much to be gained by developing accurate model reduction procedures that allow us to replace high-dimensional *a*b initio models for components within a circuit by much lower-dimensional effective models. Such model-reduction strategies could presumably be applied hierarchically. As in the classical theory of signals-and-systems there are many potential strategies for dimensional reduction of quantum input-output models. For the case of components or sub-circuits whose input-output behavior admits a simplified description as certain parameter ratios become large, a recently-derived limit theorem [12, 13, 44] can be utilized as demonstrated in [57, 58]. Here we describe an empirical approach similar to the classical strategy of approximating Markov chains [61], which utilizes numerical simulation and statistical analysis to derive a reduced $(\mathbf{S}, \mathbf{L}, H)$ model for the pseudo-NAND latch (for which no simplifying parameter limits are known).

Our approach is based on the assumption that the device state can be inferred with reasonable accuracy from a small number of observables. We can then construct a dynamic model just in terms of these parameters [82]. By generating many quantum jump trajectories for the full $(\mathbf{S}, \mathbf{L}, H)$ model, we generated many time series for the expectation values of the cavity field photon numbers $\left\langle a^\dagger a \right\rangle$ and $\left\langle b^\dagger b \right\rangle$ under the three valid input conditions $\{\mathrm{HOLD}, \mathrm{SET}, \mathrm{RESET}\}$. We can now coarse-grain the 2-dimensional space of expectation values and associate an internal model state $i \in \{1, 2, \dots, M\}$ with each bin that is actually visited during the trajectory simulations, but due to the high correlations between the cavity photon numbers, we are actually able to obtain good results by performing this coarse-graining or 'binning' procedure in terms of the single quantity

$$D \equiv \left\langle a^\dagger a \right\rangle - \left\langle b^\dagger b \right\rangle,$$

implying that within the two-dimensional configuration space our system always stays fairly close to a one-dimensional submanifold.

By analyzing the observed transitions between these reduced states for each input condition $\xi \in \{\mathrm{HOLD}, \mathrm{SET}, \mathrm{RESET}\}$, we calculate an empirical estimate $\hat{\mathbf{P}}^{(\xi)} = (p_{ij}^{(\xi)})_{i,j=1}^M$ of the conditional transition probabilities $p_{ij}^{(\xi)} = P(x_{n+1} = j | x_n = i,\, \xi)$ and thus model the system in terms of a discrete time Markov chain with a set of conditional transition probabilities for each particular input condition ξ. The time step δt of the discrete Markov chain corresponds to the interval at which we sampled our original continuous-time system. We now wish to get back to a description that is compatible with our $(\mathbf{S}, \mathbf{L}, H)$ formalism. In the following we briefly outline a procedure to do this: For a temporally homogeneous Markov jump process with an even number of states $i \in \{1, 2, \dots, M\}$ and transition rate matrix[7] $\mathbf{Q} = (\gamma_{ij})_{i,j=1}^M$ we can define a K-channel $(\mathbf{S}_0, \mathbf{L}_0, H_0)$ model with states corresponding directly to the Markov process states $\{|1\rangle, |2\rangle, \dots, |M\rangle\}$ via

---

[7] $\mathbf{Q}$ is also often referred to as the *generator matrix*.

$$\mathbf{S}_0 = \mathbf{1}_K, \tag{3.12}$$

$$\mathbf{L}_0 = \left(\sqrt{\gamma_{i_1 j_1}}\, |j_1\rangle \langle i_1|\,,\ \dots, \sqrt{\gamma_{i_K j_K}}\, |j_K\rangle \langle i_K|\right)^T \tag{3.13}$$

$$H_0 = 0, \tag{3.14}$$

where the components of $\mathbf{L}_0$ drive transitions $\{i_k \to j_k, k = 1, 2 \dots K\}$ and K is given by the number of positive transition rates $\gamma_{ij} > 0$. By construction, as one may verify by writing down the master equation, this system always collapses into a purely classical mixture of the coarse-grained states. Equivalently, in a quantum jump trajectory simulation, after the first quantum jump, the state is always given by a single such state. In fact, in such a trajectory simulation, this system behaves exactly like the original Markov jump process. Neglecting for now that our original model has three different input conditions $\xi \in \{\text{HOLD}, \text{SET}, \text{RESET}\}$ and thus three different conditional transition matrices $\hat{\mathbf{P}}^{(\xi)}$, we first discuss how to move from the discrete time Markov chain model to a continuous time Markov jump process. Rephrasing this question, we can ask the following: is there a Markov jump process with conditional transition probability matrix $\mathbf{P}(t)$ that 'looks like' our Markov chain when stroboscopically probed at fixed time intervals δt? If our Markov chain has transition matrix $\hat{\mathbf{P}}$, then we need to determine a generator matrix $\mathbf{Q}$ such that

$$\hat{\mathbf{P}} = \mathbf{P}(t = \delta t) \equiv e^{\delta t \mathbf{Q}} \approx \mathbf{1} + \delta t \mathbf{Q} + O(\delta t^2). \tag{3.15}$$

If our sample interval δt is sufficiently small, we may define

$$\hat{\mathbf{Q}} \equiv \frac{1}{\delta t}(\hat{\mathbf{P}} - \mathbf{1}) \tag{3.16}$$

as an approximation to the conditional rate matrix. We now carry out the procedure outlined above to create a model $(\mathbf{S}_0, \mathbf{L}_0, H_0)$ that realizes the HOLD condition. The transition rates of the HOLD condition alone lead to a system that has two bistable clusters of states with low state indices and high state indices, respectively.

To account for the input-controlled switching in the SET and RESET conditions, we extend our model by concatenating it with a second model that explicitly includes the input fields $(\mathbf{S}_{SR}, \mathbf{L}_{SR}, H_{SR}) = (\mathbf{S}_1, \mathbf{L}_1, H_1) \lhd (W(\overline{S}) \boxplus W(\overline{R}) \boxplus \mathbb{1}_2)$. Hence, in the SET and RESET conditions, the HOLD transitions continue, but we drive further transitions through this additional component. Here, $(\mathbf{S}_1, \mathbf{L}_1, H_1)$ is given by

$$\mathbf{S}_1 = \mathbf{1}_4 - \begin{pmatrix} \Sigma_S^\dagger \Sigma_S & 0 & -\Sigma_S^\dagger & 0 \\ 0 & \Sigma_R^\dagger \Sigma_R & 0 & -\Sigma_R^\dagger \\ -\Sigma_S & 0 & \Sigma_S \Sigma_S^\dagger & 0 \\ 0 & -\Sigma_R & 0 & \Sigma_R \Sigma_R^\dagger \end{pmatrix} \tag{3.17}$$

$$\mathbf{L}_1 = -\alpha(1 - \Sigma_S^\dagger \Sigma_S,\ 1 - \Sigma_R^\dagger \Sigma_R,\ \Sigma_S,\ \Sigma_R)^T \tag{3.18}$$

$$H_1 = 0, \tag{3.19}$$

where the 'drift' operators Σ_S and Σ_R are defined by

$$\Sigma_S \equiv |M-4\rangle \langle M-1| + |M-6\rangle \langle M-3| + \cdots + |M/2-1\rangle \langle M/2+2|\,,$$
$$\Sigma_R \equiv |5\rangle \langle 2| + |7\rangle \langle 4| + \cdots + |M/2+2\rangle \langle M/2-1|\,.$$

For our simulation we chose $M = 38 \ll N^2$, but the general ansatz works for a range of different $M = 4k+2$ with sufficiently large k. The drift operators satisfy $\Sigma_{R/S}^2 = 0$, as well as the projection relations $(\Sigma_{R/S}^\dagger \Sigma_{R/S})^2 = \Sigma_{R/S}^\dagger \Sigma_{R/S}$, $(\Sigma_{R/S}^\dagger \Sigma_{R/S})^2 = \Sigma_{R/S}^\dagger \Sigma_{R/S}$ and $\Sigma_{R/S} \Sigma_{R/S}^\dagger \Sigma_{R/S} = \Sigma_{R/S}$. These relations suffice to show that $\mathbf{S}_1$ as defined above is indeed unitary. To make sense of the effect of this extension to our model, consider now what happens for the different input conditions: In the HOLD condition $\overline{S} = \overline{R} = \alpha$ the input fields cancel out all elements of the coupling vector $\mathbf{L}_1$ and we have $(\mathbf{S}_{SR}, \mathbf{L}_{SR}, H_{SR}) = (\mathbf{S}_1, \mathbf{0}, 0)$, i.e. the transition dynamics of our full system $(\mathbf{S}_{SR}, \mathbf{L}_{SR}, H_{SR}) \boxplus (\mathbf{S}_0, \mathbf{L}_0, H_0)$ are simply given by those of $(\mathbf{S}_0, \mathbf{L}_0, H_0)$ alone.

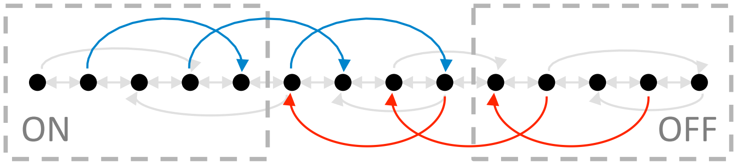


Figure 3.6.: Here we schematically visualize the state space and the transitions of the reduced model. The SET transitions (red, from right to left) introduce a drift that drives the system to the states on the left, corresponding to the logical 'ON' state of the latch. The RESET transitions (blue, from left to right) have the opposite effect. The HOLD transitions (depicted in grey) are always active, but in the absence of additional SET and RESET transitions only very rarely lead to a switch of the logical latch state.

In the SET condition, however, we have $\overline{S} = 0, \overline{R} = \alpha$ and thus

$$(\mathbf{S}_{SR}, \mathbf{L}_{SR}, H_{SR}) = \left(\mathbf{S}_1, (-\alpha(1 - \Sigma_S^\dagger \Sigma_S),\ 0,\ -\alpha\Sigma_S, 0)^T, 0\right).$$

The full system $(\mathbf{S}_{SR}, \mathbf{L}_{SR}, H_{SR}) \boxplus (\mathbf{S}_0, \mathbf{L}_0, H_0)$ now features the drift operator as an additional transition operator[8] $-\alpha\Sigma_S$ which induces transitions $\{M-1 \to M-4, M-3 \to M-6, \dots, M/2+2 \to M/2-1\}$ with constant rate $|\alpha|^2$. Together with the HOLD transitions, these lead to a drift from states with high index (corresponding to the logical 'off' state of the latch) to those with low index ('on'). On the other hand, in the RESET condition, the situation is reversed. Now the other transition operator of $\mathbf{L}_1$ is canceled out and the non-zero transition operator $-\alpha\Sigma_R$ drives transitions in the reverse direction $\{2 \to 5, 4 \to 7, \dots, M/2-1 \to M/2+2\}$, again with constant rate $|\alpha|^2$. In Figure 3.6 we visualize the transition structure of the model schematically.

Note also that we can emulate the state-dependent coherent output field(s) of the latch by concatenating a triplet $(\mathbf{S}_{out}, \mathbf{L}_{out}, H_{out})$ that re-routes bias input fields via state-dependent scattering into one or more output channels. For example we could use $(\mathbf{S}_{SR}, \mathbf{L}_{SR}, H_{SR}) \boxplus (\mathbf{S}_0, \mathbf{L}_0, H_0) \boxplus (\mathbf{S}_{out}, \mathbf{L}_{out}, 0)$ where

$$\mathbf{S}_{out} = \sum_{i=1}^{M} |i\rangle\langle i| \begin{pmatrix} e^{i\phi_{1i}}\cos\theta_i & -e^{i\phi_{1i}}\sin\theta_i \\ e^{i\phi_{2i}}\sin\theta_i & e^{i\phi_{2i}}\cos\theta_i \end{pmatrix}, \quad \mathbf{L}_{out} = \mathbf{S}_{out}(\beta', 0)^T, \tag{3.20}$$

where β' is the complex amplitude of a bias field and the parameters $\{\theta_i, \phi_{1i}, \phi_{2i}\}$ are chosen such that the outputs of $(\mathbf{S}_{out}, \mathbf{L}_{out}, 0)$ vary as desired with the internal state $|i\rangle$. Having thus created a reduced model that mimics the desired input-output behavior in $(\mathbf{S}, \mathbf{L}, H)$ form, we can use it to replace the full latch model in more complex circuits. If we had already specified a QHDL file for such a circuit, we could simply replace[9] the referenced latch component with the reduced model component. Re-parsing this modified QHDL-file would then yield a computationally more tractable model for simulations.

3.3. Conclusion

In this paper we have described the use of QHDL, a quantum hardware description language, to facilitate the analysis, design, and simulation of complex networks constructed from interconnected quantum optical components. We have also presented a parsing algorithm for obtaining quantum equations of motion from the QHDL description. QHDL can be used as the basis for a schematic capture workflow for designing quantum circuits that automates many of the conceptually challenging and computationally demanding aspects of quantum network synthesis. As QHDL inherits the hierarchical structure of VHDL, its use may facilitate the crucial development of hierarchical model reduction methods for quantum nonlinear photonics.

---

[8] The second non-zero element of $\mathbf{L}_{SR}$, which is a projection operator, does not affect the transition dynamics due to the fact that our system is never in a superposition of two states.

[9] In principle it should be possible to include the reduced model as an alternative architecture for the latch entity and to select whether or not to use it in place of the full model at compile-time using a VHDL configuration file. However this would require some enhancements to the QHDL-Parser to correctly handle the K extra (vacuum) input ports required by the reduced model to drive spontaneous transitions among the internal states.

Important future directions for QHDL research include simulation strategies for exploiting weak entanglement among components, stability analysis and design optimization of QHDL-based models [81], and the incorporation of techniques from static program analysis and formal verification to assist in the design of complex, hierarchically defined photonic components. While we have emphasized classical photonic logic [69] as a tutorial paradigm for QHDL in this paper, emerging ideas in quantum information processing and quantum sensing/metrology may provide even more compelling applications for QHDL as a convenient and extensible modeling framework.

3.4. Reduced parameters in case of signal feedback

where effective parameters are then given by [40]

$$\tilde{\mathbf{S}} = \mathbf{S}_{[\cancel{k},\cancel{l}]} + \begin{pmatrix} S_{1l} \\ S_{2l} \\ \vdots \\ S_{k-1\,l} \\ S_{k+1\,l} \\ \vdots \\ S_{nl} \end{pmatrix} (1 - S_{kl})^{-1} \begin{pmatrix} S_{k1} & S_{k2} & \cdots & S_{kl-1} & S_{kl+1} & \cdots & S_{kn} \end{pmatrix}, \tag{3.21}$$

(3.22)

$$\tilde{\mathbf{L}} = \mathbf{L}_{[\cancel{k}]} + \begin{pmatrix} S_{1l} \\ S_{2l} \\ \vdots \\ S_{k-1\,l} \\ S_{k+1\,l} \\ \vdots \\ S_{nl} \end{pmatrix} (1 - S_{kl})^{-1} L_k, \quad \tilde{H} = H + \Im\left\{ \left[\sum_{j=1}^{n} L_j^\dagger S_{jl}\right] (1 - S_{kl})^{-1} L_k \right\}. \tag{3.23}$$

Here we have written $\mathbf{S}_{[\cancel{k},\cancel{l}]}$ as a shorthand notation for the matrix $\mathbf{S}$ with the k-th row and l-th column removed and similarly $\mathbf{L}_{[\cancel{k}]}$ is the vector $\mathbf{L}$ with its k-th entry removed. These resulting parameters fulfill the conditions[10] for circuit components. Moreover, they have shown that in the case of multiple feedback loops, the result is independent of the order in which the feedback operation is applied[11].

3.5. Latch circuit library file

Listing 3.6: Python[122] source for the pseudo-NAND latch circuit library component.

```python
#!/usr/bin/env python

from qhdl_component_lib.library import retrieve_component, make_namespace_string
from qnet.qos_algebra import P_sigma, qid, FB
from sympy.core.symbol import symbols

NCHANNELS = 6
GENERIC_DEFAULT_VALUES = {}

def arch_default(name_space = '', **generic_params):
    """
    Generate a symbolic LATCH expression using the provided name_space.
    Return a 2-tuple (netlist_symbolic, var_map)
    such that calling
       >>> SLH = netlist_symbolic.substitute(var_map).evalf()
    will result in an SLH triplet, where the parameters are accessible via
       >>> SLH.S
       >>> SLH.L
       >>> SLH.H
    """
    # dictionary that stores the replacement SLH models for
    # the circuit component symbols
    var_map = {}

    # load symbolic component expressions as well
    # as actual SLH model replacements for NAND1...
    NAND1, NAND1_var_map = retrieve_component('nand', 4,
                                make_namespace_string(name_space,'NAND1'))
    var_map.update(NAND1_var_map)

    #...and NAND2
    NAND2, NAND2_var_map = retrieve_component('nand', 4,
```

---

[10] This is obvious for $\tilde{\mathbf{L}}$ and $\tilde{H}$, for a proof that $\tilde{\mathbf{S}}$ is indeed unitary see Gough and James's original paper [40].

[11] Note however that some care has to be taken with the indices of the feedback channels when permuting the feedback operation.

```
                            make_namespace_string(name_space,'NAND2'))
    var_map.update(NAND2_var_map)

    ########## symbolic circuit expression as computed by the parser
    # the '+' operator is overloaded to the concatenation-operation [+]
    # the '<<' operator is overloaded to the series-operation <|
    # P_sigma(s(0),s(1),...(s(n-1))) is a zero-based channel permutation object
    # FB(Q, k,l) is the feedback operation [Q]_{k->l} (with zero-based channel indices)
    # qid(n) is the n-channel identity
    netlist_symbolic = (P_sigma(0, 2, 3, 4, 5, 1)
                   << FB(
                      (
                         (qid(1) +
                            ((qid(3) + P_sigma(2, 0, 1))
                               << ((P_sigma(1, 2, 3, 0) << NAND2) + qid(2))))
                         << P_sigma(0, 2, 5, 6, 1, 3, 4)
                         << ((P_sigma(0, 3, 1, 2) << NAND1) + qid(3))
                      ), 6, 1)
                   << P_sigma(0, 4, 5, 3, 1, 2))

    return netlist_symbolic, var_map
```

4. Ultra-Low-Power All-Optical Computation

In this chapter we present various useful single device and circuit models for constructing components for analog or digital computation. Since I joined Hideo Mabuchi's group, there has been a shift away from typical cavity QED models towards networks of resonators featuring some optically nonlinear bulk material.

Although this turns out to be equivalent to cavity QED in the extremely nonlinear limit [70] – when the rate of interactions between two photons approaches or exceeds the linear dissipation rate – realistic assumptions of current and near-term feasible nonlinear materials and fabrication capabilities are very far from this limit.

It thus becomes an interesting question to consider what computational devices in the intermediate regime where nonlinear effects become apparent for a few tens of photons. There are some particularly interesting questions here:

1. What is the consequence of optical coherence and interference for designing circuits for computation?
2. At what energy scale is classical optical computing possible despite the effects of quantum shot noise?
3. Can even weak quantum effects provide any enhancement to the efficiency of analog optical computers?

In 2011 Hideo Mabuchi demonstrated how coherent nonlinear feedback can help reduce the spontaneous switching rate of a bi-stable system [68]. This and the work below provided some insight into the first of these questions.

In 2014 Charles Santori, Jason Pelc and Ray Beausoleil at HP Labs as well as Ryan Hamerly, Hideo Mabuchi and I derived a semi-classical coupled mode theory [98] for ultra low power photonic circuits. The 'Wigner method' lately has gained great popularity inside and outside of our group. It is based on transforming the master equation to a Wigner quasi-probability representation and approximating the resulting partial differential equation by a Fokker-Planck equation, thus allowing to stochastically simulate the mode dynamics. This result generalizes previous methods developed by Carter [16].

Extending earlier work by Hideo Mabuchi on how to implement optical sequential logic [69] we were able to simulate some fairly complex digital circuits assuming different ratios of signal levels to shot noise, or equivalently nonlinear interaction rate to linear dissipation rate. For a simple bistable Kerr-resonator we compared numerically estimated switching rates between the equilibria and found good agreement when the intra cavity photon numbers were at least $\sim 20 - 40$. For integrated photonic logic circuits operating on picosecond timescales the gate error probability could be made sufficiently small $< 10^{-18}$ at intra-resonator energy levels of around ~ 177 photons. Thus, the Wigner method has provided us with an answer to the second question posed above, at least as far as digital logic is concerned.

Our group has recently become interested in photonic schemes for neuromorphic computation, probabilistic computation and machine learning both in the low power limit and even the quantum regime. At least some neuromorphic architectures are very robust to device imperfections and noise, making them interesting candidates for optical computing. Furthermore, assuming device nonlinearity and fabrication quality continue to improve, it is interesting to consider how such circuits might behave in the quantum regime where entanglement could arise between spatially separated parts of a device. I divide these into the following broad categories:

1. *Optical Reservoir Computers*: exploiting the complex dynamics in networks featuring optical or electro-optical nonlinearity for signal processing [126, 125, 26, 127]

2. *Optical Neurons and Neural Networks*: the attempt to directly implement neural networks (with spiking or continuous wave dynamics) using photonic hardware [124, 123, 29, 101, 116]

3. Coherent Ising Machines: for solving graph-based discrete optimization problems [121, 131, 71, 47, 49, 53, 109, 118]

4. *Probabilistic computation on sparse graphs*: e.g., all optical decoding of Low-Density Parity Check Codes [88, 86].

The Wigner method allows us to investigate what sort of computational devices, beyond logic gates, might be constructed from circuits comprising static linear elements and nonlinear optical resonators. Below I describe and analyze a few such ideas. Both Dmitri Pavlichin ([86], Chapter 3) and Ryan Hamerly ([46], Chapters 7+8) have similar chapters in their theses that feature additional device models.

In the remainder of this chapter, I will present the semi-classical analysis of a few interesting nonlinear resonator models, as well as, some concrete ways in which they can be used for all optical computing and signal processing. This chapter is intended a reference, e.g. when reading the following chapter, which specifies an end-to-end design of a photonic circuit implementing one of the simplest machine learning algorithms, the perceptron algorithm for linear binary classification.

4.1. Basic Single Mode Kerr Model

Let us initially assume an SLH quantum model of the form

$$\mathbf{S} = \mathbf{1}_n, \quad \mathbf{L} = \begin{pmatrix} \sqrt{\kappa_1} a \\ \sqrt{\kappa_2} a \\ \vdots \\ \sqrt{\kappa_n} a \end{pmatrix}, \quad H = \Delta a^\dagger a + \frac{\chi}{2} a^{\dagger 2} a^2 \,. \tag{4.1}$$

The QSDE for the mode operator is given by

$$da = -\left(\frac{\kappa_T}{2} + i\Delta + i\chi a^\dagger a\right) a\mathrm{d}t - \sum_{j=1}^{n} \sqrt{\kappa_j} dA_j, \tag{4.2}$$

where the $\{dA_j,\, j = 1, 2 \ldots, n\}$ are the quantum input noise processes. Each coupling operator $L_j = \sqrt{\kappa_j} a$ corresponds to an individual loss channel of the cavity with a rate κ_j. The total loss rate is equal to the cavity linewidth $\kappa_T = \sum_j \kappa_j$. Some of these coupling operators model the coupling to external waveguide modes while other coupling operators just represent additional loss sources such as intrinsic (linear) resonator losses due to surface roughness.

Assume now that we are driving the system (via one or multiple ports) with coherent (monochromatic) light. Specifically, if each port with coupling operator L_j gets driven by a coherent amplitude ϵ_j then we can represent the overall input as a vector

$$\epsilon = \begin{pmatrix} \epsilon_1 \\ \epsilon_2 \\ \vdots \\ \epsilon_n \end{pmatrix}. \tag{4.3}$$

Feeding this into the above system and applying the SLH circuit algebra rules leads to the following changes in our model:

$$\mathbf{L} \to \mathbf{L} + \mathbf{w}, \quad H \to H + \sum_{j=1}^{n} \frac{\sqrt{\kappa_j}}{2i} \left(a^\dagger \epsilon_j - a \epsilon_j^*\right), \tag{4.4}$$

leading to the modified equation of motion

$$da = -\left(\frac{\kappa_T}{2} + i\Delta + i\chi a^\dagger a\right) a\mathrm{d}t - \sum_{j=1}^{n} \sqrt{\kappa_j}(\epsilon_j \mathrm{d}t + dA_j). \tag{4.5}$$

We see that the coherent displacement was equivalent to simply displacing the noise inputs as $dA_j \to dA_j + \epsilon_j dt$, which is an intuitive result.

4.1.1. Semi-classical steady-state analysis

Transforming now to a semi-classical model, we replace the mode operator $a \to \alpha$ with a semi-classical complex amplitude in (4.5) and the input quantum noise processes by classical complex valued noise processes $dA_j \to d\beta_j$ and find

$$d\alpha = -\left(\frac{\kappa_T}{2} + i\Delta + i\chi|\alpha|^2\right)\alpha \mathrm{d}t - \sum_{j=1}^{n}\sqrt{\kappa_j}(\epsilon_j \mathrm{d}t + d\beta_j). \tag{4.6}$$

Apart from a small shift $\Delta \to \Delta - \chi$, we would have obtained this exact same result by applying the Wigner method [98].

Assuming for simplicity only one coherent drive field $\epsilon_1 = \epsilon$ and taking only the deterministic part of Equation (4.6), we can find the following relationship for the steady state mode amplitude α and the driving field ϵ:

$$\begin{aligned} 0 &= -\left(\frac{\kappa_T}{2} + i\Delta + i\chi|\alpha|^2\right)\alpha - \sqrt{\kappa_1}\epsilon && (4.7)\\ \Rightarrow\ \kappa_1|\epsilon|^2 &= \left|\left(\frac{\kappa_T}{2} + i\Delta + i\chi|\alpha|^2\right)\right|^2 |\alpha|^2 && (4.8)\\ &= \left(\frac{\kappa_T^2}{4} + (\Delta + \chi|\alpha|^2)^2\right)|\alpha|^2. && (4.9)\end{aligned}$$

Defining $n := |\alpha|^2$ as a semi-classical measure of the intra-cavity photon number[1], and $F := \kappa|\epsilon|^2$ as a measure of power in the coherent driving field, their mutual relationship is given by

$$F(n) := \left(\frac{\kappa_T^2}{4} + (\Delta + \chi n)^2\right) n = \left(\frac{\kappa_T^2}{4} + \Delta^2\right) n + 2\Delta\chi n^2 + \chi^2 n^3. \tag{4.10}$$

We have visualized this relationship between intra-cavity energy and input power in Figure 4.1.

This third order polynomial can have two extremal points found by setting $\frac{\partial F}{\partial n}(n) = 0$

$$\begin{aligned} F'(n) &= \left(\frac{\kappa_T^2}{4} + \Delta^2\right) + 4\Delta\chi n + 3\chi^2 n^2 && (4.11)\\ &= 3\chi^2(n - n_+)(n - n_-) && (4.12)\\ \text{where } n_\pm &= -\frac{2\Delta}{3\chi} \pm \sqrt{\frac{1}{9\chi^2}\left[\Delta^2 - \frac{3}{4}\kappa_T^2\right]} && (4.13)\end{aligned}$$

---

[1] According to the Wigner moment equivalence, the correct mapping is actually given by $\langle a^\dagger a\rangle = \langle|\alpha|^2\rangle_W - \frac{1}{2}$.

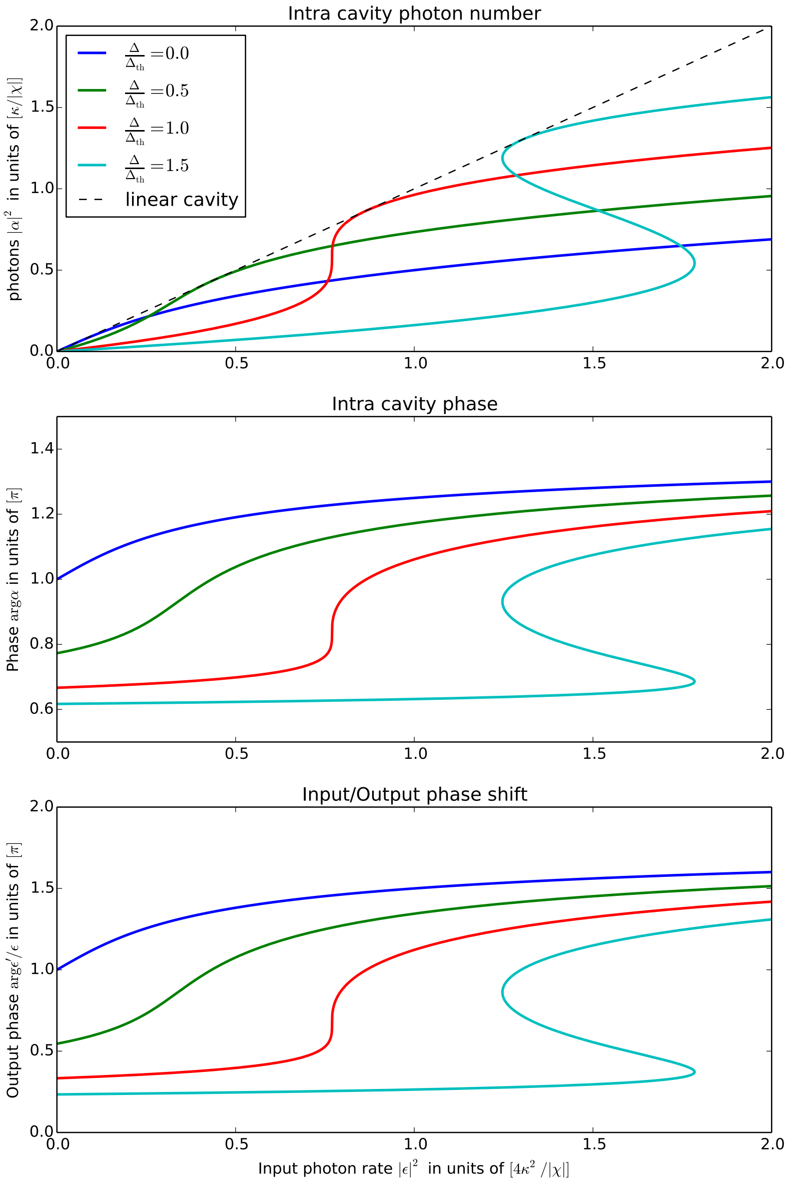


Figure 4.1.: Here we present the input-output characteristics for a single port $\kappa_T = \kappa$ Kerr cavity. The axes are normalized by $\kappa/|\chi|$. Bi-stability occurs when the cavity detuning Δ from the laser drive exceeds a threshold value Δ_{th}. The critical case $\Delta = \Delta_{th}$ is plotted in red. For a single port model, the reflected power equals incident power (in steady state) and it is therefore sufficient to plot only the imparted phase shift.

For these solutions to be real and positive-valued we need the detuning to be large enough to satisfy $|\Delta| \geq \Delta_{th} := \frac{\sqrt{3}}{2}\kappa_T$, and the detuning needs to have the opposite sign than the nonlinear coefficient, $\Delta/\chi < 0$. We thus find that the extremal points are evenly spaced about $n_0 = -\frac{2\Delta}{3\chi}$. It is possible to see that the existence of two extremal points in the input-power photon number characteristic is equivalent to the system being bistable.

If we look at the second derivative of $F(n)$, we find that for $\Delta/\chi < 0$ there exists an inflection point at n_0 (regardless of the magnitude of the detuning):

$$F''(n) = 4\Delta\chi + 6\chi^2 n = 6\chi^2(n - n_0)\,. \tag{4.14}$$

The input photon rate at this inflection point is given by

$$F(n_0) = -\frac{2\Delta}{3\chi}\frac{\chi^2}{4}\left(\frac{\kappa_T^2}{\chi^2} + \frac{4\Delta^2}{9\chi^2}\right)\,. \tag{4.15}$$

When the system is bi-stable, the inflection point is an unstable solution of the steady-state relationship, otherwise it is stable. If we introduce the shifted rate characteristic $Q(n) := F(n + n_0) - F(n_0)$, this will have the form

$$Q(n) = \chi^2 n(n^2 - q)\,, \tag{4.16}$$

where q is positive-real-valued for $|\Delta| \geq \Delta_{th}$ and $\Delta/\chi < 0$. It turns out that

$$q = \frac{\kappa_T^2}{4\chi^2}\left[\frac{\Delta^2}{\Delta_{th}^2} - 1\right]\,. \tag{4.17}$$

This implies that if the system is biased with an input rate $F(n_0)$, the bistable states are separated by $\delta n := 2\sqrt{q}$.

In general, for any complex input amplitude that leads to a steady state photon number n, the corresponding mode amplitude is given by

$$\alpha(n) = -\frac{\epsilon}{|\epsilon|}\sqrt{n\frac{\kappa_T/2 - i\Delta - i\chi n}{\kappa_T/2 + i\Delta + i\chi n}}\,. \tag{4.18}$$

In the bi-stable regime, the different mode amplitudes in the upper and lower branch can also be computed. The photon number here is equal to $n_\pm = n_0 \pm \sqrt{q}$. We then find

$$\alpha_\pm := a(n_\pm) = -\sqrt{(n_0 \pm \sqrt{q})\frac{\kappa_T/2 - i\Delta/3 \mp i\chi\sqrt{q}}{\kappa_T/2 + i\Delta/3 \pm i\chi\sqrt{q}}}\,. \tag{4.19}$$

Finally, for bi-stable conditions, we may ask about the width of the hysteresis curve, i.e. what differences in input power are necessary to switch. This can be obtained from the extremal points of $F(n)$, or more simply from $Q(n)$:

$$Q'(n) = \chi^2(n^2 - q) + 2\chi^2 n^2 = 3\chi^2 n^2 - \chi^2 q = 0 \Leftrightarrow n = \pm\sqrt{q/3} \tag{4.20}$$

At the extremal points, $Q(n)$ takes on the values $Q(\pm\sqrt{q/3}) = \mp 2\chi^2\sqrt{q^3/27}$. Thus, if the input power is biased to $F(n_0)$, it takes a change of $\delta F = 2\chi^2\sqrt{q^3/27}$ in input power (in either direction) to reliably switch the bi-stable states.

4.1.2. Linear Transfer Function

The $ABCD$ matrices in the complex amplitude representation of the above model, displaced as $a \to a + \alpha_0$ and then linearized, are given by

$$A = \breve{\Delta}\left(-\frac{\kappa_T}{2} - i\Delta - 2i\chi|\alpha_0|^2, -i\chi\alpha_0^2\right) \tag{4.21}$$

$$B = -\breve{\Delta}\left(C_-^\dagger, 0\right), \tag{4.22}$$

$$C = \breve{\Delta}\left(C_-, 0\right), \tag{4.23}$$

$$D = \mathbf{1}_{2n}, \tag{4.24}$$

where $C_- = \begin{pmatrix}\sqrt{\kappa_1} & \cdots & \sqrt{\kappa_n}\end{pmatrix}^T$, and where we are using the double-up notation introduced by [85, 140]

$$\breve{\Delta}\left(A, B\right) := \begin{pmatrix} A & B \\ B^\sharp & A^\sharp \end{pmatrix}, \tag{4.25}$$

with $A^\sharp$ being the element-wise complex conjugate of a matrix A.

In terms of these matrices, the linear transfer function relating inputs to outputs is given as [42, 140]

$$\Xi(\omega) = D - C\left(i\omega\mathbf{1} + A\right)^{-1} B. \tag{4.26}$$

We can explicitly compute the inverse appearing in the transfer function to be

$$\left(i\omega\mathbf{1} + A\right)^{-1} = \begin{pmatrix} -\frac{\kappa_T}{2} - i\Delta - 2i\chi|\alpha_0|^2 + i\omega & -i\chi\alpha_0^2 \\ i\chi\alpha_0^{*2} & -\frac{\kappa_T}{2} + i\Delta + 2i\chi|\alpha_0|^2 + i\omega \end{pmatrix} \tag{4.27}$$

$$= \frac{1}{D(\omega)}\begin{pmatrix} -\frac{\kappa_T}{2} + i\Delta + 2i\chi|\alpha_0|^2 + i\omega & i\chi\alpha_0^2 \\ -i\chi\alpha_0^{*2} & -\frac{\kappa_T}{2} - i\Delta - 2i\chi|\alpha_0|^2 + i\omega \end{pmatrix} \tag{4.28}$$

with the determinant $D(\omega) = \left|(i\omega\mathbf{1} + A)^{-1}\right|$ given by

$$D(\omega) = \left(\frac{\kappa_T}{2} - i\omega\right)^2 + \left(\Delta + 2\chi|\alpha_0|^2\right)^2 - |\chi|^2|\alpha_0|^4. \tag{4.29}$$

Since $\mathrm{Im}\left[D(\omega)\right] = -\kappa_T\omega$ we see that $D(\omega)$ cannot be zero for $\omega \neq 0$, showing that a single mode Kerr-cavity (without feedback) cannot have a conjugate pair of complex poles crossing the imaginary axis, i.e., it does not admit a Hopf-Bifurcation.

The transfer function between two ports j, k is given by

$$\Xi(\omega)_{jk} = \delta_{jk}\mathbf{1}_2 + \frac{\sqrt{\kappa_j\kappa_k}}{D(\omega)}\begin{pmatrix} -\frac{\kappa_T}{2} + i\Delta + 2i\chi|\alpha_0|^2 + i\omega & i\chi\alpha_0^2 \\ -i\chi\alpha_0^{*2} & -\frac{\kappa_T}{2} - i\Delta - 2i\chi|\alpha_0|^2 + i\omega \end{pmatrix} \tag{4.30}$$

For non-zero α_0 it is phase/quadrature sensitive and thus needs to be specified as a 2×2 matrix.

4.2. Phase Sensitive Amplifier Model

We now consider the special case of a single port $\kappa_T = \kappa_1 = \kappa$. We furthermore limit ourselves to analyzing the dc-transfer function, i.e., $\omega = 0$. Assume that we are biasing our system with an input ϵ_0. Since we assume zero internal loss, in the steady-state the bias amplitude only receives an intensity dependent phase shift $\epsilon_0' = \epsilon_0 e^{i\phi_0}$.

In this case the transfer function is given by

$$\Xi_a = \mathbf{1}_2 + \frac{\kappa}{D(0)}\begin{pmatrix} -\frac{\kappa}{2} + i\Delta + 2i\chi|\alpha_0|^2 & i\chi\alpha_0^2 \\ -i\chi\alpha_0^{*2} & -\frac{\kappa}{2} - i\Delta - 2i\chi|\alpha_0|^2 \end{pmatrix}. \tag{4.31}$$

The phase sensitive gain values for the maximally and minimally amplified orthogonal quadratures can be obtained from the singular value decomposition of this matrix. At dc, the transfer function has the especially simple (doubled up) form

$$\Xi_a = \begin{pmatrix} U & V \\ V^* & U^* \end{pmatrix}. \tag{4.32}$$

A straightforward (but somewhat tedious) calculation reveals that the singular value of such a matrix is given by

$$\begin{pmatrix} U & V \\ V^* & U^* \end{pmatrix} = \frac{1}{2}\begin{pmatrix} \eta' & i\eta' \\ \eta'^* & -i\eta'^* \end{pmatrix}\begin{pmatrix} ||U| + |V|| & 0 \\ 0 & ||U| - |V|| \end{pmatrix}\begin{pmatrix} \eta^* & \eta \\ -i\eta^* & i\eta, \end{pmatrix} \tag{4.33}$$

where $\eta = \sqrt{\frac{U^*V}{|U||V|}}$ and $\eta' = \sqrt{\frac{UV}{|U||V|}}$ are complex phase factors determining the amplified and deamplified input quadratures and their respective image.

Furthermore, it can be shown that for fixed κ, Δ and χ, the larger singular value is maximal when

$$|\alpha_0|^2 = \sqrt{\frac{\Delta^2 + \frac{\kappa^2}{4}}{3\chi^2}}. \tag{4.34}$$

It turns out that in this case the amplifier acts as a perfect squeezer in that one quadrature gets amplified by a factor g whereas the orthogonal quadrature gets deamplified with a factor g^{-1}. This factor can be computed to be

$$g = \sqrt{\frac{\sqrt{f}+\kappa}{\sqrt{f}-\kappa}}, \text{ with } f = 28\Delta^2 + 4\kappa^2 - 8\Delta\sqrt{12\Delta^2 + 3\kappa^2} \tag{4.35}$$

Below the bi-stability threshold $|\Delta| \leq \frac{\sqrt{3}}{2}\kappa$, we find $f \geq \kappa^2$. The minimum amplitude gain achievable in this fashion is $\sqrt{3}$, which is realized in the $\Delta \to 0$ limit. For $\Delta \to \Delta_{th}$ the gain diverges, as one would expect. We can invert the relationship in order to compute the detuning at which a particular amplitude gain g is realized. This yields

$$\Delta = \frac{\sqrt{3}\kappa}{2} \frac{(g-\sqrt{3})(g-\frac{1}{\sqrt{3}})}{g^2-1}\,. \tag{4.36}$$

To verify our result, note that for $g \to \infty$, we have $\Delta \to \Delta_{th}$, as one would expect.

4.2.1. Kerr Amplifier Recipe

Based on the previous results, we have a simple recipe to construct a single port Kerr-amplifier. Given a desired amplitude gain g, a system bandwidth κ, and an intrinsic non-linear coefficient χ, proceed as follows.

1. Compute Δ according to Equation (4.36).
2. Compute the intra-cavity bias amplitude $|\alpha_0|^2$ according to (4.34).
3. Compute the necessary input amplitude to achieve that intra-cavity field $\epsilon_0 = \sqrt{\frac{F(|\alpha_0|^2)}{\kappa}}$ using (4.10).

To actually use the cavity as an amplifier, it is necessary to add the correct bias to any input signal and subtract the scattered bias downstream. Expressed in the Gough-James circuit algebra this is

$$A = W_{-\beta_0 e^{i\phi_0}} \lhd C \lhd W_{\beta_0}. \tag{4.37}$$

The whole circuit should be wrapped in additional phase shifters to normalize which quadrature is (anti-)squeezed and which output quadrature it maps to.

As we can see in Figure 4.3, the dependence of the gain on the bias power is fairly peaked, suggesting that it saturates at less than ten percent of the bias power.

The sharpness of the bias-gain characteristic, and therefore the saturation effect, increases with a higher maximum gain as can be seen in Figure 4.4. A more detailed analysis of this effect would be interesting but one reasonable explanation would be that the overall possible output amplitude swings for a cavity with given κ and χ are only weakly dependent on the maximal gain.

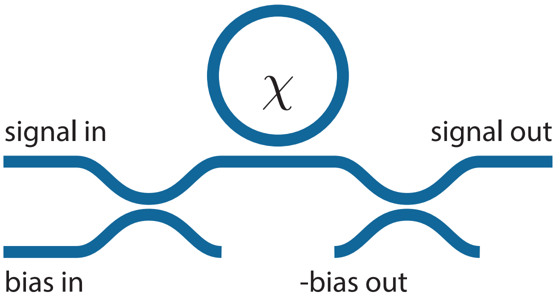


Figure 4.2.: A simple single port Kerr cavity amplifier with bias addition and subtraction elements.

The frequency bandwidth of our amplifier is necessarily bounded by the cavity linewidth κ, but it turns out to be even smaller than that as can be seen in Figures 4.5 and 4.6. The bandwidth generally decreases for increasing gain.

As we will see in the next section, there exists a nice construction, that has several advantages.

4.3. Symmetric amplifier design

A symmetric construction combining two identical Kerr-cavities (and twice the bias power) leads to a particularly elegant scheme.

The small signal u and a bias amplitude $\sqrt{2}\epsilon_0$ (multiplied by $\sqrt{2}$ because we need to bias both resonators) are mixed on a 50/50 beamsplitter. Each of the outputs $\epsilon_\pm = \epsilon_0 \pm u/\sqrt{2}$ reflects off one Kerr-cavity and then the reflected beams $\epsilon'_\pm \approx \epsilon_0 e^{i\phi_0} \pm \frac{g(\theta)u}{\sqrt{2}}$ are recombined in a beamsplitter with inverse scattering matrix. Both amplifiers are biased to the same intra-cavity photon number, but they receive the signal input with opposite signs. Thus, the amplified signal can easily be extracted on the output beamsplitter because it is simply the difference of the scattered beams:

$$u' = \frac{\epsilon'_+ - \epsilon'_-}{\sqrt{2}} = g(\theta)u, \tag{4.38}$$

where $g(\theta)$ is the signal phase dependent complex gain factor. This argument already correctly suggests that the overall circuit's transfer function from $u \to u'$ is identical to that of a single cavity, but some obvious advantages of this construction are the following:

1. The bias addition and subtraction happens automatically and without incurring additional loss or noise (at least to linear order).

2. The linear relationship $u' = g(\theta)u$ is correct up to third order in u, because the second order (and thus even) contributions are subtracted. This makes the amplifier

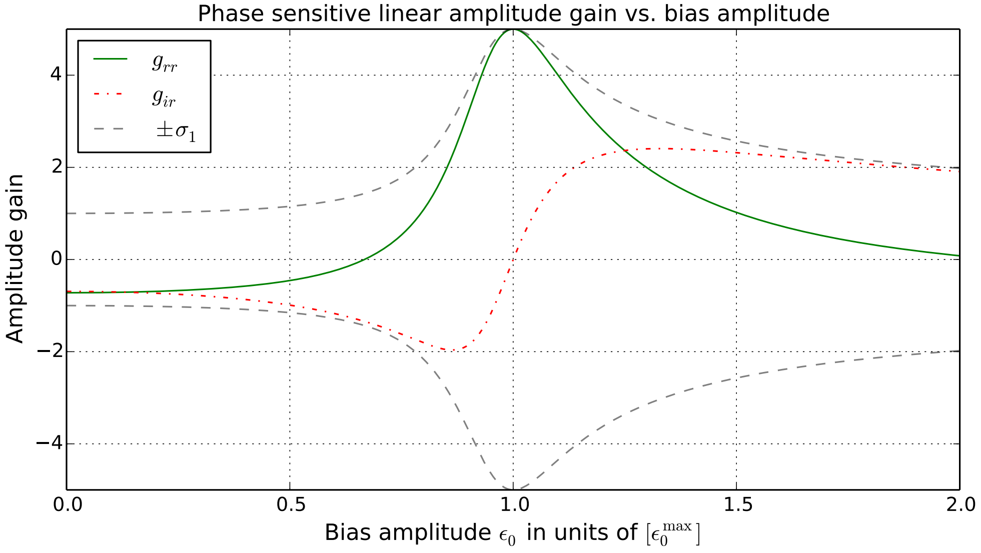


Figure 4.3.: The gain falls off quickly away from the optimal bias input ϵ_0. We see, however, that the imaginary to real quadrature gain has a zero close to the maximum real to real gain. The strongly-sloped linear behavior of this gain suggests that the amplifier acts as a quadratic map for inputs in the imaginary quadrature. The dashed envelope's amplitude equals σ_1, the largest singular value of the steady-state transfer function at each bias amplitude. It equals the maximally achievable input-output gain between any two input and output quadratures. We have wrapped the amplifier in phase shifts ensuring that the maximal gain is between the real input and output quadratures at the optimal bias.

response completely symmetric under $u \to -u$. This design principle is similar to what is often done in electronics where complementary pairs of transistors are used to increase overall linearity.

3. As we discuss below, this amplifier's gain can dynamically be tuned by reducing the bias input. The clean separation of bias and signal inputs makes this very easy to do.
4. This construction should work just as well for any other amplifier model that requires bias and signal to enter the cavity through the same physical port.

4.4. Quadrature Filter

A very similar model to the symmetric amplifier discussed in Section 4.3 allows us to construct a circuit that transmits only a single quadrature of the input and subtracts the other quadrature by transmitting it to an additional drop port. This situation is characterized by a non-full rank transfer function, or equivalently a zero determinant. This already suggests

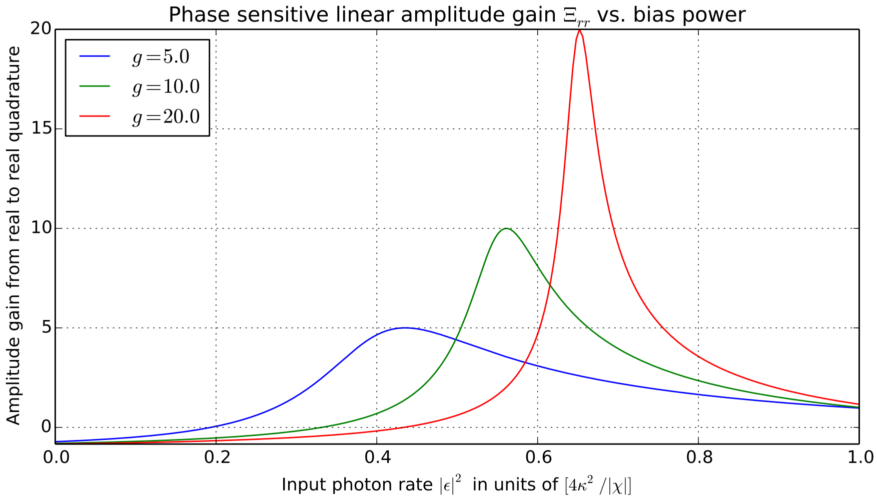


Figure 4.4.: The dc real-to-real quadrature amplitude gain for different chosen gain factors g (and therefore different detunings Δ, but constant κ, χ). We see that the peaks become narrower for increasing maximal gain suggesting stronger saturation effects.

that at minimum, a two port model is necessary, because in our formalism single input/output transfer functions are always of full rank. This is a consequence of the requirement to preserve the canonical commutation relationship of the scattered quantum field operators. The condition for the reflection transfer function off a port with coupling κ to be of rank one can be expressed as

$$\left[D(0) - \frac{\kappa\kappa_T}{2}\right]^2 + \kappa^2\left[\Delta + 2\chi|\alpha_0|^2\right]^2 = \kappa^2\chi^2|\alpha_0|^4. \tag{4.39}$$

For a two port model with matching coupling/loss rates $\kappa = \kappa_T/2$ this is equivalent to

$$\begin{aligned} 0 &= \left[\left(\Delta + 2\chi|\alpha_0|^2\right)^2 - |\chi|^2|\alpha_0|^4\right]^2 + \kappa^2\left[\left(\Delta + 2\chi|\alpha_0|^2\right)^2 - \chi^2|\alpha_0|^4\right] && (4.40)\\ &= \left[\left(\Delta + 2\chi|\alpha_0|^2\right)^2 - |\chi|^2|\alpha_0|^4\right]\left[\kappa^2 + \left(\Delta + 2\chi|\alpha_0|^2\right)^2 - |\chi|^2|\alpha_0|^4\right]. && (4.41)\end{aligned}$$

Introducing $x := \chi|\alpha_0|^2$, this can be further factored as:

$$\begin{aligned} 0 &= [\Delta^2 + 4\Delta x + 3x^2][\kappa^2 + \Delta^2 + 4\Delta x + 3x^2] && (4.42)\\ &= 9\left[\left(x + \frac{2}{3}\Delta\right)^2 - \frac{\Delta^2}{9}\right]\left[\left(x + \frac{2}{3}\Delta\right)^2 - \frac{\Delta^2 - 9\kappa^2}{9}\right] && (4.43)\\ &= 9\left[x + \frac{\Delta}{3}\right][x + \Delta]\left[x + \frac{2}{3}\Delta + \frac{\sqrt{\Delta^2 - 9\kappa^2}}{3}\right]\left[x + \frac{2}{3}\Delta - \frac{\sqrt{\Delta^2 - 9\kappa^2}}{3}\right] && (4.44)\end{aligned}$$

The first two solutions always exist when χ and Δ have opposite signs. The other pair of solutions only exist for $|\Delta| \geq 3\kappa$, which is far in the bi-stable regime and should thus be avoided. The case $x = \chi|\alpha_0|^2 = -\Delta$ equivalent to $|\alpha_0|^2 = -\frac{\Delta}{\chi}$, will from here on be

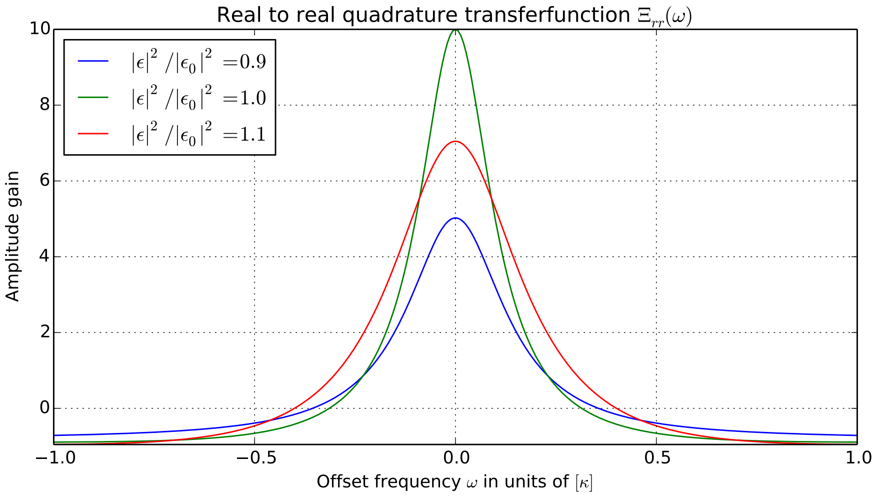


Figure 4.5.: When the input amplitude ϵ differs slightly from the optimal bias input ϵ_0 the gain decreases, but the bandwidth increases.

called the 'dynamic resonance condition', because it implies an overall 'dynamic' detuning of $\Delta(|\alpha_0|^2) = \Delta + \chi|\alpha_0|^2 = 0$. There is a second solution $x = \chi|\alpha_0|^2 = -\frac{\Delta}{3}$, but it leads to a less than unity gain on the non-filtered quadrature. We will thus limit our discussion to the dynamically resonant model.

4.4.1. Dynamically Resonant Quadrature Filter

In the dynamically resonant case the steady-state reflection transfer function is given by

$$\Xi_R(0) = -\frac{i\Delta}{\kappa}\begin{pmatrix} 1 & 1 \\ -1 & -1 \end{pmatrix} \tag{4.45}$$

which shows that an input field in the imaginary quadrature will be fully subtracted from the input mirror. A real quadrature input, on the other hand, will be reflected with an amplitude gain of $|g| = \sqrt{2}\frac{\Delta}{\kappa}$ and receive a $-\pi/2$ phaseshift. Keeping in mind that for the system to have a unique steady state for all given input powers, the gain can at most be $\sqrt{2}\frac{\Delta_{th}}{\kappa} = \sqrt{6}$ (where it is important to remember that $\kappa = \kappa_T/2$). Furthermore, the necessary bias power to achieve the correct intra-cavity bias field can be computed to be

$$|\epsilon|^2 = \kappa^{-1}F\left(-\frac{\Delta}{\chi}\right) = -\frac{\Delta\kappa}{\chi} = \frac{|g|\kappa_T}{4\sqrt{2}}\frac{\kappa_T}{|\chi|}\,. \tag{4.46}$$

In the last expression we can see that the necessary bias power (or more correctly input rate of bias photons) scales linearly with the achieved amplitude gain, the cavity linewidth and the number of photons required to shift the cavity resonance by one linewidth.

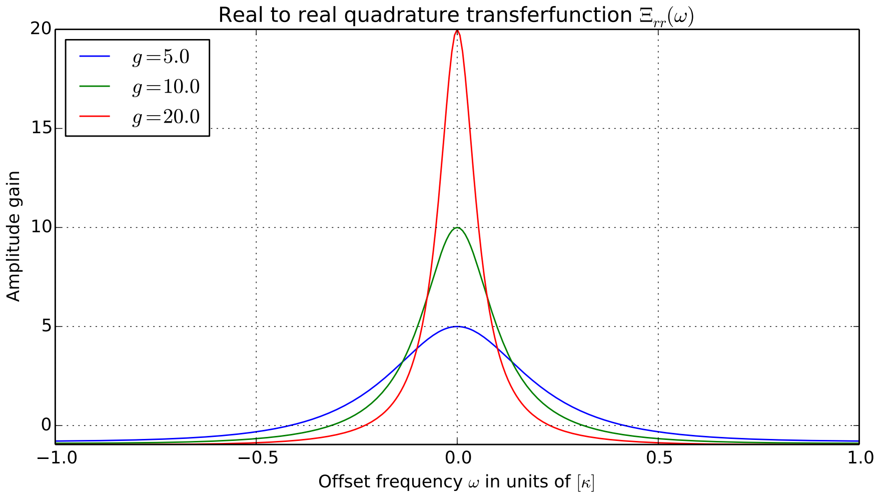


Figure 4.6.: The bandwidth of the amplifier decreases with increasing maximal gain g.

4.4.2. Noise properties of the dynamically resonant quadrature filter

Since we have designed the filter to have matching port couplings $\kappa_1 = \kappa_2$ we need not consider additional internal losses unless they exceed our chosen value for κ. Otherwise we can group all other losses (internal losses as well as additional waveguide couplings) in the second port and engineer them to have equal total coupling rate κ. The additional input port will only be driven by a vacuum but the noise from this input can be transmitted through the cavity to the port that we are reflecting our signal off of. We should therefore also compute the steady state transfer function for transmission from one port to the other. This turns out to just be

$$\Xi_T(0) = \Xi_R(0) - \mathbf{1} = \begin{pmatrix} -\frac{i\Delta}{\kappa} - 1 & -\frac{i\Delta}{\kappa} \\ \frac{i\Delta}{\kappa} & \frac{i\Delta}{\kappa} - 1 \end{pmatrix} . \tag{4.47}$$

Therefore, if we pick our gain to be unity, the primary signal output port will consist of the reflected real input quadrature (rotated by $-\pi/2$ into the imaginary quadrature) as well as transmitted (and amplified) shot noise from the other input port. The real quadrature of the signal output will only contain a vacuum noise level, while the imaginary quadrature contains the filtered real input quadrature with unit gain, its shot noise and the sum of the real and imaginary quadrature shot noise contributions from the other input port. Therefore, if the quadrature filter input's noise exceeds the shot noise level by much more than a factor of two (in power, $\sqrt{2}$ in amplitude), then the additional noise due to the filter should not play a huge role. In Figure 4.8 we show a simulated trace for this system with $\frac{|\chi|}{\kappa} = \frac{1}{10{,}000}$ and unit gain.

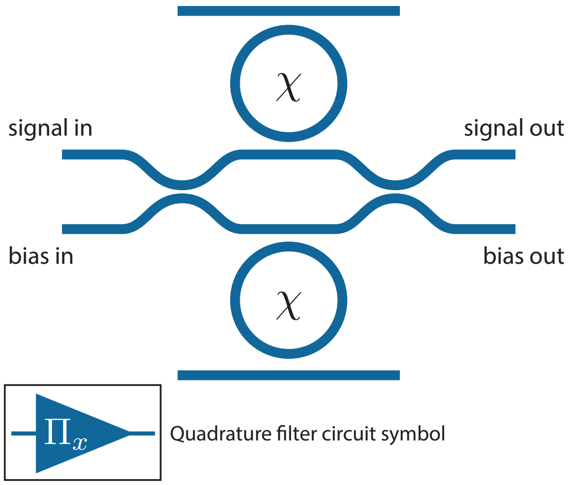


Figure 4.7.: A Kerr cavity based quadrature filter comprising a pair of two-port cavities biased to dynamic resonance.

4.5. Two-mode Kerr-models

The two-mode Kerr-model is quite similar to two copies of a single mode model but features an additional cross-Kerr-coupling term. For simplicity, we restrict ourselves to a single input-output coupling per mode. This yields a two-port, two-mode model

$$\mathbf{S} = \mathbf{1}_2, \quad \mathbf{L} = \begin{pmatrix} \sqrt{\kappa_a} a \\ \sqrt{\kappa_b} b \end{pmatrix}, \quad H = \Delta_a a^\dagger a + \Delta_b b^\dagger b + \frac{\chi_a}{2} a^{\dagger 2} a^2 + \frac{\chi_b}{2} b^{\dagger 2} b^2 + \chi_{ab} a^\dagger a b^\dagger b. \tag{4.48}$$

The QSDE's for a and b are then given by

$$da = -\left(\frac{\kappa_a}{2} + i\Delta_a + i\chi_a a^\dagger a + i\chi_{ab} b^\dagger b\right) a\mathrm{d}t - \sqrt{\kappa_a} dA_a, \tag{4.49}$$

$$db = -\left(\frac{\kappa_b}{2} + i\Delta_b + i\chi_{ab} a^\dagger a + i\chi_b b^\dagger b\right) b\mathrm{d}t - \sqrt{\kappa_b} dA_b. \tag{4.50}$$

Thus we see that, as before, the modes experience a power dependent detuning, which now depends on the photon numbers in both modes. We can carry out a similar steady state analysis of the input output behavior as in the single mode case. Assuming constant input amplitudes ϵ_a, ϵ_b to each mode we quickly find the semiclassical steady state relationship between internal mode amplitudes $\alpha \leftrightarrow \langle a \rangle, \beta \leftrightarrow \langle b \rangle$:

$$0 = -\left(\frac{\kappa_a}{2} + i\Delta_a + i\chi_a|\alpha|^2 + i\chi_{ab}|\beta|^2\right)\alpha - \sqrt{\kappa_a}\epsilon_a\,, \tag{4.51}$$

$$0 = -\left(\frac{\kappa_b}{2} + i\Delta_b + i\chi_{ab}|\alpha|^2 + i\chi_b|\beta|^2\right)\beta - \sqrt{\kappa_b}\epsilon_b \tag{4.52}$$

$$\Rightarrow \kappa_a|\epsilon_a|^2 = \left[\frac{\kappa_a^2}{4} + \left(\Delta_a + \chi_a|\alpha|^2 + \chi_{ab}|\beta|^2\right)^2\right]|\alpha|^2, \tag{4.53}$$

$$\kappa_b|\epsilon_b|^2 = \left[\frac{\kappa_b^2}{4} + \left(\Delta_b + \chi_{ab}|\alpha|^2 + \chi_b|\beta|^2\right)^2\right]|\beta|^2\,. \tag{4.54}$$

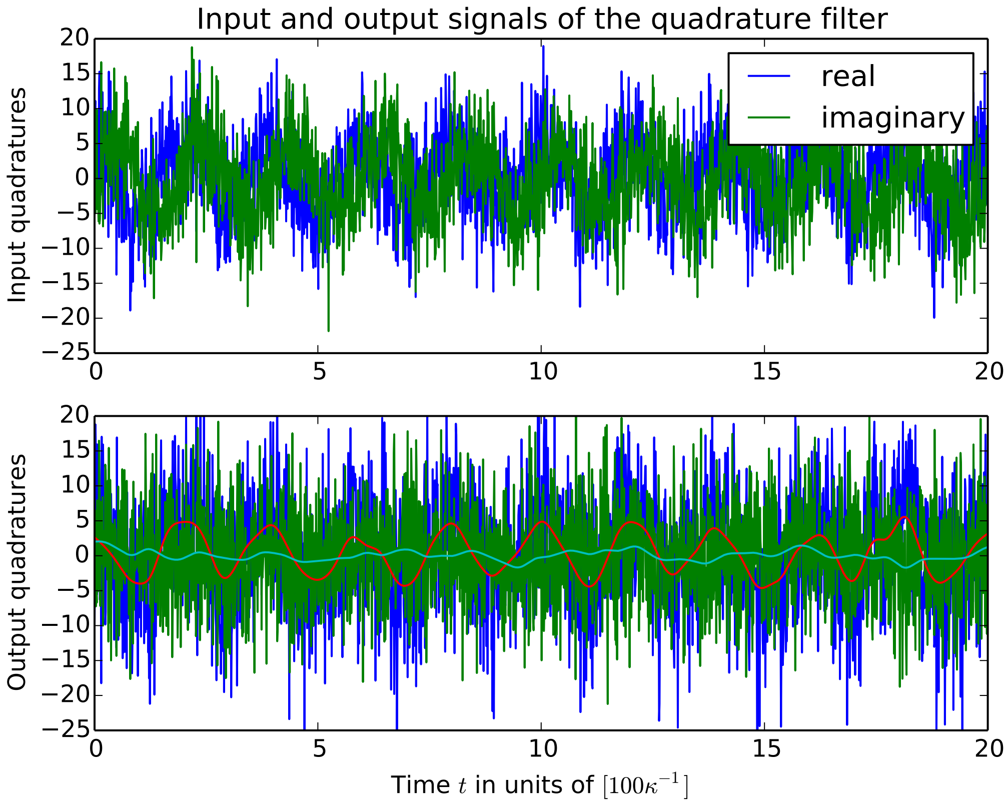


Figure 4.8.: Stochastically simulated input and output signals of a quadrature filter. For a constant amplitude input with slowly varying phase, the output contains only the real quadrature of the input (with unit gain) plus a somewhat increased noise in the imaginary output quadrature. The noise in this figure is effectively low-passed at the finite simulation stepsize $\Delta t^{-1} = 10\kappa$. Input noise at frequencies far off resonance is not affected by the system and simply sees a unit gain.

Defining again $n := |\alpha|^2, m := |\beta|^2$ as semi-classical measures of the intra-cavity photon number, and $F := \kappa_a|\epsilon_a|^2, G := \kappa_b|\epsilon_b|^2$ as measures of power in the coherent driving fields, their mutual relationship is given by

$$F(n,m) := \left(\frac{\kappa_a^2}{4} + (\Delta_a + \chi_a n + \chi_{ab} m)^2\right) n \tag{4.55}$$

$$= \left(\frac{\kappa_a^2}{4} + \Delta_a^2\right) n + 2\Delta_a\chi_a n^2 + 2\Delta_a\chi_{ab} nm + 2\chi_a\chi_{ab} n^2 m + \chi_{ab}^2 nm^2 + \chi_a^2 n^3, \tag{4.56}$$

$$G(n,m) = \left(\frac{\kappa_b^2}{4} + \Delta_b^2\right) m + 2\Delta_b\chi_b m^2 + 2\Delta_b\chi_{ab} nm + 2\chi_b\chi_{ab} nm^2 + \chi_{ab}^2 mn^2 + \chi_b^2 m^3 \tag{4.57}$$

A full discussion of when the system becomes bi-stable would exceed the scope of this work. However, we note that the cases $m = 0$ or $n = 0$ reduce exactly to the single mode problem, as one would expect. This implies that the single mode results are certainly still necessary conditions for uni-stability (if all χ coefficients have the same sign, opposite to Δ)

$$|\Delta_{a/b}| \leq \frac{\sqrt{3}\kappa_{a/b}}{2}. \tag{4.58}$$

In the following, we will usually designate a to be the signal mode, which may take on continuous values, centered around a constant offset. We take b to be the control mode, usually with only two different input conditions.

4.5.1. Cross-Kerr-effect phase modulator

Using a two-mode cross-Kerr modulation allows us to get a phase insensitive device and for small/negligible losses we can reuse the scattered control signal to drive a whole array of these devices. The easiest way to understand the general principle is as follows. If there is no direct linear coupling between the two-modes (this is if they are not degenerate in frequency or counter-propagating degenerate modes with negligible cross-scattering), then their interaction is phase insensitive and can be understood intuitively as each mode's excitation leading to an additional detuning for itself and the respective other mode $\Delta_a' = \Delta_a'(n,m) = \Delta_a + \chi_a n + \chi_{ab} m$ and $\Delta_b' = \Delta_b'(n,m) = \Delta_b + \chi_{ab} n + \chi_b m$.

If there is just a single port per mode, no cross-scattering, and we neglect internal losses, then a reflected beam coupling to either mode will receive no attenuation, but just a phase shift that depends on both modes excitation energy. The overall phase factor of a scattered signal mode field in steady state is given by

$$\epsilon_a'/\epsilon_a = e^{i\phi_a(n,m)} = -\frac{\frac{\kappa_a}{2} - i\Delta_a'(n,m)}{\frac{\kappa_a}{2} + i\Delta_a'(n,m)}. \tag{4.59}$$

The device works if for a low control input the signal mode is detuned half a linewidth above the laser driving frequency $\Delta_a' = \kappa_a/2$ and for a high control input the signal mode's detuning is shifted negatively by one linewidth to $\Delta_a' = -\kappa_a/2$ via the cross-Kerr interaction. From the above expression we can see that this leads to a phase factors of i and $-i$ for the two control input conditions, respectively.

Things are simplified (and made maximally energy efficient) by choosing the control mode's detuning such that the high power state corresponds to dynamic resonance $\Delta_b' \approx 0$ which gives the maximum build up of control photons in the cavity per input power.

Let's assume that the control input takes on the values 0 and ξ. And let's assume that the average amplitude in the signal is ϵ and that any signals are given by small modulations around this constant value (usually we will chose $\epsilon = 0$).

Then, we have the following steady state relationships

$$\sqrt{\kappa_a}\epsilon_a = -\left[\frac{\kappa_a}{2} + i\left(\Delta_a + \chi_a n + \chi_{ab} m\right)\right]\alpha\,, \tag{4.60}$$

$$\sqrt{\kappa_b}\epsilon_b = -\left[\frac{\kappa_b}{2} + i\left(\Delta_b + \chi_{ab} n + \chi_b m\right)\right]\beta\,. \tag{4.61}$$

with $n(\epsilon_a, \epsilon_b) = |\alpha(\epsilon_a, \epsilon_b)|^2$ and $m(\epsilon_a, \epsilon_b) = |\beta(\epsilon_a, \epsilon_b)|^2$ implicitly defined by the steady state relationships.

We now demand that for a zero input on the control mode, the effective detuning for the signal mode be $+\kappa_a/2$ and for a high input on the control mode it should shift to $-\kappa_a/2$ such that the relative phase difference in the transfer function for the signal (linearized about that average input amplitude) is given by π. These conditions are equivalent to

$$\Delta_a + \chi_a n(\epsilon, 0) = \kappa_a/2 \tag{4.62}$$

$$\Delta_a + \chi_a n(\epsilon, \xi) + \chi_{ab} m(\epsilon, \xi) = -\kappa_a/2 \tag{4.63}$$

Furthermore, we would like the control mode's effective detuning to be zero at high control power. This leads us to

$$\Delta_b + \chi_b m(\epsilon, \xi) + \chi_{ab} n(\epsilon, \xi) = 0. \tag{4.64}$$

Since, by construction, the effective detuning for the signal mode is of the same magnitude for either control mode state, we can assume that $n_a(\epsilon, 0) = n_a(\epsilon, \xi) =: n_\epsilon$.

The control photon number is non-zero only for the high power input and we call its value at that point $m_\xi = m(\epsilon, \xi)$ In this case we can write the above relations as

$$\Delta_a = \kappa_a/2 - \chi_a n_\epsilon\,, \tag{4.65}$$

$$\Delta_b = -\chi_b m_\xi - \chi_{ab} n_\epsilon\,, \tag{4.66}$$

$$m_\xi = \frac{-\kappa_a/2 - \Delta_a - \chi_a n_\epsilon}{\chi_{ab}}\,. \tag{4.67}$$

Inserting the first into the third equation yields

$$m_\xi = -\frac{\kappa_a}{\chi_{ab}} = \frac{\kappa_a}{|\chi_{ab}|}\,, \tag{4.68}$$

and inserting this into the second we find

$$\Delta_b = \frac{\chi_b \kappa_a}{\chi_{ab}} - \chi_{ab} n_\epsilon\,. \tag{4.69}$$

Finally, since we determined everything such that the effective signal detuning is $\mp\kappa_a/2$, we can express n_ϵ in terms of the average driving amplitude: $n_\epsilon = \frac{2|\epsilon|^2}{\kappa_a}$, and similarly, since the control mode is at dynamic resonance, we know that $m_\xi = \frac{4|\xi|^2}{\kappa_b}$ and consequently must have $|\xi| = \frac{\sqrt{\kappa_a\kappa_b}}{2\sqrt{|\chi_{ab}|}}$.

So we finally have

$$\Delta_a = \frac{\kappa_a}{2} - \frac{2\chi_a|\epsilon|^2}{\kappa_a}\,, \tag{4.70}$$

$$\Delta_b = \frac{\kappa_a\chi_b}{\chi_{ab}} - \frac{2\chi_{ab}|\epsilon|^2}{\kappa_a}\,, \tag{4.71}$$

$$\xi = \frac{\sqrt{\kappa_a\kappa_b}}{2\sqrt{|\chi_{ab}|}}\,. \tag{4.72}$$

In order to fulfill our single mode stability conditions we need to ensure that

$$\Delta_a = \frac{\kappa_a}{2} - \frac{2\chi_a|\epsilon|^2}{\kappa_a} \leq \sqrt{3}\frac{\kappa_a}{2} \tag{4.73}$$

$$\Leftrightarrow \kappa_a \geq \sqrt{\frac{4|\chi_a||\epsilon|^2}{\sqrt{3}-1}} \tag{4.74}$$

as well as

$$\Delta_b = \frac{\kappa_a\chi_b}{\chi_{ab}} - \frac{2\chi_{ab}|\epsilon|^2}{\kappa_a} \leq \sqrt{3}\frac{\kappa_b}{2} \tag{4.75}$$

$$\Leftrightarrow \kappa_b \geq \frac{2}{\sqrt{3}}\left(\frac{\kappa_a\chi_b}{\chi_{ab}} + \frac{2|\chi_{ab}||\epsilon|^2}{\kappa_a}\right). \tag{4.76}$$

Adhering to these conditions has generally yielded stable models in numerical simulations. For zero signal offset $\epsilon = 0$ we can first pick κ_a and then choose a κ_b such that $\kappa_b \geq \frac{2}{\sqrt{3}}\frac{\kappa_a\chi_b}{\chi_{ab}}$. For non-zero signal offset we can fix $r_a, r_b > 1$ and choose

$$\kappa_a = r_a\sqrt{\frac{4|\chi_a||\epsilon|^2}{\sqrt{3}-1}}\,, \tag{4.77}$$

$$\kappa_b = r_b r_a \frac{2}{\sqrt{3}}\sqrt{\frac{4|\chi_a||\epsilon|^2}{\sqrt{3}-1}}\frac{\chi_b}{\chi_{ab}}\left(1 + \frac{(\sqrt{3}-1)\chi_{ab}^2}{2r_a^2\chi_a\chi_b}\right)\,. \tag{4.78}$$

and then proceed to compute Δ_a, Δ_b and the high control input amplitude ξ from these using Equations (4.70) to (4.72).

If we combine this two-mode cavity with an additional $-\pi/2$ phase shift for the signal input, then for a low control input $\epsilon_b = 0$ any signal with amplitude close to ϵ in magnitude will just be scattered without a phase shift. For a high control input $\epsilon_b = \xi$ that same input would pick up a π phase shift. We can wrap this phase shifter in an interferometer to realize an all-optical router equivalent to a so called 'Fredkin gate'.

4.5.2. An all-optical Fredkin gate

As mentioned above, we wrap the two-mode cavity with parameters chosen according to the previous subsection 4.5.1 with $\epsilon = 0$ and a subsequent $-\pi/2$ phase shifting element in a Mach-Zehnder interferometer. In Figure 4.10 we show some numerical simulation results for a switch constructed with the above scheme.

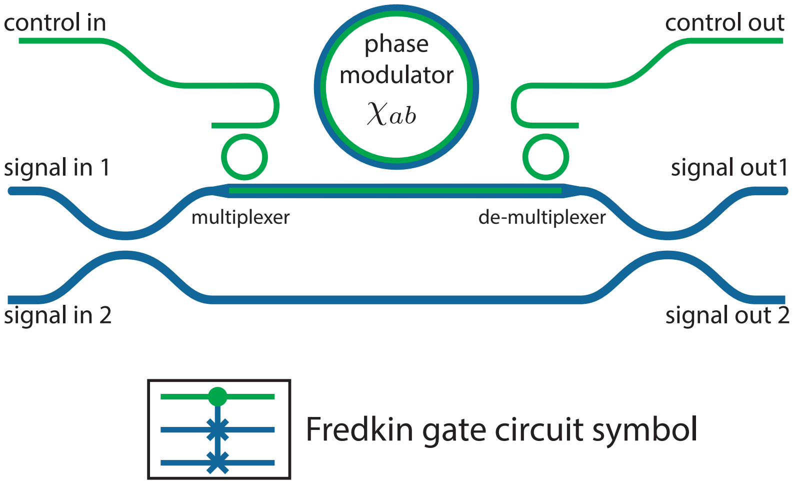


Figure 4.9.: **An optical Fredkin gate** based on having two non-degenerate modes cross phase modulate in a doubly resonant Kerr-cavity. Based on the control input, the Kerr-cavity imparts a phase shift of 0 or π. The control and signal inputs are combined by multiplexing and demultiplexing elements and the overall signal input/output path is wrapped in a Mach-Zehnder interferometer to enable controlled switching.

4.5.3. Two-mode-thresholder

The modified Fredkin-gate (with a non-zero ϵ) can act as a thresholding device. It turns out that the switching behavior is relatively robust to slight deviations in the control input power, and, more importantly, the threshold for switching is actually larger than at half the high input amplitude. This enables us to use the Fredkin gate with a single constant signal input (now acting more like a logical/binary value) and a continuously variable control input as a thresholder with two inverted signal outputs $(\epsilon_b > \epsilon_b^{th})$ and $\neg(\epsilon_b > \epsilon_b^{th}) = (\epsilon_b \leq \epsilon_b^{th})$.

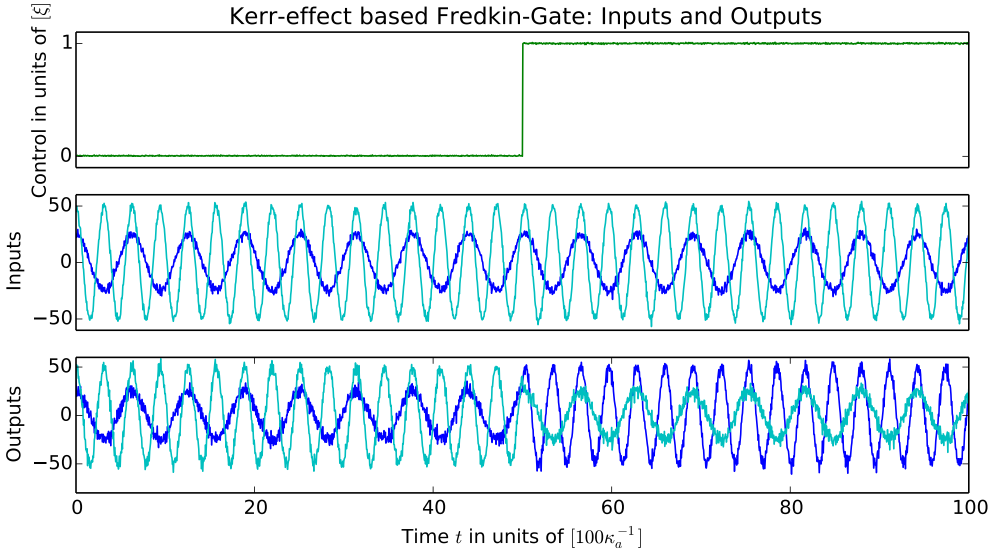


Figure 4.10.: When the control input (green trace) switches from zero to its 'on' state, the signal inputs (second plot) are switched on the outputs (last plot).

For a thresholder we may wish to bias the control input with an amplitude ξ_{bias} such that the overall circuit acts as a 50% beamsplitter for the signal ports. This marks the halfway point between the two switch states. This is the case when the signal mode is dynamically resonant with in the two-mode cavity:

$$0 = \Delta_a + \chi_a n(\epsilon, \xi_{bias}) + \chi_{ab} m(\epsilon, \xi_{\rm bias}) \quad \Rightarrow n(\epsilon, \xi_{\rm bias}) = \frac{4|\epsilon|^2}{\kappa_a} \tag{4.79}$$

$$\Rightarrow m(\epsilon, \xi_{\rm bias}) = -\frac{\Delta_a + \frac{4\chi_a|\epsilon|^2}{\kappa_a}}{\chi_{ab}} \tag{4.80}$$

$$\Rightarrow |\xi_{\rm bias}|^2 = \left[\frac{\kappa_b^2}{4} + (\Delta_b + \chi_b m(\epsilon, \xi_{\rm bias}) + \chi_{ab} n(\epsilon, \xi_{\rm bias}))^2\right] \frac{m(\epsilon, \xi_{\rm bias})}{\kappa_b} . \tag{4.81}$$

4.6. Non-degenerate optical parametric oscillators

Both degenerate and non-degenerate optical parametric oscillators (DOPOs and NOPOs, respectively) are very interesting systems to study as they exhibit critical dynamical points and multi-stability making them useful candidates for applications in signal processing and creating optical memories. We focus here on the non-degenerate case, as this will turn out to be useful in the next chapter where we construct a device with continuous memory. The math of the DOPO is very similar to the NOPO case, but it only exhibits a discrete bi-stability.

The basic model we are considering here is given by a cavity with three modes, a pump field c, a signal a and an idler mode b. The signal and idler resonance frequencies add up to the pump frequency $\omega_c = \omega_a + \omega_b$ and there is a χ_2- non-linearity that allows for conversion of a pump photon into a pair of signal and idler photons and vice versa.

A basic SLH model would then be

$$S = \mathbf{1}, \quad L_a = (\sqrt{\kappa_a} a, \sqrt{\kappa_b} b, \sqrt{\kappa_c} c)^T , \tag{4.82}$$

$$H = \Delta_a a^\dagger a + \Delta_b b^\dagger b + i \left(\chi a^\dagger b^\dagger c - \chi^* abc^\dagger \right) . \tag{4.83}$$

Going to a semi-classical description we introduce the complex mode amplitudes $(\alpha, \beta, \sigma) \leftrightarrow (\langle a \rangle , \langle b \rangle , \langle c \rangle)$. The drift component of their generally stochastic dynamics is then given by:

$$\dot{\alpha} = - \left(\frac{\kappa_a}{2} + i\Delta_a \right) \alpha + \chi \beta^* \sigma , \tag{4.84}$$

$$\dot{\beta} = - \left(\frac{\kappa_b}{2} + i\Delta_b \right) \beta + \chi \alpha^* \sigma , \tag{4.85}$$

$$\dot{\sigma} = - \frac{\kappa_c}{2} \sigma - \chi^* ab . \tag{4.86}$$

As they are, the equations of motion are invariant under any simultaneous transformation of modes and the χ parameter:

$$\alpha \to \alpha e^{i\phi}, \beta \to \beta e^{i\theta}, \sigma \to \sigma e^{i\nu}, \chi \to \chi e^{i(\phi + \theta - \nu)} . \tag{4.87}$$

We can fix $\phi + \theta = - \arg \chi$ to enforce $\chi \in \mathbb{R}_{\geq 0}$ and still retain the original $U(1) \times U(1)$ symmetry for the modes:

$$\alpha \to \alpha e^{i\phi + i\nu}, \beta \to \beta e^{-i\phi + i\nu}, \sigma \to \sigma e^{i\nu} . \tag{4.88}$$

Finally, we will want to include an external driving term for the pump. This changes $L_c \to L_c + \epsilon$ and $H \to H + \frac{\sqrt{\kappa_c}}{2i} \left(\epsilon c^\dagger - \epsilon^* c \right)$. The equation of motion for the pump is modified as

$$\dot{\sigma} \to \dot{\sigma} - \sqrt{\kappa_c} \epsilon . \tag{4.89}$$

We can now use the symmetry of the problem to fix the phase of σ such that $\epsilon \in \mathbb{R}_{\geq 0}$. With this the model becomes

$$\dot{\alpha} = - \left(\frac{\kappa_a}{2} + i\Delta_a \right) \alpha + \chi \beta^* \sigma , \tag{4.90}$$

$$\dot{\beta} = - \left(\frac{\kappa_b}{2} + i\Delta_b \right) \beta + \chi \alpha^* \sigma , \tag{4.91}$$

$$\dot{\sigma} = - \frac{\kappa_c}{2} \sigma - \chi^* ab - \sqrt{\kappa_c} \epsilon , \tag{4.92}$$

where we are free to choose the phase of χ and ϵ as we want to and we still retain a single $U(1)$ symmetry

$$\alpha \to \alpha e^{i\phi}, \beta \to \beta e^{-i\phi}. \tag{4.93}$$

4.6.1. Eliminated pump model

Often, we will consider a case in which the pump decay rate is much larger than the signal and idler decay rates $\kappa_c \gg \kappa_{a/b}$. In this case we can adiabatically eliminate the pump dynamics (as it becomes "slaved" to the signal and idler) and fix:

$$\sigma \approx -\frac{2}{\kappa_c}\left(\chi^* ab + \sqrt{\kappa_c}\epsilon\right) \, . \tag{4.94}$$

The reduced model in this case (and treating the mode variables as commutative) is given by

$$\dot{\alpha} = -\left(\frac{\kappa_a}{2} + \frac{2|\chi|^2}{\kappa_c}\beta^*\beta + i\Delta_a\right)\alpha - \frac{2\chi\epsilon}{\sqrt{\kappa_c'}}\beta^* \, , \tag{4.95}$$

$$\dot{\beta} = -\left(\frac{\kappa_b}{2} + \frac{2|\chi|^2}{\kappa_c}\alpha^*\alpha + i\Delta_b\right)\beta - \frac{2\chi\epsilon}{\sqrt{\kappa_c'}}\alpha^* \, . \tag{4.96}$$

This model retains the above remaining $U(1)$ symmetry. For convenience, we will rewrite this model as

$$\dot{\alpha} = -\left(\frac{\kappa_a}{2} + \gamma\beta^*\beta + i\Delta_a\right)\alpha - \eta\beta^* \tag{4.97}$$

$$\dot{\beta} = -\left(\frac{\kappa_b}{2} + \gamma\alpha^*\alpha + i\Delta_b\right)\beta - \eta\alpha^* \tag{4.98}$$

4.6.2. Fixpoints and stability

The general Jacobian of the full model is given by

$$J(\alpha,\alpha^*,\beta,\beta^*,\sigma,\sigma^*) = \begin{pmatrix} -\left(\frac{\kappa_a}{2}+i\Delta_a\right) & 0 & 0 & \chi\sigma & \chi\beta^* & 0 \\ 0 & -\left(\frac{\kappa_a}{2}-i\Delta_a\right) & \chi^*\sigma^* & 0 & 0 & \chi^*\beta \\ 0 & \chi\sigma & -\left(\frac{\kappa_b}{2}+i\Delta_b\right) & 0 & \chi\alpha^* & 0 \\ \chi^*\sigma^* & 0 & 0 & -\left(\frac{\kappa_b}{2}-i\Delta_b\right) & 0 & \chi^*\alpha \\ -\chi^*\beta & 0 & -\chi^*\alpha & 0 & -\frac{\kappa_c}{2} & 0 \\ 0 & -\chi\beta^* & 0 & -\chi\alpha^* & 0 & -\frac{\kappa_c}{2} \end{pmatrix} \tag{4.99}$$

The reduced model Jacobian is given by

$$J(\alpha,\alpha^*,\beta,\beta^*) = \begin{pmatrix} -\left(\frac{\kappa_a}{2}+\gamma|\beta|^2+i\Delta_a\right) & 0 & -\gamma\alpha\beta^* & -\eta-\gamma ab \\ 0 & -\left(\frac{\kappa_a}{2}+\gamma|\beta|^2-i\Delta_a\right) & -\eta^*-\gamma\alpha^*\beta^* & -\gamma\alpha^*\beta \\ -\gamma\alpha^*\beta & -\eta-\gamma ab & -\left(\frac{\kappa_b}{2}+\gamma|\alpha|^2+i\Delta_b\right) & 0 \\ -\eta^*-\gamma\alpha^*\beta^* & -\gamma\alpha\beta^* & 0 & -\left(\frac{\kappa_b}{2}+\gamma|\alpha|^2-i\Delta_b\right) \end{pmatrix} . \tag{4.100}$$

4.6.3. Below threshold

Full model

A fixpoint that always exists is given by $\alpha = \beta = 0, \sigma = -\frac{2\epsilon}{\sqrt{\kappa_c}}$. Inserting this into the Jacobian, find that if we arrange the variables in pairs as $(\delta\alpha, \delta\beta^*), (\delta\alpha^*, \delta\beta), (\delta\sigma, \delta\sigma^*)$ there is no cross-coupling between them in the sub-threshold Jacobian. The pump mode has a trivial Eigenspace with a single degenerate eigenvalue $-\frac{\kappa_c}{2}$. The Eigenvalues of the first pair are complex conjugates of the second, so it suffices to just look at one case:

$$J(\alpha, \beta^*) = \begin{pmatrix} A & C \\ C^* & B \end{pmatrix} \tag{4.101}$$

with $A = -\left(\frac{\kappa_a}{2} + i\Delta_a\right)$, $B = -\left(\frac{\kappa_b}{2} - i\Delta_b\right)$ and $C = \chi\sigma$. Then the eigenvalues are given by

$$\lambda_\pm = \frac{1}{2}\left(A + B \pm \sqrt{(A-B)^2 + 4|C|^2}\right) \tag{4.102}$$

Now the real parts of A and B and it can be shown that only λ_- can cross into the right half plane. This happens when

$$\kappa_a + \kappa_b \leq 2\mathrm{Re}\sqrt{\frac{(\kappa_a - \kappa_b)^2}{4} - (\Delta_a + \Delta_b)^2 + 4|\chi\sigma|^2 + \frac{i}{2}(\Delta_a + \Delta_b)(\kappa_a - \kappa_b)} \tag{4.103}$$

Since σ is directly proportional to the pump drive, we see that there always exists a critical pump driving strength at which this Eigenvalue becomes unstable. On resonance $\Delta_a = \Delta_b = 0$ and for equal decay rates $\kappa_a = \kappa_b = \kappa$ this simplifies to $|\sigma| \geq \frac{\kappa}{2|\chi|} \Leftrightarrow |\epsilon| \geq \frac{\kappa\sqrt{\kappa_c}}{4|\chi|}$. Off resonance and for a less symmetric model, the condition is modified.

Below we will derive a much simpler expression for the threshold.

Reduced model

The reduced model also has a fixpoint at $\alpha = \beta = 0$ and if we insert this into the reduced model Jacobian we end up with an equivalent threshold condition as above, now in terms of the effective pump parameter $|\eta| \propto |\epsilon|$.

4.6.4. Above threshold

In steady state we must have

$$0 = -\left(\frac{\kappa_a}{2} + i\Delta_a\right)\alpha + \chi\beta^*\sigma \tag{4.104}$$

$$0 = -\left(\frac{\kappa_b}{2} + i\Delta_b\right)\beta + \chi\alpha^*\sigma \tag{4.105}$$

$$0 = -\frac{\kappa_c}{2}\sigma - \chi^* ab - \sqrt{\kappa_c}\epsilon \tag{4.106}$$

From the first two equations we find

$$\alpha = \frac{|\chi\sigma|^2}{(\kappa_a/2 + i\Delta_a)(\kappa_b/2 - i\Delta_b)}\alpha \tag{4.107}$$

This is either solved by setting $\alpha = 0$ which leads to the above sub-threshold solution or by requiring that the coefficient on the RHS is one:

$$\alpha \neq 0 \Rightarrow |\chi\sigma|^2 = (\kappa_a/2 + i\Delta_a)(\kappa_b/2 - i\Delta_b) \tag{4.108}$$

$$= \frac{\kappa_a\kappa_b}{4} + \Delta_a\Delta_b + \frac{i}{2}(\Delta_a\kappa_b - \Delta_b\kappa_a) \tag{4.109}$$

The first thing to note is that this equation can only be solved if the imaginary part of the RHS is set to zero. This poses a constraint on the model parameters:

$$\Delta_a\kappa_b = \Delta_b\kappa_a \tag{4.110}$$

If the constraint is satisfied, the real part of the equation gives a more precise threshold condition

$$|\chi\sigma|^2 = \frac{\kappa_a\kappa_b}{4} + \Delta_a\Delta_b. \tag{4.111}$$

Note that this equality holds *at and above* threshold. Any additional pump driving power is reflected or converted to signal and idler photons.

If the constraint isn't satisfied, then note that we can always transform into a rotating frame where it is: Applying a tranformation $U(\phi) = \exp(-i\phi(a^\dagger a - b^\dagger b))$ to the mode operators realizes the transformation

$$U(\phi)aU^\dagger(\phi) = a - i\phi[a^\dagger a, a] + \cdots = ae^{i\phi}, \quad U(\phi)bU^\dagger(\phi) = be^{-i\phi} \tag{4.112}$$

If we now substitute a time dependent angle $\phi = \omega t$ then this modifies the effective detunings of a and b as

$$\Delta_{a/b} \to \Delta_{a/b} \pm \omega \tag{4.113}$$

Inserting this into the constraint, we can solve for the angular frequency to find the transformation that will take us in the correct rotating frame

$$(\Delta_a + \omega)\kappa_b = (\Delta_b - \omega)\kappa_a \tag{4.114}$$

$$\Leftrightarrow \omega = -\frac{\Delta_a\kappa_b - \Delta_b\kappa_a}{\kappa_a + \kappa_b}. \tag{4.115}$$

The new effective detunings are then

$$\Delta'_{a/b} = \Delta_{a/b} \mp \frac{\Delta_a\kappa_b - \Delta_b\kappa_a}{\kappa_a + \kappa_b} = \kappa_{a/b}\frac{\Delta_a - \Delta_b}{\kappa_a + \kappa_b} \tag{4.116}$$

and it's easy to see that they satisfy the constraints.

Plugging this in above gives us the threshold in terms of the original detunings:

$$|\chi\sigma|^2 = \frac{\kappa_a\kappa_b}{4} + \Delta'_a\Delta'_b \tag{4.117}$$

$$= \frac{\kappa_a\kappa_b}{4}\left[1 + 4\left(\frac{\Delta_a - \Delta_b}{\kappa_a + \kappa_b}\right)^2\right] \tag{4.118}$$

Or, reexpressing this in terms of the pump driving amplitude we find that to be at or above threshold we need

$$|\epsilon|^2 \geq \frac{\kappa_a\kappa_b\kappa_c}{16|\chi|^2}\left[1 + 4\left(\frac{\Delta_a - \Delta_b}{\kappa_a + \kappa_b}\right)^2\right]. \tag{4.119}$$

In general the detuning constraint is not satisfied and thus the system will start oscillating at frequency $\omega = -\frac{\Delta_a\kappa_b - \Delta_b\kappa_a}{\kappa_a + \kappa_b}$, but with a constant amplitude for signal and idler. The case $\omega = 0$ always corresponds to looking at the system within the appropriate rotating frame. This frame is ideally suited to studying phase locking because all explicit time dependence has been removed by the transformation.

The above threshold condition does not just mark where the oscillating solution starts to exist, but also where the sub-threshold solution becomes unstable. The oscillating solution bifurcates from this sub-threshold solution as we will now show by explicitly computing the signal/idler mode amplitudes as a function of the pump drive.

To do this, assume again that the detuning constraint is again satisfied $\Delta_a\kappa_b = \Delta_b\kappa_a$, then we can derive the following closed set of relation ship between α and β^*:

$$0 = -\left(\Gamma_a + i\Delta_a\right)\alpha - \eta\beta^* \tag{4.120}$$

$$0 = -\left(\Gamma_b - i\Delta_b\right)\beta^* - \eta^*\alpha \tag{4.121}$$

where we have introduced the power dependent decay rates $\Gamma_a = \frac{\kappa_a}{2} + \gamma|\beta|^2$ and $\Gamma_b = \frac{\kappa_b}{2} + \gamma|\alpha|^2$. Combining these two equations we find

$$\alpha = \frac{|\eta|^2}{(\Gamma_a + i\Delta_a)(\Gamma_b - i\Delta_b)}\alpha, \tag{4.122}$$

which for non-zero $\alpha \neq 0$ leads us to to find

$$\Delta_a\Gamma_b = \Delta_b\Gamma_a \tag{4.123}$$

$$\Gamma_a\Gamma_b = |\eta|^2 - \Delta_a\Delta_b \tag{4.124}$$

Using the constraint on the detunings we find

$$\Gamma_{a/b} = \kappa_{a/b}\sqrt{\frac{|\eta|^2 - \Delta_a\Delta_b}{\kappa_a\kappa_b}} \tag{4.125}$$

Using the relationship between the decay rates and the intra-cavity amplitudes we finally find

$$|\alpha|^2 = \frac{\Gamma_b - \kappa_b/2}{\gamma} = \frac{\kappa_b}{\gamma}\left(\sqrt{\frac{|\eta|^2 - \Delta_a\Delta_b}{\kappa_a\kappa_b}} - \frac{1}{2}\right) \tag{4.126}$$

$$|\beta|^2 = \frac{\Gamma_a - \kappa_a/2}{\gamma} = \frac{\kappa_a}{\gamma}\left(\sqrt{\frac{|\eta|^2 - \Delta_a\Delta_b}{\kappa_a\kappa_b}} - \frac{1}{2}\right). \tag{4.127}$$

This implies that $\frac{|\alpha|^2}{|\beta|^2} = \frac{\kappa_b}{\kappa_a}$ which can be rewritten as $\kappa_a|\alpha|^2 = \kappa_b|\beta|^2$ in which form it shows that the output power from signal and idler is equal.

The phases of α and β

The only thing left to determine are the phases of signal and idler. Due to the symmetry $\alpha \to \alpha e^{i\phi}, \beta \to \beta^{-i\phi}$ we will compute the phase of an invariant under this transformation given by ab. This can be obtained from

$$\beta^* = -\frac{\eta^*}{\Gamma_b - i\Delta_b}\alpha \tag{4.128}$$

$$\Leftrightarrow |\beta|^2 = -\frac{\eta^*}{\Gamma_b - i\Delta_b}ab \tag{4.129}$$

If we restrict ourselves to the special (but still quite general) case $\eta \in \mathbb{R}_{<0}$, then we must have

$$\Rightarrow \arg ab = \arg\left(1 - i\frac{\Delta_a}{\Gamma_a}\right) = \arg\left(1 - i\frac{\Delta_b}{\Gamma_b}\right) \tag{4.130}$$

$$= \arctan \frac{\Delta_{a/b}}{\kappa_{a/b}}\sqrt{\frac{\kappa_a\kappa_b}{\eta^2 - \Delta_a\Delta_b}} \tag{4.131}$$

This shows that for $\eta < 0$ and on resonance, we will always have $ab \in \mathbb{R}_{>0}$, i.e., the two-mode amplitudes are of opposite complex phase.

The resonant case $\Delta_{a/b} = 0, \eta, \epsilon < 0, \chi > 0$

As we've already seen, the phases of the modes are not only perfectly correlated but also exactly opposite each other. Moreover, the mode amplitudes in this case can be expressed as:

$$|\alpha|^2 = \frac{\kappa_b}{\gamma}\left(\frac{|\eta|}{\sqrt{\kappa_a\kappa_b}} - \frac{1}{2}\right) \tag{4.132}$$

$$|\beta|^2 = \frac{\kappa_a}{\gamma}\left(\frac{|\eta|}{\sqrt{\kappa_a\kappa_b}} - \frac{1}{2}\right). \tag{4.133}$$

Furthermore, we know that $\sigma = \frac{\sqrt{\kappa_a\kappa_b}}{2\chi}$.

4.7. A Bifurcating Kerr Amplifier

There has rececently been a lot of interest in realizing Coherent Ising Machines (CIM) [121, 131] in which discrete Ising 'spins' are encoded in individual phase-bistable modes or competing polarization modes of a nonlinear photonic network. The problem of finding the ground state of an Ising problem can then be related to finding the 'least stable' or 'maximum gain' collective supermode of such a network, as some gain mechanism is continuously turned on [121, 131]. Existing proposals and experiments have so far either focused on free-space networks of injection locked lasers or networks of parametric oscillators [121, 131, 71, 47, 49, 53, 109]. The most scalable approaches to date have been realized using time-multiplexed pulsed degenerate parametric oscillators with delay line coupled or measurement based feedback induced interactions.

During my work in Ray Beausoleil's Large Scale Integrated Photonics (LSIP) group at Hewlett Packard Labs, I proposed a different physical design for realizing a single 'spin' degree of freedom that I will outline in this section. The full Ising circuit model, an evaluation of its performance and a discussion of its expected robustness to fabrication error will be published in [118].

To make an Ising machine truly portable and mass-manufacturable, it would be convenient to have a realization as an integrated photonic system. Furthermore, it would be desirable (at least in initial devices) to require only a single optical wavelength for operating the Ising machine. As silicon features relatively strong dispersive nonlinearity (thermo-optic nonlinearity, free-carrier dispersion and Kerr-nonlinearity) we will here describe how to use dispersively nonlinear circuit elements for creating the individual Ising 'spin' devices. Like a degenerate parametric oscillator our design features a continuous bifurcation below which the bifurcating mode is very sensitive to external perturbing fields.

We construct our spin from an extended version of the above described symmetric amplifier (cf. Section 4.3) with an additional lossy self-feedback path on the bias input and output ports. As we will describe below, the bias feedback ensures that only two symmetric states exist above this bias power threshold with anti-correlated internal resonator states. In the absence of fabrication errors, such a device would in principle exhibit a pure pitchfork bifurcation as the pump drive increases above its threshold level. This is analogous to the bifurcation occurring in a degenerate optical parametric oscillator. Thus, our design mimics that of [131], but with the advantage that the 'pump' input field is of the same wavelength as the 'signal'.

If the resonators are detuned beyond the bi-stability threshold detuning $|\Delta| > \Delta_{\text{th}} = \frac{\sqrt{3}\kappa_T}{2}$ the linear gain diverges at a specific bias power threshold beyond which the resonators exhibit bi-stability. We label the low and high power bi-stable resonator states as $|\alpha_j| \in \{\alpha_{\text{lo}}, \alpha_{\text{hi}}\}$, $j = 1, 2$. The amplifier MZI is modified by an additional feedback path connecting the bias output back to bias input. Without feedback, an independent pair of bi-stable resonators would exhibit $2 \times 2 = 4$ different meta-stable states $(|\alpha_1|, |\alpha_2|) \in \{(\alpha_{\text{lo}}, \alpha_{\text{lo}}), (\alpha_{\text{hi}}, \alpha_{\text{lo}}), (\alpha_{\text{lo}}, \alpha_{\text{hi}}), (\alpha_{\text{hi}}, \alpha_{\text{hi}})\},$. By an appropriately chosen bias feedback phase, two of the meta-stable states are removed such that the two resonators can only assume anti-correlated internal states $(|\alpha_1|, |\alpha_2|) \in \{(\alpha_{\text{hi}}, \alpha_{\text{lo}}), (\alpha_{\text{lo}}, \alpha_{\text{hi}})\}$, Figure 4.11 has a schematic visualizing its construction.

We assume the symmetric open loop amplifier model with ODE's given by

$$\dot{\alpha}_1 = -\left[\kappa_T/2 + i\left(\Delta + \chi|\alpha_1|^2\right)\right]\alpha_1 - \sqrt{\kappa/2}\left(\beta_1 + \beta_2\right) - \sqrt{\kappa_L}\eta_1 \tag{4.134}$$

$$\dot{\alpha}_2 = -\left[\kappa_T/2 + i\left(\Delta + \chi|\alpha_2|^2\right)\right]\alpha_2 - \sqrt{\kappa/2}\left(\beta_1 - \beta_2\right) - \sqrt{\kappa_L}\eta_2, \tag{4.135}$$

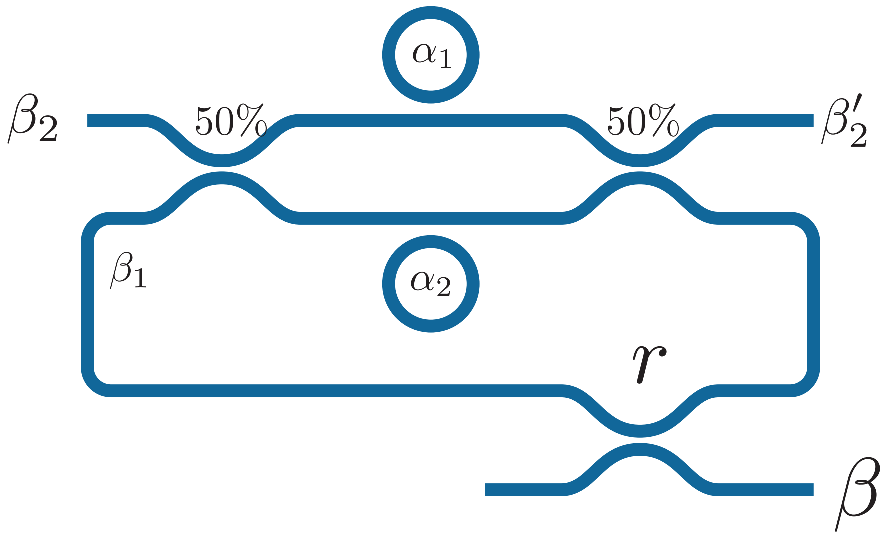


Figure 4.11.: Schematic for a tunable optical amplifier with self-feedback (TAFB): Two identical microring resonators with dispersive optical non-linearity are placed in the arms of a Mach-Zehnder interferometer formed by waveguides and either directional couplers (DCs) or multimode interference devices (MMIs). One of the interferometer's outputs is fed back to the input to selectively modify the resonance properties of the symmetric and asymmetric supermodes of the resonator pair. Additional couplers on the second input and output allow injecting bias fields for the spin variable as well as measure the current spin amplitude.

and the input-output relationship given by

$$\beta_1' = \beta_1 + \sqrt{\kappa/2}\,(\alpha_1 + \alpha_2) \tag{4.136}$$

$$\beta_2' = \beta_2 + \sqrt{\kappa/2}\,(\alpha_1 - \alpha_2)\,, \tag{4.137}$$

where we have assumed that the first MMI's scattering matrix is given by $S_1 = \frac{1}{\sqrt{2}}\begin{pmatrix}1 & 1\\ 1 & -1\end{pmatrix}$ and the second MMI's scattering matrix is given by $S_2 = S_1^{-1}$. For theoretical simplicity we assume no scattering losses in the MMIs but in our numerical analysis we allow for such losses as well.

We *do* assume internal resonator losses κ_L and summarize the total decay rate as the sum of the internal loss and the waveguide coupling $\kappa_T = \kappa + \kappa_L$. The $\eta_{1/2}$ are pure vacuum noise inputs whereas $\beta_{1/2}$ are a constant bias and time varying signal input to the system.

We see that $d\beta_1$ couples to the common mode $\alpha_+ = \frac{\alpha_1+\alpha_2}{\sqrt{2}}$ and $d\beta_2$ couples to the difference mode $\alpha_- = \frac{\alpha_1 - \alpha_2}{\sqrt{2}}$. We can re-express the SDEs in terms of these supermodes as

$$\begin{aligned} \dot{\alpha}_+ &= -\left[\kappa_T/2 + i\Delta + i\chi/2\left(|\alpha_+|^2 + |\alpha_-|^2\right)\right]\alpha_+ \\ &\quad - i\chi \mathrm{Re}\,(\alpha_+\alpha_-^*)\alpha_- dt - \sqrt{\kappa}\beta_1 - \sqrt{\kappa_L}\eta_+ \qquad (4.138) \\ \dot{\alpha}_- &= -\left[\kappa_T/2 + i\Delta + i\chi/2\left(|\alpha_+|^2 + |\alpha_-|^2\right)\right]\alpha_- \\ &\quad - i\chi \mathrm{Re}\,(\alpha_+\alpha_-^*)\alpha_+ - \sqrt{\kappa}\beta_2 - \sqrt{\kappa_L}\eta_- \qquad (4.139) \end{aligned}$$

and the input-output relationship given by

$$\beta_1' = \beta_1 + \sqrt{\kappa}\alpha_+ \tag{4.140}$$

$$\beta_2' = \beta_2 + \sqrt{\kappa}\alpha_-. \tag{4.141}$$

We now apply feedback of the common mode to itself by imposing $\beta_1 = r\underbrace{\left(\sqrt{\kappa}\alpha_+ + \beta_1\right)}_{\beta_1'} + \sqrt{1-|r|^2}\beta$, this corresponds to inserting a beamsplitter with unitary mixing matrix $\begin{pmatrix} r & \sqrt{1-|r|^2} \\ -\sqrt{1-|r|^2} & r^* \end{pmatrix}$ (with $r \in \mathbb{C}, |r| < 1$) in the feedback and adding an input $d\beta$ to the second beamsplitter input.

Solving this for the in-loop field amplitude we find

$$\beta_1 = \frac{r\sqrt{\kappa}\alpha_+ + \sqrt{1-|r|^2}\beta}{1-r}. \tag{4.142}$$

and inserting this into the SDE for α_+ yields

$$\begin{aligned} \dot{\alpha}_+ &= -\left[\kappa_T/2 + \frac{\kappa r}{1-r} + i\Delta + i\chi/2\left(|\alpha_+|^2 + |\alpha_-|^2\right)\right]\alpha_+ \\ &\quad - i\chi \mathrm{Re}\,(\alpha_+\alpha_-^*)\alpha_- - \frac{\sqrt{\kappa}\sqrt{1-|r|^2}\beta}{1-r} - \sqrt{\kappa_L}\eta_+ \qquad (4.143) \end{aligned}$$

Thus, the feedback can be understood as modifying the input coupling rate and detuning of the common mode as

$$\kappa \to \kappa\left(1 + 2\mathrm{Re}\frac{r}{1-r}\right) = \kappa \mathrm{Re}\frac{1+r}{1-r} = \kappa' \tag{4.144}$$

$$\Delta \to \Delta + \kappa \mathrm{Im}\frac{r}{1-r} = \Delta' \tag{4.145}$$

We see that for a real scattering parameter r the detuning is unmodified. Furthermore, the effective coupling rate of the cavity can be made arbitrarily large or small.

Assuming a vanishing signal input $\beta_2 = 0$ and constant bias input $\beta = \beta_0$, if the system is stable with a *unique* fixpoint, then the symmetry of (4.143) requires that $\alpha_- = 0$.

We can then solve for the steady state common mode amplitude via

$$0 = -\left[\kappa_T'/2 + i\Delta' + i\chi|\alpha_+|^2/2\right]\alpha_+ - \sqrt{\kappa'}\,\beta_0 \tag{4.146}$$

which is identical to the steady state relationship of a single Kerr-cavity with somewhat modified parameters. We express bias amplitudes relative to the amplitude at which the inflection point of the common mode photon number vs. drive amplitude occurs:

$$\beta_{\max} = \sqrt{\frac{{\kappa_T'}^2/4 + (\Delta' + \chi/2n_{+,0})^2)n_{+,0}}{\kappa'}}, \tag{4.147}$$

where $n_{+,0} = \frac{4|\Delta|}{3|\chi|}$.

A fixpoint with $\alpha_- = 0$ solving the above relations always exists, but it is not necessarily stable. The linearization $\alpha_+ \to \alpha_+ + \delta\alpha_+, \quad \alpha_- \to \delta\alpha_-$ of the ODE yields the decoupled equations

$$\begin{pmatrix} \dot{\delta\alpha_+} \\ \dot{\delta\alpha_-} \end{pmatrix} = \begin{pmatrix} -\left[\kappa_T'/2 + i\Delta' + i\chi|\alpha_+|^2\right]\delta\alpha_+ - i\chi{\alpha_+}^2/2\delta\alpha_+^* \\ -\left[\kappa_T/2 + i\Delta + i\chi|\alpha_+|^2\right]\delta\alpha_- - i\chi{\alpha_+}^2/2\delta\alpha_-^* \end{pmatrix} \tag{4.148}$$

Comparing these with the single Kerr cavity case we see that the fixpoint is stable at all input amplitudes if either $\Delta/\chi \geq 0$ or if we have both $|\Delta| < \sqrt{3}\kappa_T/2$ and $|\Delta'| < \sqrt{3}\kappa_T'/2$.

When these conditions are violated, there generally exists a range of input amplitudes in which the symmetric fix point is unstable, but it is always stable for sufficiently small or large input amplitudes $|\beta_0|$.

Let now investigate fixpoints where α_+ is constant and α_- is bi-stable. From the symmetry of the steady state equations we can infer that (α_+, α_-) is a fixpoint iff $(\alpha_+, -\alpha_-)$ is a fixpoint. Furthermore, we can assume that $\alpha_+ \in \mathbb{R}_{>0}$ since we can always adjust the bias input phase to achieve this. We then find that

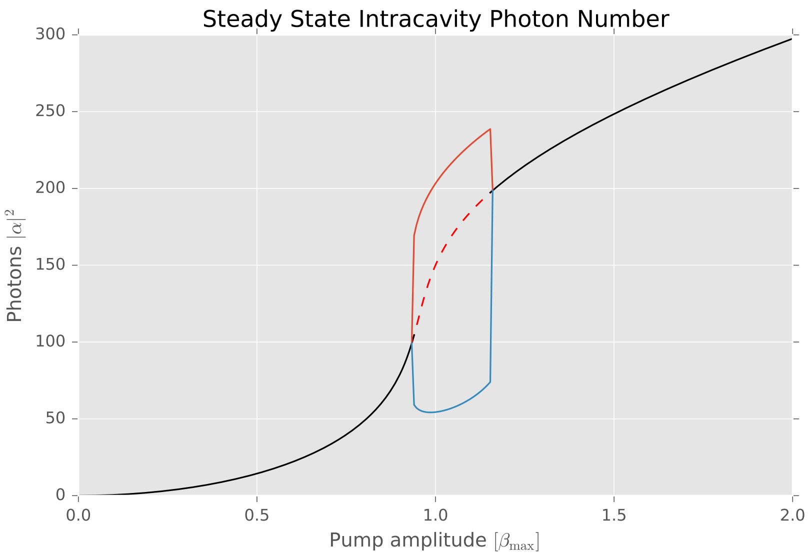
$$0 = \left[\kappa_T/2 + i\Delta + i\chi/2\left(2|\alpha_+|^2 + |\alpha_-|^2\right)\right]\alpha_- + i\chi/2\alpha_+^2\alpha_-^*$$

Using this relationship we can show that a solution with non-zero α_- exists for $\left||\alpha_+|^2 - \frac{4|\Delta|}{3|\chi|}\right| \leq \frac{\sqrt{4\Delta^2 - 3\kappa_T^2}}{3|\chi|}$. For the RHS to be real we furthermore need that $\Delta/\chi \leq 0$ and $|\Delta| \geq \sqrt{3}\kappa_T/2$. If these conditions are met, then we have

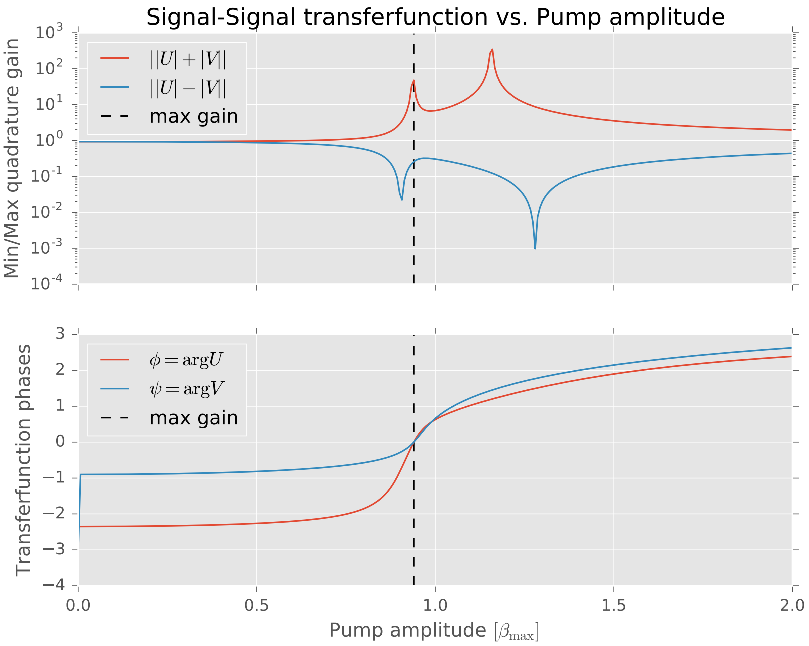
$$|\alpha_-|^2 = 2\left|\frac{\Delta}{\chi}\right| + \sqrt{|\alpha_+|^4 - \frac{\kappa_T^2}{\chi^2}} - 2|\alpha_+|^2. \tag{4.149}$$

We already know that $|\Delta| \geq \sqrt{3}\kappa_T/2$ guarantees that the $\alpha_- = 0$ solution becomes unstable. In principle we could also expect α_+ to have multiple possible solutions with potentially different stability, but we intuitively expect that this is not the case as long as $|\Delta'| < \sqrt{3}\kappa_T'/2$.

In Figures 4.12 (a) and (b) I have plotted the bi-stable steady state amplitudes as well as the magnitude and phase of the coefficients of the steady state signal-signal transfer function of our device. In particular, the output signal mode satisfies $|\beta_2|' = |U\beta_2 + V\beta_2^*|$ and thus exhibits phase sensitive gain that diverges at a specific pump amplitude.



(a) Steady States



(b) Transfer Function

Figure 4.12.: The steady states of the tunable amplifier with self-feedback exhibits a bifurcation in which the differential mode becomes unstable. The blue and green curves show the upper and lower power state that exists in each respective resonator. The linear gain diverges at the bifurcation points. The TAFB is wrapped in phase shifters such that the real signal input quadrature is scattered (and amplified) to the real signal output quadrature right at the first bifurcation point.

4.8. Final remarks on optical computing

In this chapter we have studied a variety of optical components and circuits and provided some insight into their dynamical behavior. A popular design pattern for engineering computational devices is to synthesize a system with certain dynamical attractors that then play a computational role, e.g., linear or non-linear (saturable) amplification or discrete or continuous memory. Obviously this strategy has been around since the early days of engineering, but there are good reasons to assume that biological computational systems behave quite similarly. Computational studies of neural networks trained for specific tasks have revealed a rich variety of dynamical behavior featuring separated timescales with fast (computational) dynamics and slow (memory) dynamics [106].

In the final chapter of this thesis I present a method for deriving coupled quantum-classical equations of motion between such oscillator mode amplitudes and a particular representation of quantum system's quantum state. In the presence of dissipation (and thus decoherence), many of the dynamical features of the semi-classical version of a given system are still present though certainly modified by the coupling to the quantum state.

5. A coherent perceptron for all-optical learning

This chapter was also published as [116].

Recent progress in integrated nanophotonic engineering [60, 25, 65, 95, 27, 32, 83, 20, 127, 98] has motivated follow-up proposals [69, 87] of nanophotonic circuits for all-optical information processing. While most of these focus on implementations of digital logic, we present here an approach to all-optical analog, *neuromorphic* computation and propose design schemes for a set of devices to be used as building blocks for large scale circuits.

Optical computation has been a long-time goal [1, 102], with research interest surging regularly after new engineering capabilities are attained [76, 77], but so far the parallel progress and momentum of CMOS based integrated electronics has outperformed all-optical devices.

In recent years we have seen rapid progress in the domain of machine learning, and artificial intelligence in general. Although most current 'big data'-applications are realized on digital computing architectures, there is now an increasing amount of computation done in specialized hardware such as GPUs. Specialized analog computational devices for solving specific subproblems more efficiently than possible with either GPUs or general purpose computers are being considered or already implemented by companies such as IBM, Google and HP and in academia, as well. [2, 80, 104, 131] Specifically in the field of neuromorphic computation, there has been impressive progress on CMOS based analog computation platforms [19, 17].

Several neuromorphic approaches to use complex nonlinear optical systems for machine learning applications have recently been proposed [26, 125, 124, 24] and some initial schemes have been implemented [62, 127]. So far, however, all of these 'optical reservoir computers' have still required digital computers to prepare the inputs and process the output of these devices with the optical systems only being employed as static nonlinear mappings for dimensional lifting to a high dimensional feature space [22], in which one then applies straightforward linear regression or classification for learning an input-output map. [129]

In this work, we address how the final stage of such a system, i.e., the linear classifier could be realized all-optically. We provide a universal scheme, i.e., independent of which particular kind of optical nonlinearity is employed, for constructing *tunable* all-optical, phase-sensitive amplifiers and then outline how these can be combined with self-oscillating systems to realize an optical amplifier with *programmable* gain, i.e., where the gain can be set once and is then fixed subsequently.

Using these as building blocks we construct an all-optical *perceptron* [96, 97], a system that can classify multi-dimensional input data and, using pre-classified training data learn the correct classification boundary 'on-line', i.e., incrementally. The perceptron can be seen as a highly simplified model of a neuron. While the idea of all-optical neural networks has been proposed before [75] and an impressive scheme using electronic, measurement-based feedback for spiking optical signals has been realized [30], to our knowledge, we offer the first complete description for how the synaptic weights can be stored in an optical memory and programmed via feedback.

The physical models underlying the employed circuit components are high intrinsic-Q optical resonators with strong optical nonlinearities. For theoretical simplicity we assume resonators with either a χ_2 or a χ_3 nonlinearity, but the design can be adapted to depend on only one of these two or alternative nonlinearities such as those based on free carrier effects or optomechanical interactions.

The strength of the optical nonlinearity and the achievable Q-factors of the optical resonators determine the overall power scale and rate at which a real physical device could operate. Both a stronger nonlinearity and higher Q allow operating at lower overall power.

We present numerical simulations of the system dynamics based on the semi-classical Wigner-approximation to the full coherent quantum dynamics presented in [98]. For photon numbers as low as ($\sim 10-20$) this approximation allows us to accurately model the effect of optical quantum shot noise even in large-scale circuits.

In the limit of both very high Q and very strong nonlinearity, we expect quantum effects to become significant as entanglement can arise between the field modes of physically separated resonators. In the appendix, we provide full quantum models for all basic components of our circuit. The possibility of a quantum speedup is being addressed in ongoing work. Recently, D-Wave Systems has generated a lot of interest in their own superconducting qubit based quantum annealer. Although the exact benefits of quantum dynamics in their machines has not been conclusively established [8], recent results analyzing the role of tunneling in a quantum annealer [9] are intriguing and suggest that quantum effects can be harnessed in computational devices that are not unitary quantum computers.

5.1. The Perceptron algorithm

The perceptron is a machine learning algorithm that maps an input $x \in \mathbb{R}^n$ to a single binary class label $\hat{y}_w[x] \in \{0, 1\}$. Binary classifiers generally operate by dividing the input space into two disjoint sets and identifying these with the class labels. The perceptron is a linear classifier, meaning that the surface separating the two class label sets is a linear space, a hyperplane, and its output is computed simply by applying a step function $\theta(u) := \mathbb{1}_{u \geq 0}$ to the inner product of a single data point x with a fixed *weight vector* w:

$$\hat{y}_w[x] := \theta(w^T x) = \begin{cases} 1 \text{ for } w^T x \geq 0, \\ 0 \text{ otherwise.} \end{cases} \tag{5.1}$$

Geometrically, the weight vector w parametrizes the hyperplane $\{z \in \mathbb{R}^n : \; w^T z = 0\}$ that forms the decision boundary.

In the above parametrization the decision boundary always contains the origin $z = 0$, but the more general case of an affine decision boundary $\{\tilde{z} \in \mathbb{R}^n : \; \tilde{w}^T \tilde{z} = b\}$ can be obtained by extending the input vector by a constant $z = (\tilde{z}^T, 1)^T \in \mathbb{R}^{n+1}$ and similarly defining an extended weight vector $w = (\tilde{w}^T, -b)^T$.

The perceptron converges in a finite number of steps for all linearly separable problems [96] by randomly iterating over a set of pre-classified training data $\{(y^{(j)}, x^{(j)}) \in \{0, 1\} \otimes \mathbb{R}^n, \; j = 1, 2, \ldots, M\}$ and imparting a small weight correction $w \to w + \Delta w$ for each falsely classified training example $x^{(j)}$

$$\Delta w = \tilde{\alpha} \left(y^{(j)} - \hat{y}_w[x^{(j)}] \right) x^{(j)}. \tag{5.2}$$

The *learning rate* $\tilde{\alpha} > 0$ determines the magnitude of the correction applied for each training example. The expression in parentheses can only take on the values $\{0, -1, 1\}$ with the zero corresponding to a correctly classified example and the non-zero values corresponding to the two different possible classification errors.

Usually there exist many separating hyperplanes for a given linear binary classification problem. The standard perceptron is only guaranteed to find one that works for the training set. It is possible to introduce a notion of optimality to this problem by considering the minimal distance (“margin”) of the training data to the found separating hyperplane. Maximization of this margin naturally leads to the “support vector machine” (SVM) algorithm [21]. Although the SVM outperforms the perceptron in many classification tasks it does not lend itself to a hardware implementation as readily because it cannot be trained incrementally. It is this that makes the perceptron algorithm especially suited for a hardware implementation: We can convert the discrete update rule (5.2) to a differential equation

$$\dot{w}(t) = \alpha \left\{ y(t) - \hat{y}_{w(t)}(t) \right\} x(t), \tag{5.3}$$

and then construct a physical system that realizes these dynamics. In this continuous-time version the inputs are piece-wise constant $x(t) = x^{(j_t)}$, $y(t) = y^{(j_t)}$ and take on the same discrete values as above indexed by $j_t := \lceil \frac{t}{\Delta t} \rceil \in \{1, 2, \dots, M = \frac{T}{\Delta t}\}$.

5.1.1. The circuit modeling framework

Circuits are fully described via Quantum Hardware Description Language (QHDL) [117] based on Gough and James' SLH-framework [41, 40]. To carry out numerical simulations for large scale networks, we derive a system of semi-classical Langevin equations based on the Wigner-transformation as described in [98]. Note that there is a perfect one-to-one correspondence between nonlinear cavity models expressed via SLH and the Wigner method as long as the nonlinearities involve only oscillator degrees of freedom. There is ongoing research in our group to establish similar results for more general nonlinearities [48].

Both the Wigner method and the more general SLH framework can be used to model networks of quantum systems where the interconnections are realized through bosonic quantum fields. The SLH framework describes a system interacting with n independent input fields in terms of a unitary scattering matrix S parametrizing direct field scattering, a coupling vector $L = (L_1, L_2, \dots, L_n)^T$ parametrizing how external fields couple into the system and how the system variables couple to the output and a Hamilton operator inducing the internal dynamics. We summarize these objects in a triplet (S, L, H). L and H are sufficient to parametrize any Schrödinger picture simulation of the quantum dynamics, e.g., the master equation for a mixed system state ρ is given by

$$\dot{\rho} = -i[H, \rho] + \sum_{j=1}^{n} \left(L_j \rho L_j^\dagger - \frac{1}{2}\{L_j^\dagger L_j, \rho\} \right). \tag{5.4}$$

The scattering matrix S is important when composing components into a network. In particular, the input-output relation in the SLH framework is given by

$$dA_{out} = S\, dA_{\mathrm{in}} + L\, dt, \tag{5.5}$$

where the $dA_{\mathrm{in}/out,j}$, $j = 1, 2, \dots, n$ are to be understood as quantum stochastic processes whose differentials can be manipulated via a quantum Ito calculus [41]. The Wigner method provides a simplified, approximate description which is valid when all non-linear resonator modes are in strongly displaced states [98]. The simulations presented here were carried out exclusively at energy scales for which the Wigner method is valid, allowing us to scale to much larger system sizes than we could in a full SLH-based quantum simulation. This is because the computational complexity of the Wigner method scales at most quadratically (and in sparsely interconnected systems nearly linearly) with the number of components as opposed to the exponential state space scaling of a quantum mechanical Hilbert space. We nonetheless provide our models in both Wigner-method form and SLH form in anticipation that our component models will also be extremely useful in the full quantum regime.

In the Wigner-based formalism, a system is described in terms of time-dependent complex coherent amplitudes $\alpha(t) = (\alpha_1(t), \alpha_2(t), \ldots, \alpha_m(t))^T$ for the internal cavity modes and external inputs $\beta_{\mathrm{in}}(t) = (\beta_{\mathrm{in},1}(t), \beta_{\mathrm{in},2}(t), \ldots, \beta_{\mathrm{in},n}(t))^T$. These amplitudes relate to quantum mechanical expectations as $\langle \alpha_j \rangle \approx \langle a_j \rangle_{\mathrm{Q}M}$, where $\langle \cdot \rangle$ denotes the expectation with respect to the Wigner quasi distribution and $\langle \cdot \rangle_{\mathrm{Q}M}$ a quantum mechanical expectation value. See [98] for the corresponding relations of higher order moments.

To simplify the analysis, we exclusively work in a rotating frame with respect to all driving fields. As in the SLH case we define output modes $\beta_{\mathrm{out}}(t)$ that are algebraically related to the inputs and the internal modes. The full dynamics of the internal and external modes are then governed by a multi-dimensional Langevin equation

$$\dot{\alpha}(t) = [\mathbf{A}\alpha(t) + \mathbf{a} + A_{\mathrm{N}L}(\alpha, t)] + \mathbf{B}\beta_{\mathrm{in}}(t), \tag{5.6}$$

as well as a purely algebraic, linear input-output relationship

$$\beta_{\mathrm{out}}(t) = [\mathbf{C}\alpha(t) + \mathbf{c}] + \mathbf{D}\beta_{\mathrm{in}}(t). \tag{5.7}$$

The complex matrices $\mathbf{A}, \mathbf{B}, \mathbf{C}, \mathbf{D}$ as well as the constant bias input vectors $\mathbf{a}$ and $\mathbf{c}$ parametrize the linear dynamics, whereas the function $A_{\mathrm{NL}}(\alpha, t)$ gives the nonlinear contribution to the dynamics of the internal cavity modes.

Each input consists of a coherent, deterministic part and a stochastic contribution $\beta_{\mathrm{in},j}(t) = \bar{\beta}_{\mathrm{in},j}(t) + \eta_j(t)$. The stochastic terms $\eta_j(t) = \eta_{j,1}(t) + i\eta_{j,2}(t)$ are assumed to be independent complex Gaussian white noise processes with correlation function $\langle \eta_{j,s}(t)\eta_{k,r}(t') \rangle = \frac{1}{4}\delta_{jk}\delta_{sr}\delta(t - t')$.

The linearity of the input-output relationship in either framework (5.5) and (5.7) in the external degrees of freedom leads to algebraic rules for deriving reduced models for whole circuits of nonlinear optical resonators by concatenating component models and algebraically solving for their interconnections. [40, 98] To see the basic component models used in this work see Appendix 5.5. Netlists for composite components and the whole circuit will be made available at [111].

5.2. The Coherent Perceptron Circuit

The full perceptron's circuit is visualized in Figure 5.1. The input data x to the perceptron circuit is encoded in the real quadrature of N coherent optical inputs. Equation (5.3) informs us what circuit elements are required for a hardware implementation by decomposing the necessary operations:

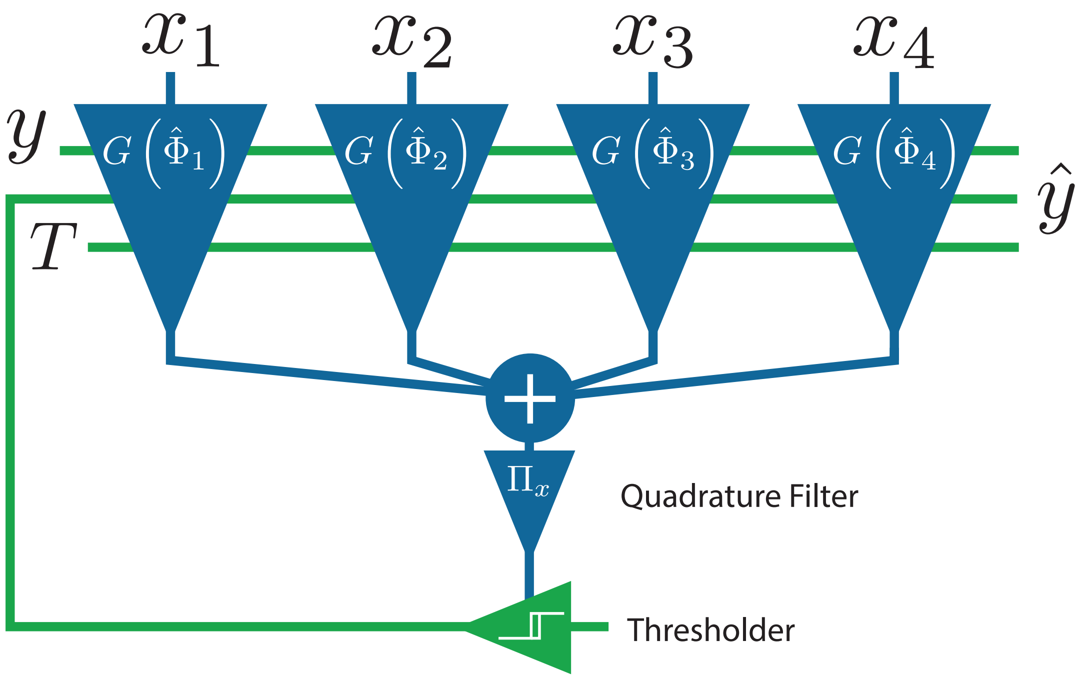

x_1
x_2
x_3
x_4
y
T
$G\left(\hat{\Phi}_1\right)$
$G\left(\hat{\Phi}_2\right)$
$G\left(\hat{\Phi}_3\right)$
$G\left(\hat{\Phi}_4\right)$
$\hat{y}$
+
Π_x
Quadrature Filter
Thresholder

Figure 5.1.: An example perceptron circuit consisting of $N = 4$ programmable amplifiers for the coherent input vector $x = (x_1, x_2, x_3, x_4)^T$, a static mixing element that sums their output, a quadrature filter to remove the imaginary quadrature and a final thresholding element to generate the estimated binary class label $\hat{y}$. The additional binary input T controls whether the system is in *training mode*, in which case the estimated class label $\hat{y}$ is compared to the true class label Y which is provided as an additional input. When they differ, the programmable amplifiers receive a feedback signal to adjust their internal weights.

1. Each input x_j is multiplied by a weight w_j.
2. The weighted inputs are coherently added.
3. The sum drives a thresholding element to generate the estimated class label $\hat{y}$.
4. In the training phase (input $T = 1$) the estimated class label $\hat{y}$ is compared with the true class label (input Y) and based on the outcome, feedback is applied to modify the weights $\{w_j\}$.

The most crucial element for this circuit is the system that multiplies an input x_j with a programmable weight w_j. This not only requires having a linear amplifier with tunable gain, but also a way to encode and store the continuous weights w_j. In the following we outline one way how such systems can be constructed from basic nonlinear optical cavity models: Section 5.2.1 presents an elegant way to construct a phase sensitive linear optical amplifier where the gain can be tuned by changing the amplitude of a bias input. In Section 5.2.2 we propose using an above threshold non-degenerate optical parametric amplifier to store a continuous variable in the output phase of the signal (or idler) mode. In Section 5.2.3 these systems are combined to realize an optical amplifier with *programmable* gain, i.e., a control input can program its gain, which then stays constant even after the control has been turned off. Finally, we present a simple model for all-optical switches based on a cavity with two modes that interact via a cross-Kerr-effect in Section 5.2.4. This element is used both for the feedback logic as well as the thresholding function to generate the class label $\hat{y}$.

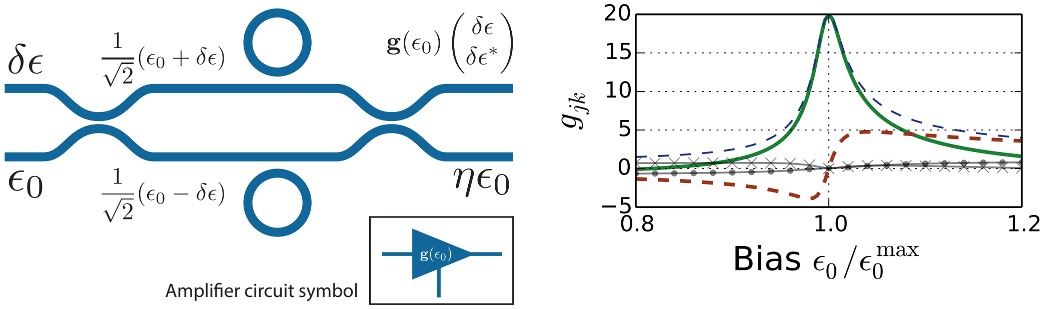
5.2.1. Tunable Gain Kerr-amplifier

A single mode Kerr-nonlinear resonator driven by an appropriately detuned coherent drive ϵ can have a strongly nonlinear dependence of the intra-cavity energy on the drive power. When the drive of a single resonator is given by the sum of a constant large bias amplitude and a small signal $\epsilon = \frac{1}{\sqrt{2}}(\epsilon_0 + \delta\epsilon)$, the steady state reflected amplitude is $\epsilon' = \frac{1}{\sqrt{2}}(\eta\epsilon_0 + g_-(\epsilon_0)\delta\epsilon + g_+(\epsilon_0)\delta\epsilon^*) + O(\delta\epsilon^2)$, where $|\eta| \leq 1$ with equality for the ideal case of negligible intrinsic cavity losses. The small signal thus experiences phase sensitive gain dependent on the bias amplitude and phase. We provide analytic expressions for the gain in Appendix 5.5.2.

Placing two identical resonators in the arms of an interferometer allows for isolating the signal and bias outputs even if their amplitudes vary by canceling the scattered bias in one output and the scattered signal in the other (cf. Figure 5.2). This highly symmetric construction, which generalizes to any other optical nonlinearity, ensures that the the signal output is linear in $\delta\epsilon$ up to third order[1]. If the system parameters are well-chosen, the amplifier gain depends very strongly on small variations of the bias amplitude. This allows to tune the gain from close to unity to its maximum value, which, for a given waveguide coupling κ and Kerr coefficient χ depends on the drive detuning from cavity. For Kerr-nonlinear resonators there exists a critical detuning beyond which the system becomes bi-stable and exhibits hysteresis. This can be used for thresholding type behavior though as shown in [108] in this case it may be advantageous to reduce the symmetry of the circuit. It is convenient to engineer the relative propagation phases such that at maximum gain, a real quadrature input signal $x \in \mathbb{R}$ leads to an amplified output signal $x' = g_{rr}^{\max} x$ with no imaginary quadrature component (other than noise and higher order contributions). However, for different bias input amplitudes and consequently lower gain values the output will generally feature a linear imaginary quadrature component $x' = [g_{rr}(\epsilon_0) + i g_{ir}(\epsilon_0)]\, x$ as

---

[1] One can easily convince oneself that all even order contributions are scattered into the bias output.



(a) Amplifier circuit

(b) Gain vs. bias

Figure 5.2.: (a) shows two identical single mode Kerr-nonlinear optical resonators symmetrically placed in the two arms of an interferometer. (b) gives the phase sensitive amplifier gain $g_{rr}(\epsilon_0)$ (green, solid) and the $g_{ir}(\epsilon_0)$ (red, dashed) as a function of the bias photon input rate normalized by the drive power at which dynamic resonance occurs. For completeness we also provide g_{ri} (black X's) and g_{ii} (black dots). The detuning has been chosen such that $g_{rr}^{\max} = g_{rr}(\epsilon_0^{\max}) = 20$. The dashed blue envelope gives the maximal input output gain achievable between any two signal quadratures at that bias. Note that g_{rr} vanishes at $\epsilon_0/\epsilon_0^{max} \approx 0.8$.

well. Figure 5.2(b) demonstrates this for a particular choice of maximal gain. We note that there exist previous proposals of using nonlinear resonator pairs inside interferometers to achieve desirable input-output behavior [108], but to our knowledge, no one has proposed using these for signal/bias isolation and tunable gain. To first order the linearized Kerr model is actually identical to a sub-threshold degenerate OPO model. This implies that it can be used to generate squeezed light and also that one could replace the Kerr-model by an OPO model.

An almost identical circuit, but featuring resonators with additional internal loss equal to the wave-guide coupling[2] and constantly biased to *dynamic resonance* $\langle|\alpha|^2\rangle_{\mathrm{ss}} = -\Delta/\chi$ can be used to realize a *quadrature filter*, i.e., an element that has unity gain for the real quadrature and zero for the imaginary one. Now the quadrature filtered signal still has an imaginary component, but to linear order this only consists of transmitted noise from the additional internal loss. While it would be possible to add one of these downstream of every tunable Kerr amplifier, in our specific application it is more efficient to add just a single one downstream of where the individual amplifier outputs are summed (cf. Section 5.2.5). This also reduces the total amount of additional noise introduced to the system.

---

[2] In the photonics community this is referred to as *critically coupled,* whereas the amplifier circuit would ideally be strongly *overcoupled* such that additional internal losses are negligible.

5.2.2. Encoding and Storing the Gain

In the preceding section we have seen how to realize a tunable gain amplifier, but for programming and *storing* this gain (or equivalently its bias amplitude) an additional component is needed. Although it is straightforward to design a multi-stable system capable of outputting a discrete set of different output powers to be used as the amplifier bias, such schemes would likely require multiple nonlinear resonators and it would be more cumbersome to drive transitions between the output states.

An alternative to such schemes is given by systems that have a continuous set of stable states. Recent analysis of continuous time recurrent neural network models trained for complex temporal information processing tasks has revealed multi-dimensional stable attractors in the internal network dynamics that are used to store information over time. [106]

A simple semi-classical nonlinear resonator model to exhibit this is given by a non-degenerate optical parametric oscillator (NOPO) pumped above threshold; for low pump input powers this system allows for parametric amplification of a weak coherent signal (or idler) input. In this case vacuum inputs for the signal and idler lead to outputs with zero expected photon number. Above a critical threshold pump power, however, the system down-converts pump photons into pairs of signal and idler photons.

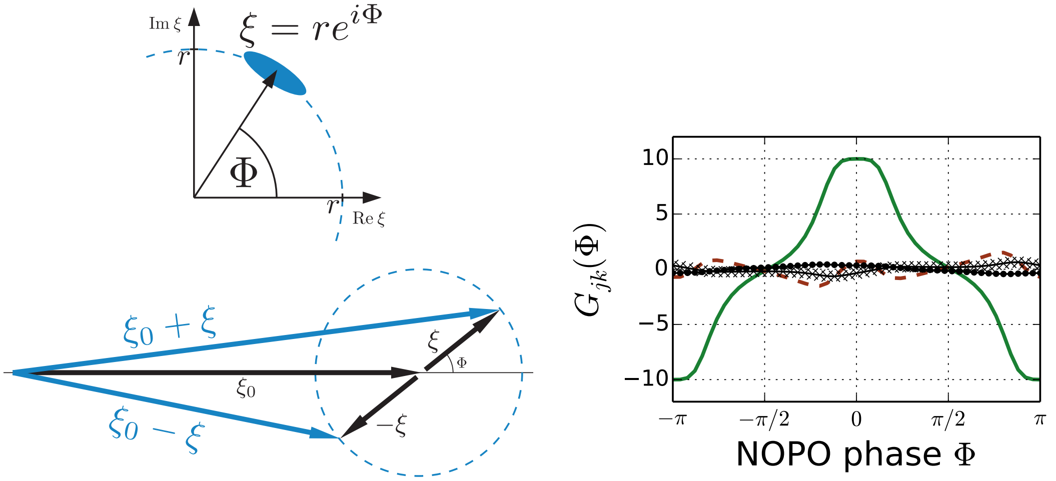
Due to an internal $U(1)$ symmetry of the underlying Hamiltonian (cf. Appendix 5.5.2), the signal and idler modes spontaneously select phases that are dependent on each other but independent of the pump phase. This implies that there exists a whole manifold of fix-points related to each other via the symmetry transformation $(\alpha_s, \alpha_i) \to (\alpha_s e^{i\phi}, \alpha_i e^{-i\phi})$, where α_s and α_i are the rotating frame signal and idler mode amplitudes, respectively. Consequently the signal output of an above threshold NOPO lives on a circular manifold (cf Figure 5.3).

Vacuum shot noise on the inputs leads to phase diffusion with a rate of $\gamma_\Phi = \frac{\kappa}{8n_0}$, where κ is the signal and idler line width and n_0 is the steady state intra cavity photon number in either mode. We point out that this diffusion rate does not directly depend on the strength of the nonlinearity which only determines how strongly the system must be pumped to achieve a given intra cavity photon number n_0.

A weak external signal input breaks the symmetry and biases the signal output phase towards the external signal's phase. This allows for changing the programmed phase value.

Finally, we note that parametric oscillators can also be realized in materials with vanishing χ_2 nonlinearity. They have been successfully realized via four-wave mixing (i.e., exploiting a χ_3 nonlinearity) in [60, 99, 25] and even in opto-mechanical systems [20] in which case the idler mode is given by a mechanical degree of freedom.

In principle any nonlinear optical system that has a stable limit cycle could be used to



(a) Combined bias

(b) Gain vs. OPO phase

Figure 5.3.: The NOPO's signal output $\xi = \sqrt{\kappa}\alpha_s$ lives on a circular manifold parametrized by Φ (a, upper figure). Vacuum input shot noise leads to small fluctuations perpendicular to the manifold and diffusion along it. Mixing this signal output with a constant bias offset on a beamsplitter produces two outputs with anti-correlated total amplitude (a, lower figure). When both outputs are used to drive a complementary pair of tunable amplifiers whose outputs are subtracted, the overall real-to-real quadrature gain (green) of the system varies from positive to negative values (b). We can also see that the real-to-imaginary gain (dashed red) stays small for all NOPO phases, which allows us to efficiently subtract it downstream by the quadrature filter. The imaginary to real and imaginary gains are also plotted.

store and encode a continuous value in its oscillation phase. Non-degenerate parametric oscillators stand out because of their theoretical simplicity allowing for a 'static' analysis inside a rotating frame.

5.2.3. Programmable Gain Amplifier

Combining the circuits described in the preceding sections allows us to construct a fully programmable phase sensitive amplifier. In Figure 5.2(b) we see that there exists a particular bias amplitude at which the real to real quadrature gain vanishes $g_{rr}(\epsilon_0^{\text{min}}) = 0$. We combine the NOPO signal output $\xi = re^{i\Phi}$ with a constant phase bias input ξ_0 (cf. Figure 5.3(a)) on a beamsplitter such that the outputs vary between zero gain and the maximal gain bias values $\left|\frac{\xi_0 \pm re^{i\Phi}}{\sqrt{2}}\right| \in [\epsilon_0^{\text{min}}, \epsilon_0^{\text{max}}]$. To realize both positive and negative gain, we use the second output of that beamsplitter to bias another tunable amplifier. The two amplifiers are always biased oppositely meaning that one will have maximal gain when the other's gain vanishes and vice versa. The overall input signal is split and sent through both amplifiers and then re-combined with a relative π phase shift. This complementary setup leads to an overall effective gain tunable within $G_{rr}(\Phi) \in [-\frac{g_{rr}^{max}}{2}, \frac{g_{rr}^{max}}{2}]$ (cf. Figure 5.3(b)).

In Figure 5.4 we present both the complementary pair of amplifiers and the NOPO used for storing the bias as well as some logic elements (described in Section 5.2.4) used for implementing conditional training feedback. We call the full circuit a synapse because it features programmable gain and implements the perceptron's conditional weight update rule.

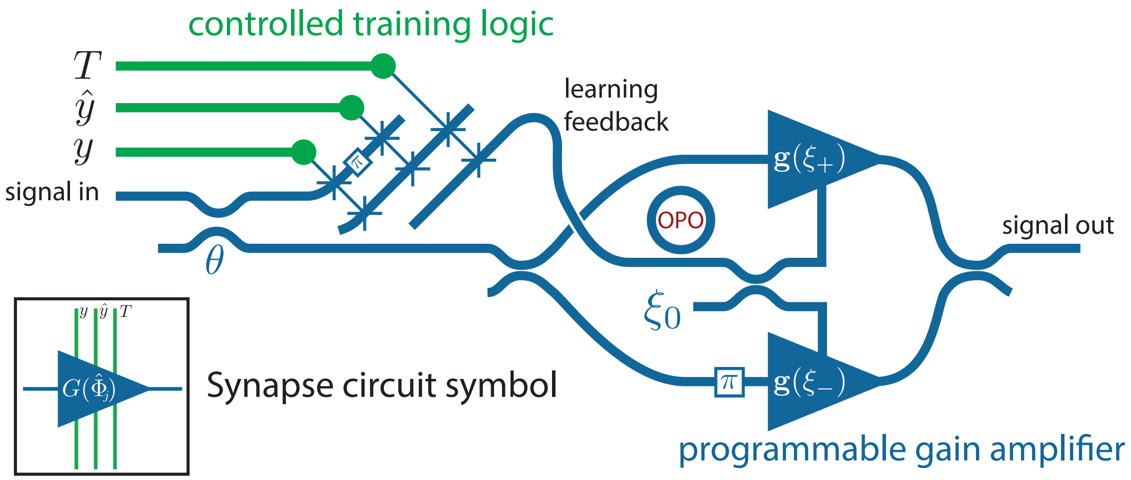
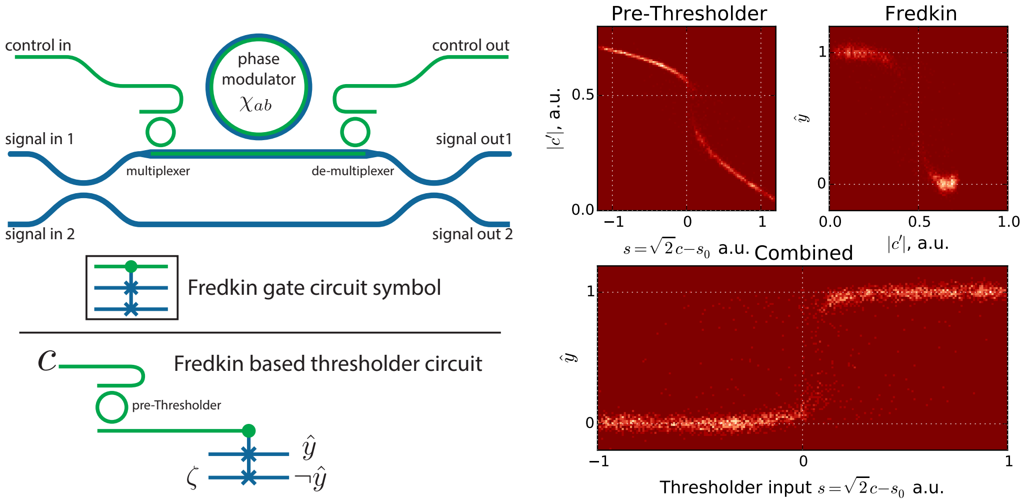


Figure 5.4.: Synapse circuit composed of a programmable amplifier and feedback logic (cf. Section 5.2.4) that implements the perceptron learning feedback (5.3) for a single weight. The upper amplifier when biased optimally leads to positive gain whereas the lower amplifier leads to negative gain due to the additional π phase shift.

The resulting synapse model is quite complex and certainly not optimized for a minimal component number but rather the ease of theoretical analysis. A more resource efficient programmable amplifier could easily be implemented using just two or three nonlinear resonators. E.g., inspecting the the real to imaginary quadrature gain $g_{ir}(\epsilon_0)$ in Figure 5.2(b) we see that close to $\epsilon_0^{\max}$ it passes through zero fairly linearly and with an almost symmetric range. This indicates that we could use a single tunable amplifier to realize both positive and negative gain. Using only a single resonator for the tunable amplifier could work as well, but it would require careful interferometric bias cancellation and more tedious upfront analysis. We do not think it is feasible to use just a single resonator for both the parametric oscillator and the amplifier because any amplified input signal would have an undesirable back-action on the oscillator phase.

5.2.4. Optical Switches

The feedback to the perceptron weights (cf. Equation (5.3)) is conditional on the binary values of the given and estimated class labels y and $\hat{y}$, respectively. The logic necessary for implementing this can be realized by means of all-optical switches. There have been various proposals and demonstrations [92, 83] of all-optical gates/switches and quantum optical switches [74].



(a) Fredkin gate and thresholder

(b) Thresholder input/output

Figure 5.5.: In the upper graphic of (a) we present a schematic for Fredkin gate based on a two mode cross-Kerr-nonlinear resonator. The lower graphic shows how this circuit can be pre-pended with a single mode nonlinear resonator to better approximate a thresholding response. In (b) we present the input output characteristic of the prepended resonator (upper left), the Fredkin gate (upper right) and the combined input output relationship between the inner product amplitude s and the estimated state label $\hat{y}$.

The model that we assume here (cf. Figure 5.5) is to use two different modes of a resonator that interact via a cross-Kerr-effect, i.e., power in the control mode leads to a refractive index shift (or detuning) for the signal mode. The index shift translates to a control mode dependent phase shift of a scattered signal field yielding a controlled optical phase modulator. Wrapping this phase modulator in a Mach-Zehnder interferometer then realizes a controlled switch: If the control mode input is in one of two different states $|\xi| \in 0, \xi_0$, the signal inputs are either passed through or switched. This operation is often referred to as a *controlled swap* or Fredkin gate [31] which was originally proposed for realizing reversible computation. This dispersive model has the advantage that the control input signal can be reused.

Note that at control input amplitudes significantly different from the two control levels the outputs are coherent mixtures of the inputs, i.e., the switch then realizes a tunable beamsplitter.

Finally, we point out that using two different (frequency non-degenerate) resonator modes has the advantage that the interaction between control and signal inputs is phase insensitive which greatly simplifies the design and analysis of cascaded networks of such switches.

5.2.5. Generation of the Estimated Label

The estimated classifier label $\hat{y}$ should be a step function applied to the inner product of the weight vector and the input. In the preceding sections we have shown how individual inputs x_j can be amplified with programmable gain to give $\tilde{s}_j = \tilde{G}(\Phi_j)x_j$, thus realizing the individual contributions to the inner product. These are then summed on an n-port beamsplitter that has an output which gives the uniformly weighted sum $\tilde{s} := \frac{1}{\sqrt{N}}\sum_{k=1}^{N}\tilde{G}(\Phi_k)x_k$.

The gain factors $\tilde{G}(\Phi_k) = G_{rr}(\Phi_k)+iG_{ir}(\Phi_k)$ generally have an unwanted imaginary part which we subtract by passing the summed output through a *quadrature filter* circuit (cf. the last paragraph of Section 5.2.1), which has unit gain for the real quadrature and zero gain for the imaginary quadrature leading to an overall output $s = \mathrm{R}e\,\tilde{s} = \frac{1}{\sqrt{N}}\sum_{k=1}^{N}G_{rr}(\Phi_k)x_k$. The thresholding circuit should now produce a high output if $s > 0$ and a zero output if $s \leq 0$.

It turns out that the optical Fredkin gate described in the previous section already works almost as a two mode thresholder, where the control input leads to a step-like response in the signal outputs: A constant signal input amplitude which encodes the logical ‘1’ state is applied to one of the signal inputs. When the control input amplitude is varied from zero to ξ_0, the signal output turns on fairly abruptly at some threshold $\xi_{th} < \xi_0$. To make the thresholding phase sensitive, the control input is given by the sum of s and a constant offset s_0 that provides a phase reference: $c = \frac{1}{\sqrt{2}}(s + s_0)$.

For a Fredkin gate operated with continuous control inputs the signal output is almost zero for a considerable range of small control inputs. However, for very high control inputs, i.e., significantly above ξ_0, the signal output decreases instead of staying constant as would be desirable for a step-function like profile. We found that this issue can be addressed by transmitting the control input through a single mode Kerr-nonlinear cavity, with resonance frequency chosen such that the transmission gain $|c'/c|$ is peaked close to $c' = \xi_0$. For input amplitudes larger than c, the transmission gain is lower (although $|c'|$ still grows monotonically with $|c|$) which extends the input range over which the subsequent Fredkin gate stays in the on-state.

5.3. Results

The perceptron's SDEs where simulated using a newly developed custom software package named QHDLJ [112] implemented in Julia [7] which allows allows for dynamic compilation of circuit models to LLVM [63] bytecode that runs at speed comparable to C/C++. All individual simulations can be carried out on a laptop, but the results in Figure 5.8 were obtained by averaging over the results of 100 stochastic simulation run on an HP ProLiant server with 80 cores. The current version of QHDLJ uses one process per trajectory, but the code could easily be vectorized.

In Figure 5.6 we present an example of a single application of an $N = 8$ perceptron including both a learning stage with pre-labeled training data and a classification testing stage in which the perceptron's estimated class labels are compared with their correct values. The data to be classified here are sampled from a different $8-$ dimensional Gaussian distribution for each class label with their mean vectors separated by a distance $\|\mu_1 - \mu_0\|_2/\sigma = 2$ relative to the standard deviation of both individual clusters. For each sample the input was held constant for a duration $\Delta t = 2\kappa^{-1}$ where κ is the NOPO signal and idler line width. The perceptron was first trained with $M_{train} = 100$ training examples and subsequently tested on $M_{test} = 100$ test examples with the learning feedback turned off.

In Figure 5.7 we visualize linear projections of the testing data as well as the estimated classification boundaries. We can see that the classifier performs very well far away from the decision boundary. Close to the decision boundary there are some misclassified examples. We proceed to compare the performance of the classifier to the theoretically optimal performance achievable by any classifier and with the optimal classifier for this scenario, Gaussian Discriminant Analysis (GDA) [28, 73], implemented in software. Using the identical perceptron model as above and an identical training/testing procedure, we estimate the error rate $p_{err} = \mathbb{P}[y \neq \hat{y}]$ of the trained perceptron as a function of the cluster separation $\|\mu_1 - \mu_0\|_2/\sigma$. The results are presented in Figure 5.8(a). Identically distributed training and testing data was used to evaluate the performance of the GDA algorithm and both results are compared to the theoretically optimal error rate for this discrimination task, which can be computed analytically to be $p_{err,\,optim.} = \frac{1}{2}erfc\left(\frac{\|\mu_1-\mu_0\|_2}{\sqrt{8}\sigma}\right)$, where $erfc(x) = \frac{2}{\sqrt{\pi}}\int_x^\infty e^{-u^2}du$ is the complementary error function. We see that the all-optical perceptron's performance is comparable to GDA's performance for this problem and both algorithms attain performance close to the theoretical optimum.

The learning rate of the perceptron is determined by two things, the overall strength of the learning feedback as well as the time for which each example is presented to the circuit. In Figure 5.8(b) we plot the estimated error rate for varying feedback strength and duration. As can be expected intuitively, we find that there are trade-offs between speed (smaller Δt

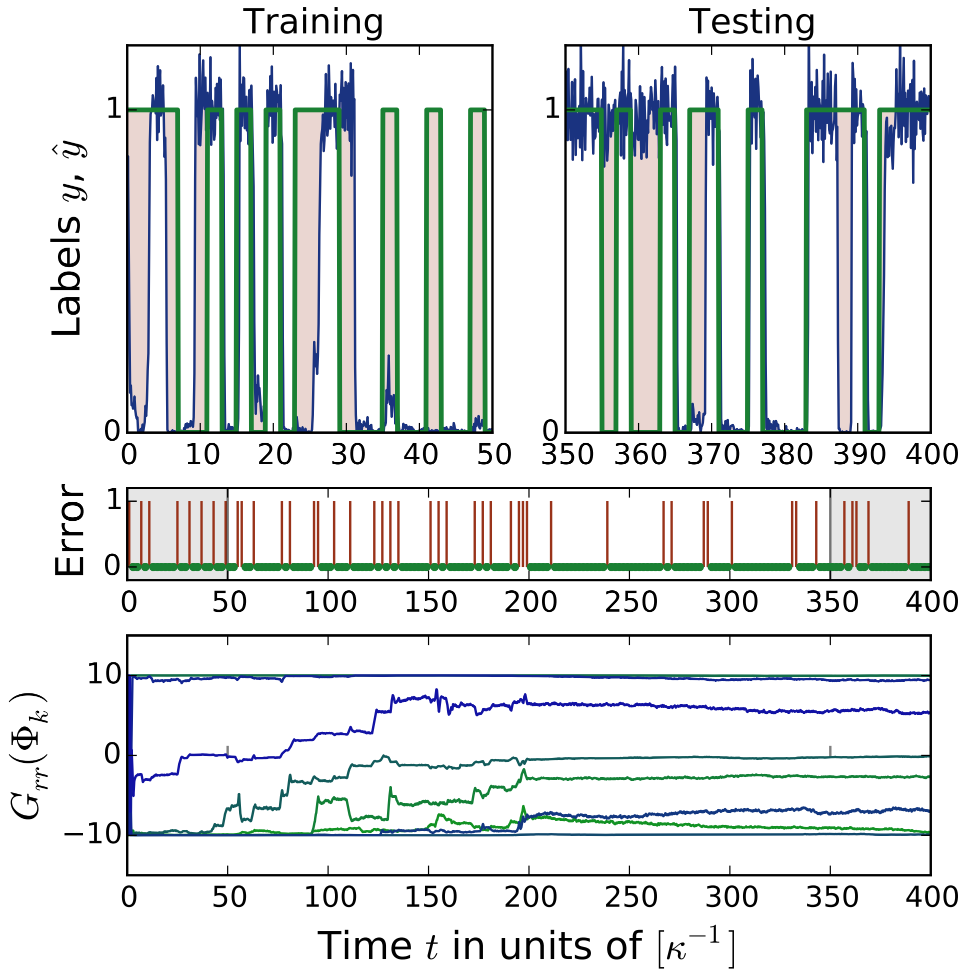


Figure 5.6.: Single trajectory divided into a training interval $0 \leq t \leq M_{train}\Delta t$ during which the learning feedback is active and a test interval $M_{train}\Delta t < t \leq M_{test}\Delta t$. During training and testing, respectively, the system is driven by $M_{train} = M_{test} = 100$ separate input states which are held constant for an interval $\Delta t = 2\kappa^{-1}$. The estimated class label is discretized by averaging the output intensity over each input interval, dividing the result by the intensity $|\zeta|^2$ corresponding to the logical '1' output state and rounding. The upper panel compares the correct class label y (green) with the estimated class label $\hat{y}$ (black) during training and testing, respectively. The area between them indicates errors or at least lag of the estimator and is shaded in light red. The second panel shows occurrences of classification errors (red vertical bars). The slight shading near the beginning and the end of the trajectory in the second panel visualizes the segments corresponding to the upper left and right panel, respectively. The third panel shows the learned linear amplitude gains for each synapse. After the learning feedback is turned off at $t = M_{train}\Delta t$, they diffuse slightly due to optical shot noise.

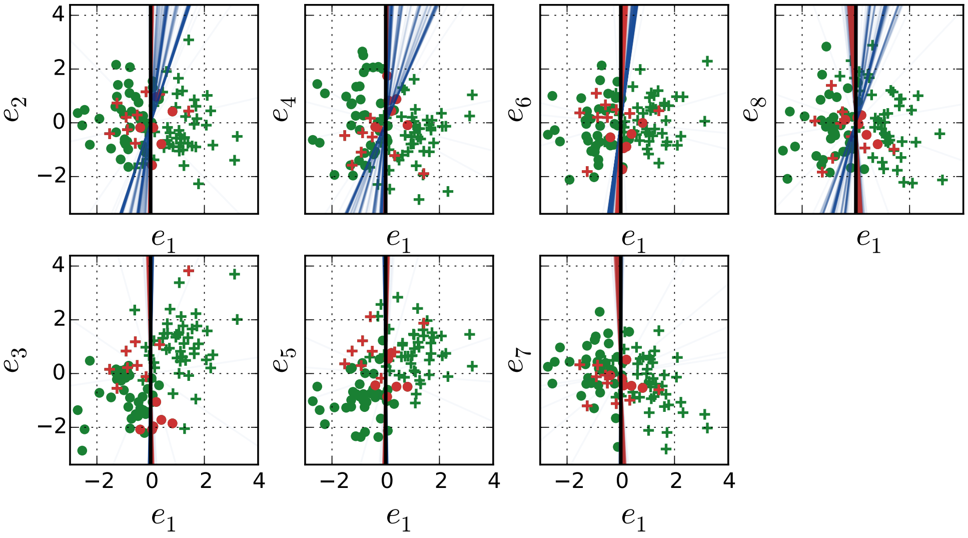


Figure 5.7.: Projection of training data and classification boundaries. The data has been rotated such that the s_1 coordinate lines up with the learned normal vector of the separating hyperplane. Incorrectly classified data are plotted in red. The faint blue (red) lines visualize the evolution of the classifier boundary during training (testing).

preferable) and energy consumption (smaller α preferable).

5.3.1. Time scales and power budget

Here we roughly estimate the power consumption of the whole device and discuss how to scale it up to a higher input dimension.

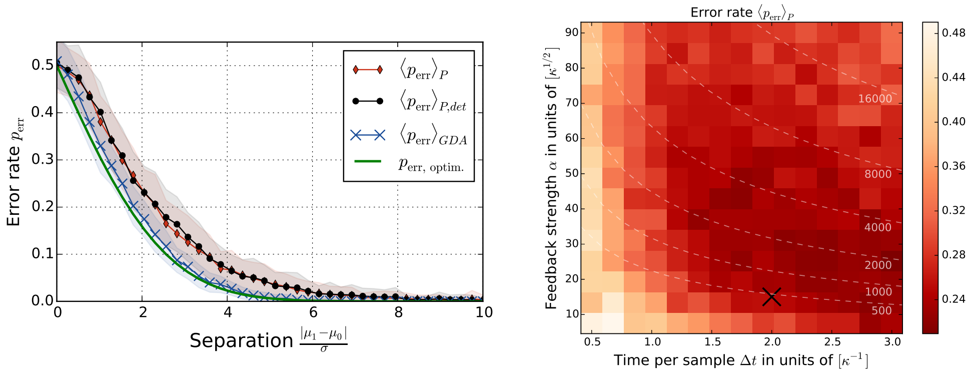
Any real-world implementation will depend strongly on the engineering paradigm, i.e., the choice of material/nonlinearity as well as the engineering precision, but based on recently achieved progress in nonlinear optics we will estimate an order of magnitude range for the input power.

The signal and feedback input power to the circuit will scale linearly in the number of synapses N.

The bias inputs for the amplifiers has to be larger than the signal to ensure linearly operation, but it should be expected that some of the scattered bias amplitudes can be reused to power multiple synapses.

In our models we have defined all rates relative to the line width of the signal and idler mode of the NOPO, because this is the component that should necessarily have the smallest decay rate to ensure a long lifetime for the memory.

All other resonators are employed as nonlinear input-output transformation devices and



(a) Error rate vs. hardness

(b) Error rate vs. learning paramaters

Figure 5.8.: The perceptron's error rate vs the difficulty of the classification task and as a function of the parameters determining the learning rate. In Figure (a) we compare the unoptimized performance of the perceptron circuit (red diamonds) to the optimal performance bound (solid, green) as well as a GDA (blue X's) trained on the same number of training examples. We show averages over 100 trials at each cluster separation. The GDA data was similarly averaged over 100 trials. The transparent envelopes indicate the sample standard deviation. The black dots show the perceptron performance when simulated without shot noise. We see that the shot noise has very little effect. In Figure (b) we plot the average error rate (averaged over 50 trials) at fixed cluster separation $\|\mu_1 - \mu_0\|_2/\sigma = 2$ for various values of the time interval Δt for which each data sample is presented to the circuit as well as the strength of the training feedback α. The total number of feedback photons $N_{fb} = |\alpha|^2 \Delta t$ per sample is constant along the faint dashed lines and the actual value is indicated on the right. A good choice of parameters is characterized both by low feedback power (small $|\alpha|^2$) and high input rate (low sample time Δt) while still resulting in a low classification error rate. The X marks the parameters used for the results in (a) and the previous Figures.

therefore a high bandwidth (corresponding to much lower loaded quality factor) is necessary for achieving a high bit rate. For our simulations we typically assumed quality factors that were lower than the NOPO's by 1-2 orders of magnitude.

Based on self-oscillation threshold powers reported in [60, 25, 64, 95] and the switching powers of [83] we estimate the necessary power per synapse to be in the range of $\sim 10 - 100\mu\mathrm{W}att$. By re-using the scattered pump and bias fields it should be possible to reduce the power consumption per amplifier even further. Even for the continuous wave signal paradigm we have assumed (as opposed to pulsed/spiking signals such as considered in [124]) the devices proposed here could be competitive with the current state of the art CMOS-based neuromorphic electrical circuits [17].

In the simulations for the 8−dimensional perceptron our input rate for training data was set to $\Delta t^{-1} = \frac{\kappa}{2}$. This value corresponds to roughly ten times the average feedback delay time between arrival of an input pattern and the conditional switching of the feedback logic upon arrival of the generated estimated state label $\hat{y}$. This time can be estimated as $\tau_{fb}(n) \approx G_{\max}\kappa_A^{-1} + \kappa_{QF}^{-1} + \kappa_{thresh}^{-1} + n\kappa_F^{-1}$, where n is the index of the synaptic weight, G_{max} is the amplifier gain range and $\kappa_A, \kappa_{QF}, \kappa_{thresh}$ and κ_F are the line widths of the amplifier, quadrature filter, the combined thresholding circuit (cf. Figure 5.5) and the feedback Fredkin gates. There is a contribution scaling with n because the feedback traverses the individual weights sequentially to save power.

When scaling up the perceptron to a higher dimension while retaining approximately the same input signal powers, it is intuitively clear that the combined 'inner product' signal amplitude s scales as $s \propto \sqrt{N}s_1$, where s_1 is the signal amplitude for a single input. This allows to similarly scale up the amplitude ζ_0 of the signal encoding the generated estimated state label $\hat{y}$ and consequently the bandwidth of the feedback Fredkin gates that it drives. A detailed analysis reveals that the Fredkin gate threshold scales as $\sqrt{N}$, in particular we find that $\sqrt{|\chi|}\zeta_0 \propto \kappa_F \propto \sqrt{|\chi|}\xi_0 \propto \kappa_{thresh} \propto \sqrt{|\chi|}s \propto \sqrt{N|\chi|}s_1$. The first two scaling relationships are due to the constraints on the Fredkin gate construction (cf. Appendix 5.5.2), the next two scaling relationships follow from demanding that the additional thresholding resonator be approximately dynamically resonant at the highest input level (cf. Appendices 5.5.2 and 5.5.2). The last proportionality is simply due to the amplitude summation at the N-port beamsplitter.

This reveals that when increasing N the perceptron as constructed here would have to be driven at a lower input bit rate scaling as $\Delta t^{-1} \propto N^{-\frac{1}{2}}$ or alternatively be driven with higher signal input powers. A possible solution that could greatly reduce the difference in arrival time $\sim \kappa_F^{-1}$ at each synapse could be to increase the waveguide-coupling to the control signal and thus decrease the delay per synapse. The resulting increase in the required control amplitude ζ_0 can be counter-acted with feedback, i.e., by effectively creating a large cavity around the control loop. When even this strategy fails one could add fan-out stages for $\hat{y}$ which introduce a delay that grows only logarithmically with N.

Finally, we note that the bias power of all the Kerr-effect based models considered here scales inversely with the respective nonlinear coefficient $\{|\zeta_0|^2, |s|^2\} \times |\chi| \sim const$ when keeping the bandwidth fixed. This implies that improvements in the non-linear coefficient translate to lower power requirements or alternatively a faster speed of operation.

5.4. Conclusion and Outlook

In conclusion we have shown how to design an all-optical device that is capable of supervised learning from input data, by describing how tunable gain amplifiers with signal/bias isolation can be constructed from nonlinear resonators and subsequently combined with self-oscillating resonators to encode the programmed amplifier gain in their oscillation phase. By considering a few additional nonlinear devices for thresholding and all-optical switching we then show how to construct a perceptron, including the perceptron feedback rule. To our knowledge this is the first end-to-end description of an all-optical circuit capable of learning from data. We have furthermore demonstrated that despite optical shot-noise it nearly attains the performance of the optimal software algorithm for the classification task that we considered. Finally, we have discussed the relevant time-scales and pointed out how to scale the circuit up to large input dimensions while retaining the signal processing bandwidth and a low power consumption per input.

Possible applications of an all-optical perceptron are as the trainable output filter of an optical reservoir computer or as a building block in a multi-layer all-optical neural network.

The programmable amplifier could be used as a building block to construct other learning models that rely on continuously tunable gain such as Boltzmann machines and hardware implementations of message passing algorithms.

An interesting next step would be to design a perceptron that can handle inputs at different carrier frequencies. In this case wavelength division multiplexing (WDM) might allow to significantly reduce the physical footprint of the device.

A simple modification of the perceptron circuit could autonomously learn to invert linear transformations that were applied to its input signals. This could be used for implementing a circuit capable of solving linear regression problems. In combination with a multi-mode optical fibers such a device could also have applications for all-optical sensing.

Finally, an extremely interesting question is whether harnessing quantum dynamics could lead to a performance increase. We hope to address these ideas in future work.

5.5. Basic Component Models

Here we present the component models used to build the perceptron circuit. We will first describe the static components such as beamsplitters, phase shifts and coherent displacements, then proceed to describe the different Kerr-nonlinear models and finally the NOPO model.

5.5.1. Static, Linear Circuit Components

All of these components have in common that they have no internal dynamics, implying that the A, B and C matrices and the a-vector have zero elements, and A_{NL} is not defined.

Constant Laser Source

The simplest possible static component is given by single input/output coherent displacement with coherent amplitude η. This model is employed to realize static coherent input amplitudes. The D matrix is trivially given by $D = (1)$ and the coherent amplitude is encoded in $c = (\eta)$. This leads to the desired input-output relationship $\beta_{out} = \eta + \beta_{\text{in}}$. For completeness we also provide the SLH [41] model $((1), (\eta), 0)$.

Static Phase Shifter

The static single input/outputs phase shifter has $D = (e^{i\phi})$ and $c = (0)$, leading to an input output relationship of $\beta_{out} = e^{i\phi}\beta_{\text{in}}$. Its SLH model is $((e^{i\phi}), (0), 0)$.

Beamsplitter

The static beamsplitter mixes (at least) two input fields and can be parametrized by a mixing angle θ. It has $D = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}$ and $c = (0, 0)^T$. This leads to an input output relationship

$$\begin{pmatrix} \beta_{out,1} \\ \beta_{out,2} \end{pmatrix} = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix} \begin{pmatrix} \beta_{in,1} \\ \beta_{in,2} \end{pmatrix} \tag{5.8}$$

Its SLH model is $\left(\begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \end{pmatrix}, 0 \right)$.

5.5.2. Resonator Models

We consider resonator models with m internal modes and n external inputs and outputs. We assume for simplicity that $a = \mathbf{0}$ and $c = \mathbf{0}$ meaning that we will model all coherent displacements explicitly in the fashion described above. We also assume that their scattering matrices are trivially given by $D = \mathbf{1}_n$ which means that far off-resonant input fields are simply reflected without a phase shift. Furthermore, none of our assumed models feature *linear* coupling between the internal cavity modes. This implies that the A-matrix is always diagonal. We are always working in a rotating frame.

Single mode Kerr-nonlinear Resonator

A Kerr-nonlinearity is modeled by the nonlinear term $A_{\text{NL}}^{\text{K}err}(\alpha) = -i\chi|\alpha|^2\alpha$ which can be understood as an intensity dependent detuning. The A-matrix is given by $(-\frac{\kappa_T}{2} - i\Delta)$, its B-matrix is $-(\sqrt{\kappa_1}, \sqrt{\kappa_2}, \dots, \sqrt{\kappa_n})$, where the total line width is given by $\sum_{j=1}^{n} \kappa_j = \kappa_T$

and the cavity detuning from any external drive is given by Δ. The C-matrix is given by $C = -B^T$. The corresponding SLH model is

$$\left(\mathbf{1}_n, \begin{pmatrix} \sqrt{\kappa_1} a \\ \vdots \\ \sqrt{\kappa_n} a \end{pmatrix}, \tilde{\Delta} a^\dagger a + \frac{\chi}{2} a^{2\dagger} a^2 \right), \tag{5.9}$$

where the detuning differs slightly $\tilde{\Delta} = \Delta + \chi$ as can be shown in the derivation of the Wigner-formalism. [98]

The special case of a single mirror with coupling rate κ and negligible internal losses is of interest for constructing the phase sensitive amplifier described in Section 5.2.1. Considering again an input given by a large static bias and a small signal $\epsilon = \frac{1}{\sqrt{2}}(\epsilon_0 + \delta\epsilon)$, the steady state reflected amplitude is to first order

$$\epsilon' \approx \frac{1}{\sqrt{2}} \left[\eta \epsilon_0 + g_-(\epsilon_0) \delta\epsilon + g_+(\epsilon_0) \delta\epsilon^* \right]. \tag{5.10}$$

For negligible internal losses we can give provide exact expressions for η, g_+ and g_-. Rather than parametrizing these by the bias ϵ_0 we parametrize them by the mean coherent intra-cavity amplitude α_0. When the system is not bi-stable (see below) relationship (5.14) defines a one-to-one map between ϵ_0 and α_0.

$$\eta = -\frac{\kappa/2 - i(\Delta + \chi|\alpha_0|^2)}{\kappa/2 + i(\Delta + \chi|\alpha_0|^2)} \quad \Rightarrow |\eta| = 1, \tag{5.11}$$

$$g_- = 1 + \frac{\kappa \left[-\frac{\kappa}{2} + i\Delta + 2i\chi|\alpha_0|^2 \right]}{\left(\frac{\kappa}{2}\right)^2 + (\Delta + 2\chi|\alpha_0|^2)^2 - |\chi|^2 |\alpha_0|^4}, \tag{5.12}$$

$$g_+ = \frac{i\kappa\chi\alpha_0^2}{\left(\frac{\kappa}{2}\right)^2 + (\Delta + 2\chi|\alpha_0|^2)^2 - |\chi|^2 |\alpha_0|^4}, \tag{5.13}$$

$$\epsilon_0 = -\frac{1}{\sqrt{\kappa}} \left[\frac{\kappa}{2} + i(\Delta + i\chi|\alpha_0|^2) \right] \alpha_0. \tag{5.14}$$

The Kerr cavity exhibits bistability for a particular interval of bias amplitudes if and only if $\Delta/\chi < 0$ and $|\Delta| \geq \frac{\sqrt{3}\kappa}{2} = \Delta_{th}$.

At any fixed bias amplitude and corresponding internal steady state mode amplitude the maximal gain experienced by a small signal is given by $g^{max} = |g_-| + |g_+|$. Here maximal means that we maximize over all possible signal input phases relative to the bias input. To experience this gain, the signal has to be in an appropriate quadrature defined by $\arg \delta\epsilon = \frac{\arg g_- - \arg g_+}{2}$. The orthogonal quadrature is then maximally de-amplified by a gain of $||g_-| - |g_+||$ and it is possible to show that for negligible losses the perfect squeezing relationship $(|g_-| + |g_+|) \, ||g_-| - |g_+|| = \left| |g_-|^2 - |g_+|^2 \right| = 1$ holds for any bias amplitude. Furthermore, for fixed cavity parameters g^{max} is maximized at a particular non-zero intra-cavity photon amplitude

$$|\alpha_0^{max}|^2 = \sqrt{\frac{\Delta^2 + \frac{\kappa^2}{4}}{3\chi^2}} \tag{5.15}$$

$$\Rightarrow g^{\max} = \sqrt{\frac{\sqrt{f}+\kappa}{\sqrt{f}-\kappa}}, \text{ with } f = 28\Delta^2 + 4\kappa^2 - 8\Delta\sqrt{12\Delta^2 + 3\kappa^2}. \tag{5.16}$$

Note that the maximal gain does not depend on the strength of the non-linearity. The relationship between g^{max} and Δ can be inverted:

$$\Delta = \frac{\sqrt{3}\kappa}{2} \frac{\left(g^{\max} - \sqrt{3}\right)\left(g^{\max} - \frac{1}{\sqrt{3}}\right)}{g^{\max 2} - 1} \tag{5.17}$$

Using all this it is straightforward to construct a tunable Kerr-amplifier. The symmetric construction proposed in Section 5.2.1 provides the additional advantage that one does not have to cancel the scattered bias. It is also convenient to prepend and append phase shifters to the signal input and output that ensure $g_-, g_+ > 0$ at maximum gain, implying that the maximally amplified quadrature is the real one.

The quadrature filter construction relies on the presence of additional cavity losses that are equal to the input coupler $\kappa_2 = \kappa_1 = \kappa$. In this case the gain coefficients for reflection of the first port are given by

$$g_- = 1 + \frac{\kappa\left[-\kappa + i\Delta + 2i\chi|\alpha_0|^2\right]}{\kappa^2 + \left(\Delta + 2\chi|\alpha_0|^2\right)^2 - |\chi|^2|\alpha_0|^4}, \tag{5.18}$$

$$g_+ = \frac{i\kappa\chi\alpha_0^2}{\kappa^2 + \left(\Delta + 2\chi|\alpha_0|^2\right)^2 - |\chi|^2|\alpha_0|^4}, \tag{5.19}$$

$$\epsilon_0 = -\frac{1}{\sqrt{\kappa}}\left[\kappa + i(\Delta + i\chi|\alpha_0|^2)\right]\alpha_0. \tag{5.20}$$

and one may easily verify that for dynamic resonance, i.e., $\chi|\alpha_0|^2 = -\Delta$, the gain coefficients are equal in magnitude $|g_-| = |g_+|$ which implies that there exists an input phase for which the reflected signal vanishes.

Two mode Kerr-nonlinear resonator

We label the mode amplitudes as α_1 and α_2. In this case the nonlinearity includes a cross-mode induced detuning

$$A_{\mathrm{NL}}^{\mathrm{Kerr}2}(\alpha) = \begin{pmatrix} -i\chi_a|\alpha_1|^2\alpha_1 - i\chi_{ab}|\alpha_2|^2\alpha_1 \\ -i\chi_{ab}|\alpha_1|^2\alpha_2 - i\chi_b|\alpha_2|^2\alpha_2 \end{pmatrix} \tag{5.21}$$

The model matrices are

$$A = \begin{pmatrix} -\frac{\kappa_{a,T}}{2} - i\Delta_a & 0 \\ 0 & -\frac{\kappa_{b,T}}{2} - i\Delta_b \end{pmatrix}, \tag{5.22}$$

$$B = -\begin{pmatrix} \sqrt{\kappa_{a,1}} & \sqrt{\kappa_{a,2}} & \dots & \sqrt{\kappa_{a,n_a}} & 0 & \dots & 0 \\ 0 & 0 & \dots & 0 & \sqrt{\kappa_{b,1}} & \sqrt{\kappa_{b,2}} & \dots & \sqrt{\kappa_{b,n_b}} \end{pmatrix}, \tag{5.23}$$

$$C = -B^T, \tag{5.24}$$

and the corresponding SLH model is

$$\left(\mathbf{1}_{n_a+n_b}, C\begin{pmatrix} a \\ b \end{pmatrix}, \tilde{\Delta}_a a^\dagger a + \tilde{\Delta}_b b^\dagger b + \frac{\chi_a}{2} a^{2\dagger} a^2 + \frac{\chi_b}{2} b^{2\dagger} b^2 + \chi_{ab} a^\dagger a b^\dagger b\right), \tag{5.25}$$

with $\tilde{\Delta}_{a/b} = \Delta_{a/b} + \chi_{a/b} + \frac{\chi_{ab}}{2}$ and where the Wigner-correspondence[3] is $\langle \alpha_1 \rangle_{\mathrm{W}} = \langle a \rangle$, $\langle \alpha_2 \rangle_{\mathrm{W}} = \langle b \rangle$.

We briefly summarize how to construct a controlled phase shifter using an ideal two-mode Kerr cavity with a single input coupling to each mode and negligible additional internal losses. We exploit that in this case the reflected steady state signal amplitude ζ' is identical to the input amplitude ζ up to a power dependent phase shift

$$\zeta' = -\frac{\frac{\kappa_a}{2} - i\left(\Delta_a + i\chi_a|\alpha_0|^2 + i\chi_{ab}|\beta_0|^2\right)}{\frac{\kappa_a}{2} + i\left(\Delta_a + i\chi_a|\alpha_0|^2 + i\chi_{ab}|\beta_0|^2\right)}\zeta \quad \Rightarrow |\zeta'| = |\zeta|. \tag{5.26}$$

We assume that the control input amplitude takes on two discrete values $\xi = 0$ or $\xi = \xi_0$ and that variations of the signal input amplitude are small $|\zeta| \approx |\zeta_0|$. In this case a good choice of detunings and coupling rates is given by

$$\Delta_a = \frac{\kappa_a}{2} - \frac{2\chi_a|\zeta_0|^2}{\kappa_a} \tag{5.27}$$

$$\Delta_b = \frac{\kappa_a \chi_b}{\chi_{ab}} - \frac{2\chi_{ab}|\zeta_0|^2}{\kappa_a} \tag{5.28}$$

$$\xi_0 = \frac{\sqrt{\kappa_a \kappa_b}}{2\sqrt{|\chi_{ab}|}} \tag{5.29}$$

in addition to two inequality constraints

$$\Delta_a \leq \sqrt{3}\frac{\kappa_a}{2} \tag{5.30}$$

$$\Delta_b \leq \sqrt{3}\frac{\kappa_b}{2} \tag{5.31}$$

that ensure that the system is stable. This construction ensures that $\frac{\zeta'|_{\xi=\xi_0}}{\zeta'|_{\xi=0}} = -1$ and in fact it can easily be generalized to the more realistic case of non-negligible internal losses.

---

[3] In this appendix we denote expectations with respect to the Wigner function as $\langle \cdot \rangle_{\mathrm{W}}$ and quantum mechanical expectations as $\langle \cdot \rangle$.

Finally note that the inequality constraints imply that the lower bounds for the input couplings scale as $\kappa_a^{\rm min}, \kappa_b^{\rm min} \propto |\zeta_0|$ which is important for our power analysis in Section 5.3.1. This, in turn implies that $\xi_0 \propto |\zeta_0|$ which is a fairly intuitive result.

The controlled phase shifter can now be included in one arm of a Mach-Zehnder interferometer to create a Fredkin gate (cf. Section 5.2.4).

To realize a thresholder, the control mode input is prepended with a two port Kerr-cavity with parameters chosen such that it becomes dynamically resonant with maximal differential transmission gain close to where its output gives the correct high control input ξ_0.

Overall, we remark that even when we account for the prepended cavity, the relationship $c \propto |\zeta_0|$ still holds, where c is the input to the thresholder. To see how the total decay rate of the thresholding cavity κ_{thresh} scales consider first that to get maximum differential gain or contrast, we ought pick a detuning right at or below the Kerr stability threshold $\Delta \approx \Delta_{\rm th} = \sqrt{3}\kappa_{thresh}/2$.

We choose the maximum input amplitude such that it approximately achieves dynamic resonance within the prepended thresholding cavity. This occurs when $\Delta = -\chi|\alpha_0|^2$ (cf. Appendix 5.5.2) and at an input amplitude of $c \propto \sqrt{\kappa_{thresh}\left|\frac{\Delta}{\chi}\right|} \propto \kappa_{thresh}$.

NOPO model

The NOPO model has consists of three modes, the signal and idler modes α_s, α_i and the pump mode α_p. We assume a triply resonant model[4] and that $\omega_s + \omega_i = \omega_p$, allowing for resonant conversion of pump photons into pairs of signal and idler photons and vice versa. The nonlinearity is given by

$$A_{\rm NL}^{NOPO}(\alpha) = \begin{pmatrix} \chi\alpha_i^*\alpha_p \\ \chi\alpha_s^*\alpha_p \\ -\chi\alpha_s\alpha_i \end{pmatrix} \tag{5.32}$$

and the model matrices are

$$A = \begin{pmatrix} -\frac{\kappa}{2} & 0 & 0 \\ 0 & -\frac{\kappa}{2} & 0 \\ 0 & 0 & -\frac{\kappa_p}{2} \end{pmatrix}, \quad B = -\begin{pmatrix} \sqrt{\kappa} & 0 & 0 \\ 0 & \sqrt{\kappa} & 0 \\ 0 & 0 & \sqrt{\kappa_p} \end{pmatrix}, \tag{5.33}$$

$$C = -B^T. \tag{5.34}$$

---

[4] It is possible to drop this resonance assumption for the pump.

Here, the SLH model is given by

$$\left(\mathbf{1}_3, C\begin{pmatrix} a \\ b \\ c \end{pmatrix}, i\chi\left(abc^\dagger - a^\dagger b^\dagger c\right)\right) \tag{5.35}$$

where now a, b and c correspond to α_s, α_i and α_p.

A steady state analysis of the system driven only by a pump input amplitude ϵ reveals that below a critical threshold $|\epsilon| < \epsilon_{th} = \frac{\kappa\sqrt{\kappa_p}}{4\chi}$ the system as a unique fixpoint with $\alpha_s = \alpha_i = 0$ and $\alpha_p = -\frac{2\epsilon}{\sqrt{\kappa_p}}$. Above threshold $|\epsilon| \geq \epsilon_{th}$, the intra-cavity pump amplitude stays constant at the threshold value $\alpha_p = -\frac{2\epsilon_{th}\epsilon/|\epsilon|}{\sqrt{\kappa_p}} = -\frac{\kappa\epsilon/|\epsilon|}{2\chi}$ and the signal and idler mode obtain non-zero magnitude

$$|\alpha_s| = |\alpha_i| = \sqrt{\frac{4\epsilon_{th}}{\kappa}(|\epsilon| - \epsilon_{th})}. \tag{5.36}$$

As an interesting consequence of the model's symmetry there exists not a single above threshold state but a whole manifold of fixpoints parametrized by a correlated signal and idler phase

$$\alpha_s = \sqrt{\frac{4\epsilon_{th}}{\kappa}(|\epsilon| - \epsilon_{th})}e^{i\phi + i\phi_0} \tag{5.37}$$

$$\alpha_i = \sqrt{\frac{4\epsilon_{th}}{\kappa}(|\epsilon| - \epsilon_{th})}e^{-i\phi + i\phi_0} \tag{5.38}$$

where the common phase ϕ_0 is fixed by the pump input phase via

$$\alpha_s\alpha_i = -\frac{4\epsilon_{th}}{\kappa}(|\epsilon| - \epsilon_{th})\frac{\epsilon}{|\epsilon|}. \tag{5.39}$$

In particular, for $\epsilon < 0$ we have $\alpha_i = \alpha_s^*$. Above threshold the system will rapidly converge to a fixpoint of well-defined phase ϕ. Without quantum shot noise ϕ would remain constant. With noise, however, the system can freely diffuse along the manifold. When the pump bias input is sufficiently large compared to threshold and consequently there are many signal and idler photons present in the cavity at any given time ($|\alpha_{s/i}|^2 \gg 1$) one can analyze the dynamics along the manifold and of small orthogonal deviations from the manifold. In the symmetric case considered here where signal and idler have equal decay rates, the differential phase degree of freedom $\phi = \frac{\arg\alpha_i - \arg\alpha_s}{2}$ decouples from all other variables and approximately obeys the SDE

$$d\phi = \sqrt{\gamma_\phi}dW_t, \quad dW_t^2 = dt \tag{5.40}$$

$$\text{with } \gamma_\phi = \frac{\kappa}{8|\alpha_s|^2} = \frac{\kappa^2}{32\epsilon_{th}(|\epsilon| - \epsilon_{th})}. \tag{5.41}$$

It is relatively straightforward to generalize these results to a less symmetric model with different signal and idler couplings and even non-zero detunings, but for a given nonlinearity the model considered here provides the smallest phase diffusion and thus the best analog memory. For a very thorough analysis of this model we refer to [45].

6. Exact co-simulation of semi-classical and quantum dynamics

The following work is as of yet unpublished and was partially done in collaboration with Nina Hadis Amini. It will be published soon [114].

6.1. Motivation

A given quantum mechanical system can be described in more than one way. Our choice of description is usually motivated by the insight it provides, its economy and, when dealing with sufficient complexity, the accuracy and efficiency with which it can be numerically simulated and analyzed.

The description in terms of a full quantum state $|\psi_t\rangle$ (or more generally a mixed state ρ_t) is the most complete because it allows for predicting arbitrary expectation values and correlations. On the other hand, the exponential scaling of the state space severely limits the kind of systems one may study numerically. For highly ordered systems such as one- or even two-dimensional lattices with local interactions, Matrix Product States (MPS) and more generally Projected Entangled Pair States (PEPS) [128] have proven immensely useful in accurately representing ground states as well as the low-lying excitations that are accessible in the dynamicsof such systems. The numerical representation of MPS and PEPS scales only polynomially in the number of degrees of freedom at the expense of losing the nice linear structure of a full Hilbert space vector representation.

In this work we propose a method that even works for systems that are unconstrained in the symmetry of their interactions and that may even be strongly excited as long as they remain fairly close to a set of near-classical states.

Quantum quasi-probability distributions [132, 52, 105] can be used to derive semi-classical stochastic dynamics suitable for sampling operator expectation values. These, however, are are sometimes numerically unstable and in general require approximations that can lead to significant discrepancies with simulations based on a full quantum state description [59]. Furthermore, there does not exist a general approach to incrementally increase the accuracy of these methods. Both our approach and the semi-classical methods based on quasi-probabilities are most useful for systems in which some observables become well localized.

Open quantum systems such as linear or non-linear optical resonators that arise in the context of cavity QED often feature dissipative dynamics that lead to localization of certain operators. This enables the description and simulation of systems containing, e.g., n coupled oscillators in a moving basis [90, 103] parametrized by coherent displacement coordinates $\alpha \in \mathbb{C}^n$. This approach and the quantum state diffusion (QSD) software package [100] published by the same authors is very useful for simulating quantum optical models, but as it is, the method does not easily generalize to non-bosonic degrees of freedom (such as ensembles of spins or fermions), it does not yield equations of motion of the coordinates, and as we demonstrate it can actually lead to higher computational complexity when applied to certain problems.

In this work we introduce a set of methods that significantly generalizes the QSD method in multiple ways. It can also be applied to non-bosonic physical degrees of freedom and even for oscillator modes it allows exploiting more general transformations than coherent displacements. By reformulating the problem of finding good basis coordinates as an optimization problem, we can derive improved update rules for the adaptive basis coordinates and even derive a system of combined stochastic differential equations for the basis coordinates and the quantum state.

6.2. Quantum state compression

Before turning to the problem of simulating a quantum system in an adaptive basis, we discuss different options for quantifying the efficiency of representing a given state in a particular basis and we introduce a corresponding optimization problem. We will always assume that our adaptive Hilbert Space bases are related to the original fixed basis by means of a unitary transformation U_θ, $U_\theta U_\theta^\dagger = \mathbb{1}$ smoothly parametrized by a set of coordinates θ. Although our method allows for arbitrary unitary representations of Lie groups some Lie group manifolds cannot be fully parametrized by a single coordinate patch. This can lead to additional technical difficulties which we will usually avoid by limiting ourselves to a single open coordinate patch $\theta \in D \subset \mathbb{R}^n$ that includes the origin 0 which we always assume to map to the identity $U_0 = \mathbb{1}$. This poses no serious limitation as closure under group multiplication is not generally required for our scheme. For any $\theta \in D$ the partial derivatives of the transform with respect to the coordinates are given by $\frac{\partial U_\theta}{\partial \theta^j} = -iU_\theta F_j^>(\theta_t)$ which implicitly defines the *right generators* $F_j^>(\theta_t) := iU_\theta^\dagger \frac{\partial U_\theta}{\partial \theta^j}$ that locally describe the

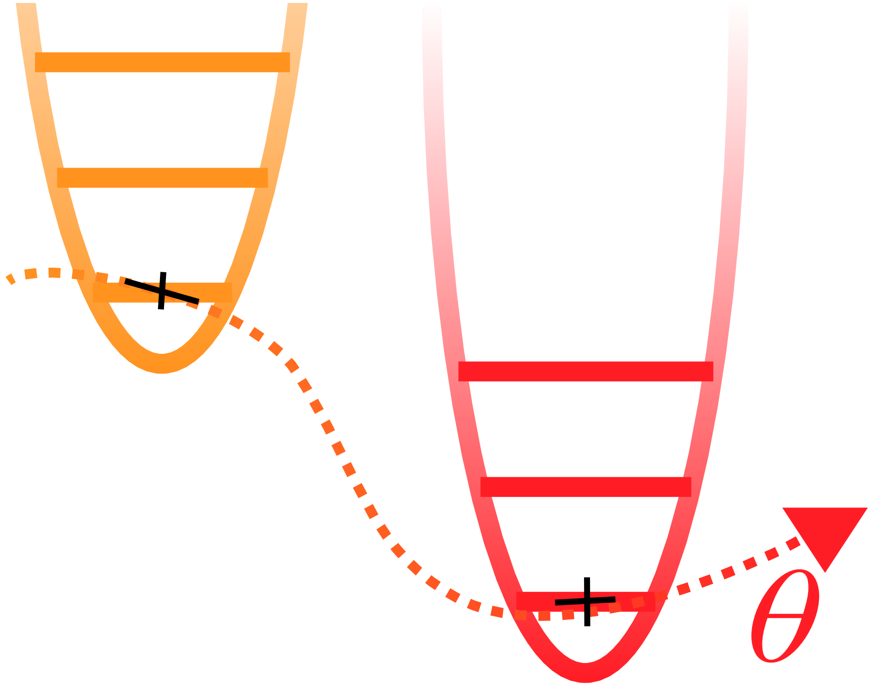


Figure 6.1.: Cartoon visualizing how a unitary Lie group representation can induce a coordinate manifold with a localized orthonormal basis set defined at each point. The group coordinates correspond to semi-classical phase space variables. If the group representation is irreducible and the group semi-simple, then the 'ground state' attached to each manifold point can be understood as a (generalized) coherent state [91]. For open, diffusive quantum dynamics, quantum states often localize in the vicinity of such generalized coherent states.

transformation

$$U_{\theta+d\theta} = U_\theta \left[1 - i \sum_{j=1}^{n} F_j^{>}(\theta_t) d\theta^j \right]. \tag{6.1}$$

The right generators are hermitian elements of the group's Lie algebra and their explicit form depends on the choice of parametrization of the transform. We present several examples and alternate constructions in the Appendix 6.7.1. We note that any such parametrization is not unique. We can smoothly re-parametrize the coordinates and then derive the transformed generators via the chain rule. Our use of upper indices for the coordinates and lower indices for the generators is thus motivated by their covariant and contravariant transformation under such re-parametrizations.

If our state in the original, fixed basis is $|\psi_t\rangle$ we assume that it can be related to a *reduced complexity state* $|\phi_t\rangle$ via

$$|\psi_t\rangle = U_{\theta_t} |\phi_t\rangle \Leftrightarrow |\phi_t\rangle = U_{\theta_t}^\dagger |\psi_t\rangle \tag{6.2}$$

By itself, equation (6.2) gives an over-parametrization of the original state $|\psi_t\rangle$ in terms of $(|\phi_t\rangle, \theta_t)$. We now outline how to remove this redundancy by specifying additional constraints on $|\phi_t\rangle$ (and thus θ_t). In doing this we attempt to make the description of $|\phi_t\rangle$ less complex, in the very simple sense that for a given numerical accuracy, it can be represented in as small a basis as possible.

6.2.1. The complexity functional

Consider first a canonical example: for a single bosonic degree of freedom with lowering operator a, and a transformation group given by coherent displacements $U_\theta = D(\theta) = e^{\theta a^\dagger - \theta^* a}$, an intuitive constraint would be to demand that $\langle a\rangle_{\phi_t} = 0$, or equivalently $\langle a\rangle_{\psi_t} = \theta$. This fully fixes the coordinates and removes the redundancy. This is precisely the constraint which the QSD software package implements. As we demonstrate later, this method works very well for nearly coherent states $|\psi_t\rangle$ but can actually *increase* the complexity when $|\psi_t\rangle$ has significant non-Gaussian features.

Below we re-derive this as the result of an optimization problem, which will allow us to generalize and improve the approach. To formulate the problem, we introduce a complexity functional $\mathcal{C}(\theta;\psi)$ which, given a state ψ, attains a unique global minimum, i.e., we define our coordinates at all times to be

$$\theta_t := \theta_*^{(\mathcal{C})}(\psi_t) = \mathrm{argmin}_\theta\, \mathcal{C}(\theta;\psi). \tag{6.3}$$

Excitation minimization

The simplest choice of complexity functional is obtained by evaluating the expectation of a lower bounded operator N that penalizes population of undesired basis levels in the transformed basis.

$$\mathcal{C}_N(\theta;\psi) := \underbrace{\left\langle U_\theta N U_\theta^\dagger \right\rangle_\psi}_{\langle N\rangle_\phi}. \tag{6.4}$$

For a bosonic degree of freedom, N could simply be the canonical number operator $a^\dagger a$. As we will see below, this particular choice leads to the QSD scheme $\theta^* = \langle a\rangle_{\psi_t}$.

The example of a counting operator suggests that we might generally introduce a partial order for our basis levels and that according to such an order, a lower complexity state is characterized by being confined to a subspace spanned by basis states of very low order.

While in principle any full enumeration $|s_0\rangle, |s_1\rangle, \dots$ of the basis implies such an ordering, it turns out that given some specific physical degrees of freedom there exist certain canonical orderings related to the existence of counting operators and associated raising and lowering operators. We will generally refer to this class of optimization problems, i.e., minimizing the expectation of counting operators, as *excitation minimization.*

As it turns out, excitation minimization can also be interpreted as minimizing the quantum relative entropy between $|\psi_t\rangle$ and a transformed thermal state $\chi_{\theta_t,\beta} := U_{\theta_t}\rho_\beta U_{\theta_t}^\dagger$ where $\rho_\beta := Z(\beta)^{-1}e^{-\beta N}$ and $Z(\beta) = \mathrm{Tr}\left(e^{-\beta N}\right)$. The quantum relative entropy between $|\psi_t\rangle\langle\psi_t|$ and $\chi_{\theta_t,\beta}$ is given by

$$\begin{aligned} S\left(|\psi_t\rangle\langle\psi_t| \,\|\chi_{\theta_t,\beta}\right) &= \mathrm{Tr}\left(|\psi_t\rangle\langle\psi_t|\left[\log|\psi_t\rangle\langle\psi_t| - \log\chi_{\theta_t,\beta}\right]\right) = -\langle\psi_t|\log\chi_{\theta_t,\beta}|\psi_t\rangle &\quad (6.5)\\ &= -\langle\psi_t|\log e^{-\beta U_{\theta_t}NU_{\theta_t}^\dagger}|\psi_t\rangle + \log Z(\beta) &\quad (6.6)\\ &= \beta\left\langle U_{\theta_t}NU_{\theta_t}^\dagger\right\rangle_{\psi_t} + \log Z(\beta) &\quad (6.7)\end{aligned}$$

Here the last term $\log Z(\beta)$ does not depend on θ and therefore minimizing the quantum relative entropy is exactly equivalent to excitation minimization as discussed above. In the second equality we exploited that pure states have zero entropy, but even in the mixed state case, its entropy would not depend on θ_t and thus not play a role in the optimization problem.

CGF minimization

A useful alternative is to more generally minimize the following *cumulant generating function* (CGF) of any such counting operator

$$\mathcal{C}_N^{\mathrm{cgf}}(\theta;\psi,\lambda) := \log\left[\underbrace{\left\langle U_\theta e^{\lambda N}U_\theta^\dagger\right\rangle_\psi}_{\left\langle e^{\lambda N}\right\rangle_\phi}\right], \quad (6.8)$$

which, for very small $0 < \lambda \ll 1$ reduces to

$$\mathcal{C}_N^{\mathrm{cgf}}(\theta;\psi,\lambda) \approx \lambda\,\langle N\rangle_\phi + \frac{\lambda^2}{2}\,\mathrm{var}\,(N)_\phi + O(\lambda^3). \quad (6.9)$$

Note that when N is unbounded there may generally exist normalizable states $|\phi\rangle$ for which $\left\langle e^{\lambda N}\right\rangle_\phi$ diverges for any $\lambda > 0$, but this is true even for the unexponentiated counting operator[1]. We will assume here without proof that such states do not actually arise in the dynamical evolution of open quantum systems if one starts from a well behaved initial state. In general if there exists an $N_0 \in \mathbb{N}$ and a constant $0 \le \alpha < 1$ such that $\forall n \ge N_0$ we find $\frac{|\phi_{n+1}|^2}{|\phi_n|^2} \le \alpha$, then $\left\langle e^{\lambda N}\right\rangle_\phi$ exists if $\lambda < \log 1/\alpha$.

---

[1] Consider $N = a^\dagger a$ and $\phi_n = \frac{\sqrt{6}}{n\pi}$, $n \in \mathbb{N}$ then $1 = \sum_n |\phi_n|^2$, but $\langle N\rangle_\phi = \infty$.

This objective yields a Chernoff bound[2] on the population of highly excited levels:

$$\mathbb{P}_\phi\left[N \geq N_0\right] \leq \frac{\left\langle e^{\lambda N}\right\rangle_\phi}{e^{\lambda N_0}} = e^{\mathcal{C}_N^{\mathrm{cgf}}(\lambda)-\lambda N_0}. \tag{6.10}$$

Since the equality is satisfied for any λ we can minimize the right hand side over λ to achieve the most restrictive bound:

$$\log \mathbb{P}_\phi\left[N \geq N_0\right] \leq \tilde{\mathcal{C}}_N^{\mathrm{cgf}}(N_0) := \min_\lambda \mathcal{C}_N^{\mathrm{cgf}}(\lambda) - \lambda N_0. \tag{6.11}$$

The negated left hand side in the equation above is proportional to the digits of relative accuracy obtained when truncating the basis at the level N_0. Given N_0 the optimal $\lambda(N_0)$ leading to the lowest bound is given via a Legendre transformation

$$\left.\frac{\partial \mathcal{C}_N^{\mathrm{cgf}}(\lambda)}{\partial \lambda}\right|_{\lambda(N_0)} = N_0. \tag{6.12}$$

We thus see that for a fixed N_0 our complexity functional specifies with what accuracy a given quantum state can be represented in a low-dimensional subspace of the overall state space. This is visualized in Figure 6.2.

We will generally refer to this family of optimization problems as *CGF minimization.*

Given a sufficiently smooth functional, we may expand it as

$$\mathcal{C}(\theta_t + \delta\theta; \psi_t) = \mathcal{C}(\theta_t; \psi_t) + \sum_{j=1}^n y_j(\theta_t; \psi_t)\delta\theta^j \tag{6.13}$$

$$+ \frac{1}{2}\sum_{j,k=1}^n m_{jk}(\theta_t; \psi_t)\delta\theta^j\delta\theta^k + O\left(\delta\theta^3\right) \tag{6.14}$$

If the complexity functional is strictly convex then the Hessian $m(\theta_t) = (m_{jk}(\theta_t))_{j,k=1}^n$ is positive definite everywhere and an appropriate variant of Newton's method can be applied to find the optimal coordinates which are implicitly defined by requiring the gradient to vanish $y_j(\theta_t; \psi_t) = 0,\ j = 1, 2, \ldots, n$.

For the case of expectation minimization, the explicit expressions for the gradient and Hessian are

$$y_j(\theta_t) = \left\langle -i[F_j^>(\theta_t), N]\right\rangle_{\phi_t} \tag{6.15}$$

$$m_{jk}(\theta_t) = m_{kj}(\theta_t) = \left\langle \left[F_j^>(\theta_t), \left[N, F_k^>(\theta_t)\right]\right] - i\left[\frac{\partial F_k^>(\theta_t)}{\partial \theta^j}, N\right]\right\rangle_{\phi_t}. \tag{6.16}$$

---

[2] This follows from Markov's inequality and the convexity of $\exp x$.

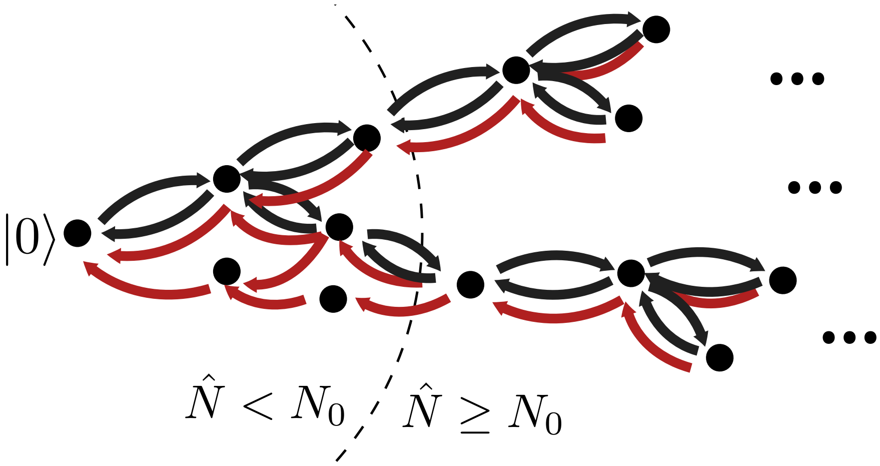


Figure 6.2.: Many physically relevant Lie groups admit unitary representations in which there exists a natural ordering of the basis states according to the spectrum of some generalized energy or counting operator. Here we visualize such basis levels by black dots and suggest that they are (partially) ordered from left to right. The black arrows represent coherent transitions induced by a Hamilton operator, while the red arrows indicate dissipation induced transitions. Transforming the dynamics to a parametrized basis induces a mapping from this graph to one with potentially more transitions, e.g., under a squeezing transformation $a \to \cosh ra + \sinh ra^\dagger$ which generally increases the terms of the Hamilton operator, but if the transformation coordinates are chosen wisely, the system state can be kept close to the left side of this graph, i.e., it is trapped in a low-dimensional subspace of the overall transformed basis.

In the case of cgf minimization one finds

$$\mathcal{C}_N^{\mathrm{cgf}}(\theta_t;\psi_t,\lambda) = \log \left\langle e^{\lambda N} \right\rangle_{\phi_t} \tag{6.17}$$

$$y_j(\theta_t) = \left\langle -i[F_j^{>}(\theta_t), e^{\lambda N - \mathcal{C}_N^{\mathrm{cgf}}(\theta_t;\psi_t,\lambda)}] \right\rangle_{\phi_t} \tag{6.18}$$

$$\begin{aligned} m_{jk}(\theta_t) = m_{kj}(\theta_t) = {}& \left\langle \left[F_j^{>}(\theta_t), \left[e^{\lambda N - \mathcal{C}_N^{\mathrm{cgf}}(\theta_t;\psi_t,\lambda)}, F_k^{>}(\theta_t)\right]\right] - i \left[\frac{\partial F_k^{>}(\theta_t)}{\partial \theta^j}, e^{\lambda N - \mathcal{C}_N^{\mathrm{cgf}}(\theta_t;\psi_t,\lambda)}\right] \right\rangle_{\phi_t} \\ & - y_j(\theta_t) y_k(\theta_t), \end{aligned} \tag{6.19}$$

However, since cgf minimization is fully equivalent to minimizing $\exp \mathcal{C}_N^{\mathrm{cgf}}(\theta_t;\psi_t,\lambda) = \left\langle e^{\lambda N} \right\rangle_{\phi_t}$, one could also use

$$y_j(\theta_t) = \left\langle -i[F_j^{>}(\theta_t), e^{\lambda N}] \right\rangle_{\phi_t} \tag{6.20}$$

$$m_{jk}(\theta_t) = m_{kj}(\theta_t) = \left\langle \left[F_j^{>}(\theta_t), \left[e^{\lambda N}, F_k^{>}(\theta_t)\right]\right] - i \left[\frac{\partial F_k^{>}(\theta_t)}{\partial \theta^j}, e^{\lambda N}\right] \right\rangle_{\phi_t}, \tag{6.21}$$

though in either case some care must be taken to avoid numerical overflow errors.

When $|\psi_t\rangle$ evolves with time we can now either solve the minimization problem (6.3) at each time and use this to obtain the coordinates θ_t or alternatively derive explicit (stochastic) differential equations for the coordinates. While the former will allow us to adapt our scheme to arbitrary stochastic dynamics (jump equations and Ito or Stratonovich diffusions) the latter method can provide us with more insight into the dynamics and open up interesting opportunities for designing control schemes.

6.2.2. Example application: The degenerate parametric oscillator

Before we describe how to co-simulate the quantum state and the manifold coordinates we illustrate the above principles by applying to one of the simplest yet highly nontrivial nonlinear oscillator models: the degenerate parametric oscillator which describes a resonator in which a high-Q resonant signal mode couples to a strongly driven pump mode at twice the frequency through a χ_2 non-linear parametric interaction that allows conversion of signal photon pairs to pump photons and vice versa. In the limit of strong nonlinearity and a small pump mode quality factor, the pump mode can be adiabatically eliminated yielding the following SLH model

$$\left(\mathbb{1}_2, \begin{pmatrix} \sqrt{\kappa} a \\ \sqrt{\beta} a^2 \end{pmatrix}, i\frac{\epsilon}{2}\left[a^{\dagger 2} - a^2\right] \right). \tag{6.22}$$

For positive $\epsilon > 0$ the dynamics are primarily captured by the evolution of the signal mode's real quadrature $q = (a + a^\dagger)/2$. A semi-classical differential equation for the quadrature amplitude $Q = \langle q \rangle$ is approximately given by

$$\dot{Q} \approx -\left(\frac{\kappa}{2} - \epsilon + \beta|Q|^2\right) Q \tag{6.23}$$

where we have assumed factorizing moments $\left\langle a^\dagger a^2 \right\rangle \approx \langle a \rangle^* \langle a \rangle^2$ as well as $\mathrm{Im}\left[\langle a \rangle\right] = 0$. Equation (6.23) is actually identical to the normal form of a pitchfork bifurcation up to some re-scaling. We visualize the bifurcation diagram in Figure 6.3. The bifurcation exists for any non-zero two-photon loss rate $\beta > 0$, however the magnitude of the two-photon strongly affects how non-classical (which in this context we take to mean non-Gaussian) the state of the signal mode becomes. When the system is pumped only slightly above the threshold, random switching or tunneling between the two equilibria is possible. We present such a trajectory in Figure 6.4. The switching rate strongly depends on the ratio of linear to two-photon loss. Additionally, in the strongly non-linear case $\beta \geq \kappa$, the system can spontaneously evolve into cat-like states that feature exhibit significant simultaneous overlap with coherent states centered at either equilibrium.

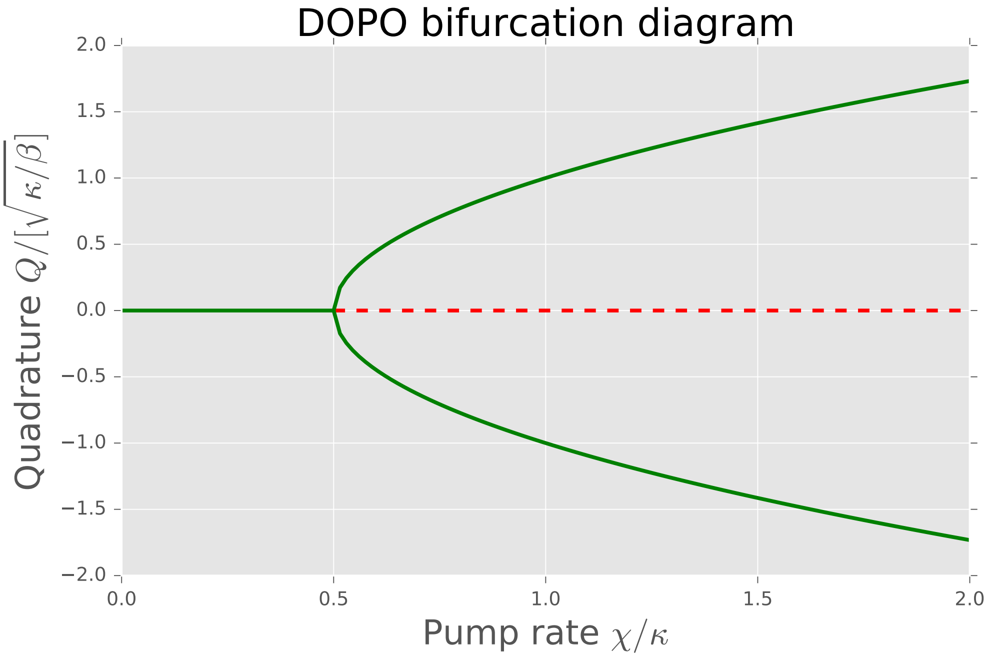
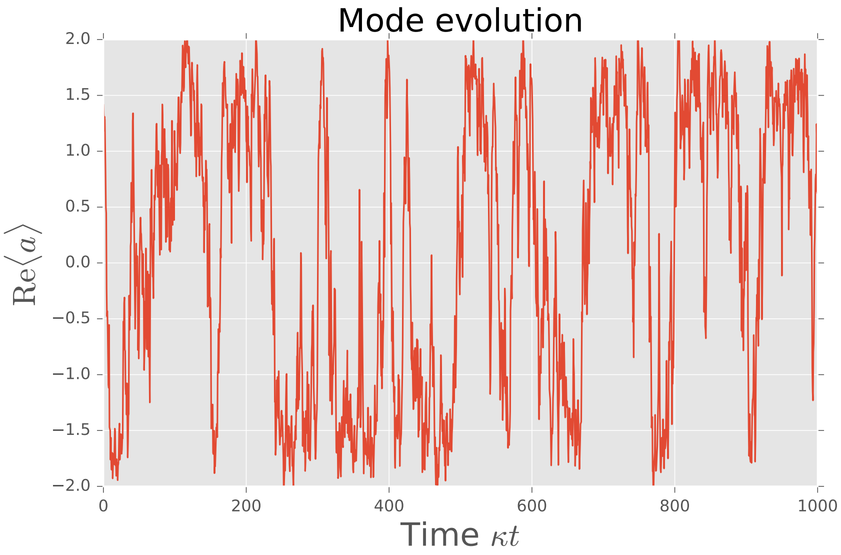
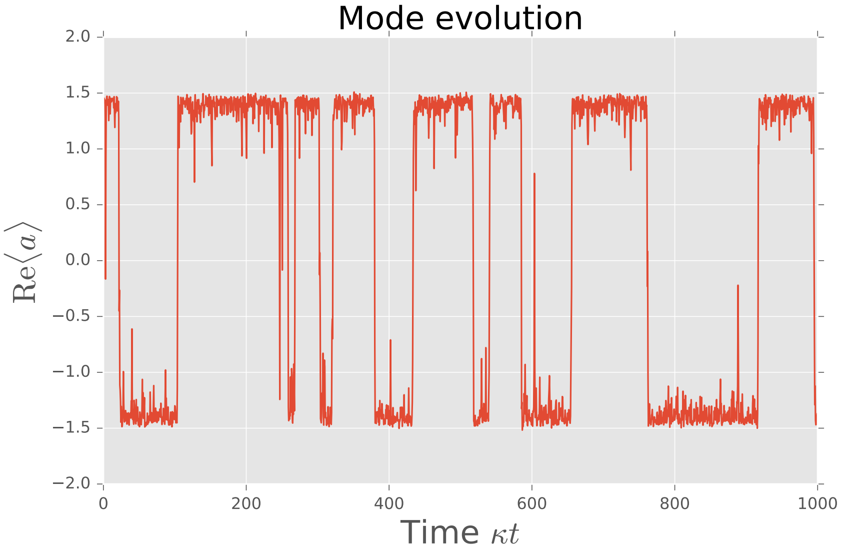


Figure 6.3.: When the linear loss is larger than the gain $\kappa/2 < \epsilon$ the real quadrature Q remains stably at 0, above a critical pump $\epsilon \geq \kappa/2$ this fixpoint bifurcates into two stable symmetric solutions and an unstable solution that is the continuation of the below threshold $Q = 0$ solution.

In the limit of vanishing *linear* loss $\kappa/\beta \to 0$, the system has a decoherence-free submanifold spanned by the two equilibrium amplitude coherent states. In [78] Mirrahimi et



(a) $\beta/\kappa = 1/12$

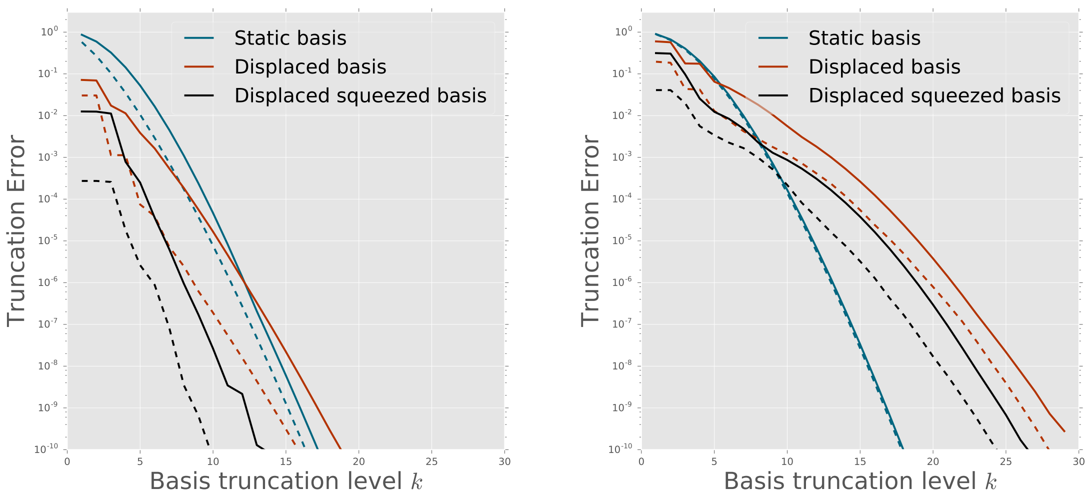


(b) $\beta/\kappa = 1$

Figure 6.4.: Stochastic switching dynamics of a DOPO above threshold. Figure (a) shows an example with very weak non-linear loss $\beta \ll \kappa$ whereas (b) shows the strongly nonlinear case $\beta = \kappa$. In both cases we have chosen the parameters β, κ, χ such that the bi-stable mode amplitude equals approximately $\alpha_{r,ss} = \pm 1/\sqrt{2}$. There is a visible reduction in the switching rate and we can also see quite clearly that the magnitude of fluctuations in either bi-stable state is strongly reduced in the case of very strong nonlinearity. Specifically, the simulation parameters were $\beta = \kappa, \chi = 5\kappa/2$ and $\beta = \kappa/12, \chi = 2\kappa/3$ for the strongly and weakly nonlinear case, respectively.

al. outline a scheme to encode quantum information in such a system. A detailed study of the switching dynamics was carried out in [59].

In Figure 6.2.2 we compare how each basis level contributes to a whole trajectory of states when represented in the original fixed basis to the excitation minimization when using either a coherently displaced basis or a displaced and squeezed basis. We see that the adaptive schemes perform well in the case of strong linear dissipation but not so well in the case of strong two-photon loss. We can understand this better by inspecting typical

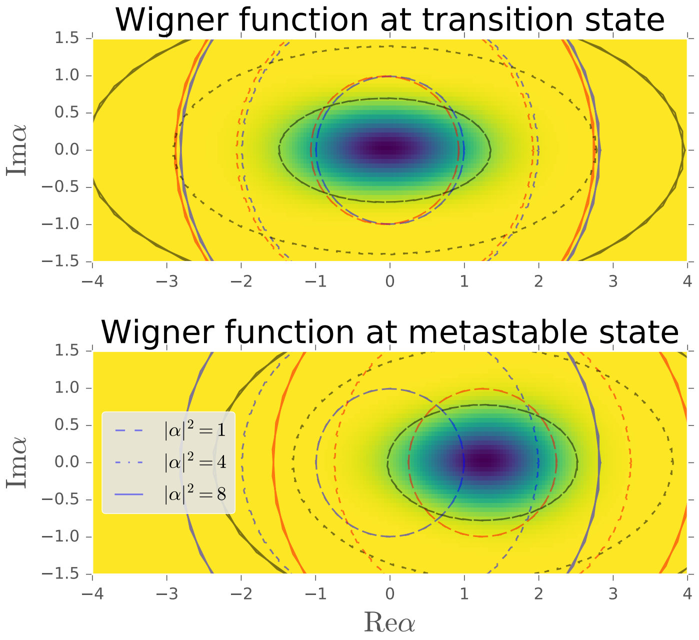


(a) $\beta/\kappa = 1/12$

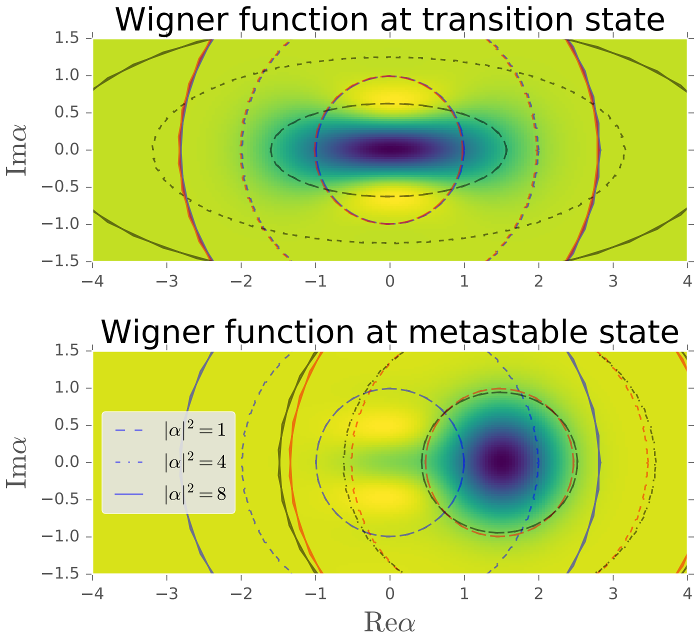
(b) $\beta/\kappa = 1$

a̧ptionFor the weakly nonlinear case (a) both the displaced basis and the displaced and squeezed basis perform fairly well, although the displaced basis truncation error falls off less rapidly than either the static or the displaced, squeezed basis. For strongly nonlinear case (b), however, we find that the static basis outperforms both the displaced and the displaced, squeezed basis. This indicates that the system dynamics depart significantly from the squeezed and displaced coherent state manifold.

states that occur in each evolution. In Figure 6.5 we present snapshots of the signal mode's Wigner function. For strong linear dissipation, the Wigner function of the signal mode typically appears quite Gaussian in shape, whereas in the strong two-photon loss case we see significant non-Gaussian features both in the transition states and when the mode is at one of the equilibria. The bad performance of the excitation minimization functional in the non-Gaussian case is much improved by the cgf minimization approach. In Figure 6.6 we compare the efficiency of the fixed basis with a coherently displaced basis where the coordinates are determined either by excitation minimization or by cgf minimization. We find that the cgf minimization (for $\lambda = 3/2$) outperforms both the fixed basis and the



(c) $\beta/\kappa = 1/12$



(d) $\beta/\kappa = 1$

Figure 6.5.: Comparing the Wigner functions of either system in typical transitions states and typical meta-stable states we see clearly that the Wigner functions of the strongly nonlinear system (b) appear much less Gaussian in shape than for the system dominated by linear dissipation. We have furthermore indicated the support set of different bases. The blue circles correspond to the fixed basis, the red circles to a coherently displaced basis and the black ellipses to a displaced and squeezed basis.

excitation minimization method (which is equivalent to the QSD package's approach). Here we have not even exploited the additional advantages that a displaced *and squeezed* basis may yield.

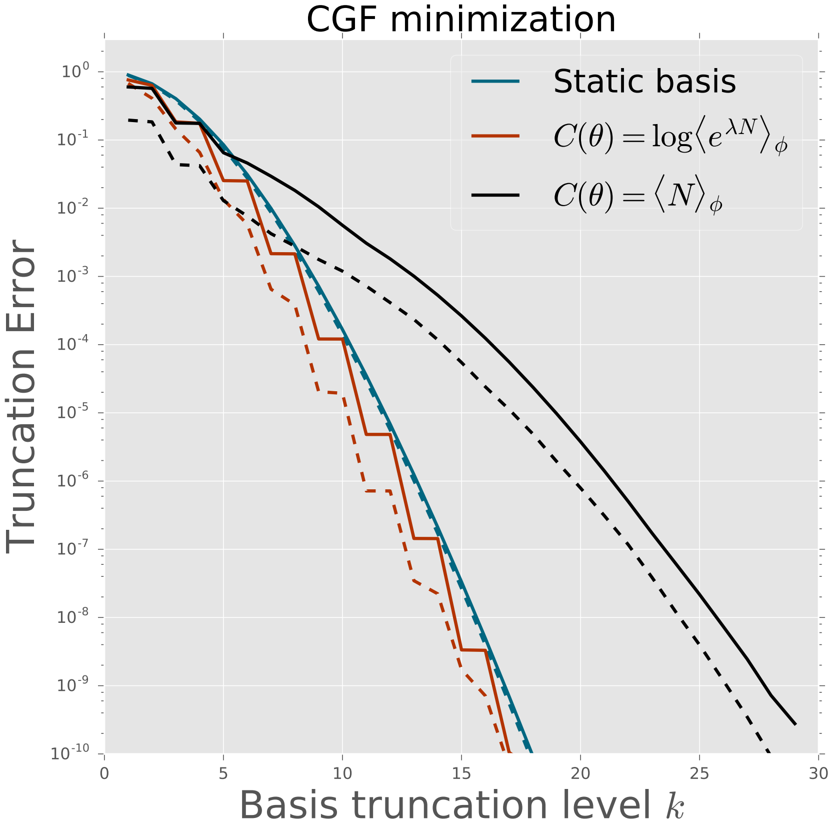


Figure 6.6.: When changing the optimization problem to CGF minimization, we see that achieve higher representation efficiency (but not by very much!) than the static basis while so far only using a displaced, non-squeezed basis.

While excitation minimization will always enforce $\langle a\rangle_{\phi_t} \equiv 0 \Leftrightarrow \langle a\rangle_{\psi_t} = \frac{Q+iP}{\sqrt{2}}$ CGF optimization generally does not lead to such a linear relationship as can be seen in Figure 6.7. We can also see that different regions in phase space lead to different complexity as measured by the CGF (cf Figure 6.8). This motivates using a simulation method in which even the basis size is adapted to the inherent complexity of the current dynamics. Finally, we note that even in the displaced basis there appear to be additional attractors for $|\phi_t\rangle$. In Figure 6.9 we have visualized the distribution of the first three moving basis level populations when transforming to the CGF optimal basis.

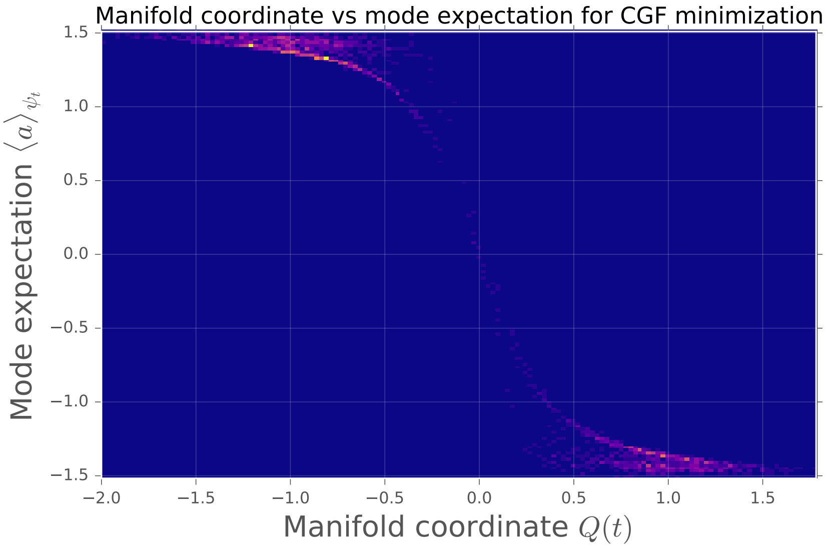


Figure 6.7.: The optimal manifold coordinate $Q(t)$ under CGF minimization appears mostly monotonically but not linearly related to the mode expectation $\langle a \rangle_{\psi_t}$.

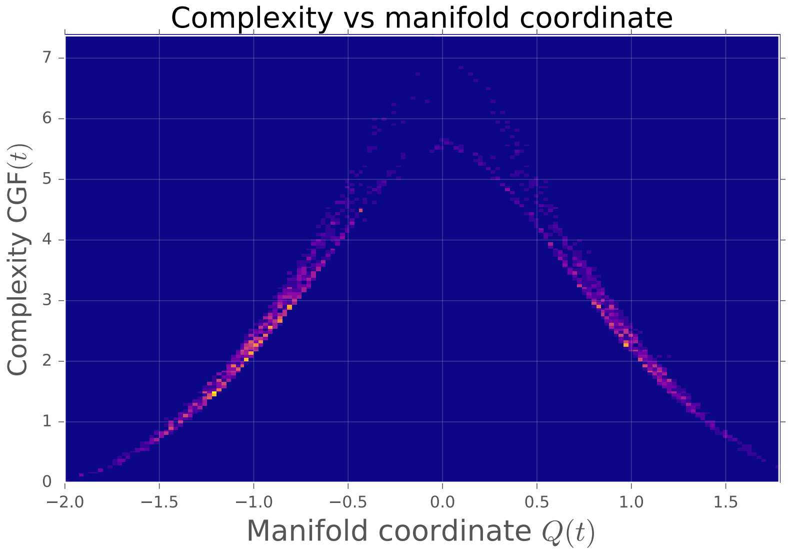


Figure 6.8.: When the system state localizes near the ‘origin’ i.e., $Q = 0$, the complexity increases, i.e., more basis levels are necessary for accurate representation.

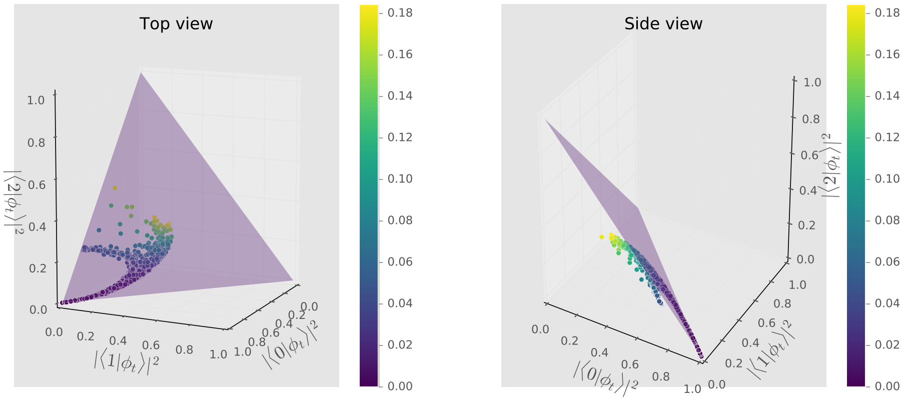


Figure 6.9.: The probability simplex spanned by the excitation probability of the first three basis levels. For the optimal CGF trajectory the basis level populations remain nearly confined to this simplex, but diverge slightly from it, especially when the first and second excited basis levels are nontrivially populated. The points are color coded according to their 'missing-probability-distance' d from the simplex, i.e. $d := 1 - \langle \Pi_0 + \Pi_1 + \Pi_2 \rangle_{\phi_t}$

6.3. Dynamics in a moving basis

Assume now that the fixed basis state vector $|\psi_t\rangle$ evolves according to $|d\psi_t\rangle = -idG\,|\psi_t\rangle$. For a closed system and deterministic dynamics we simply have $dG = \hbar^{-1} H dt$, for an open system evolving according to an unnormalized stochastic Schrödinger equation (SSE) we might have

$$dG = \left[\hbar^{-1} H - \frac{i}{2} \sum_k L_k^\dagger L_k\right] dt + i \sum_k dM^{k*} \circ L_k, \tag{6.24}$$

where the L_k are the Lindblad collapse operators. This particular complex diffusive stochastic unraveling of the open system dynamics can be interpreted as a system whose output bath modes are all measured using heterodyne detection with perfect fidelity. The associated complex heterodyne measurement processes are given by $dM^k = \langle L_k\rangle\, dt + dW_k$, with complex Wiener processes $dW_k dW_l = 0$, $dW_k^* dW_l = \delta_{kl} dt$. Note that our method naturally generalizes to mixed quantum states and deterministic or stochastic Lindblad master equations.

We take the SSE to be in Stratonovich form to ensure that the normal chain and product rules of calculus can be applied. Above, we write $X \circ dY$ to indicate a Stratonovich type SDE whereas XdY indicates an Ito SDE. For deterministic differentials the distinction is unnecessary $X \circ (Y dt) = XY dt$.

A slightly different definition of dG can be given for modeling homodyne measurements and while it is possible to extend our approach to discontinuous quantum jump equations we limit ourselves to diffusive unravellings here. For the heterodyne SSE, we actually find that the SDE assumes identical form in both the Ito and Stratonovich formalism, but this is not generally true. In the following we will assume all SDEs to be in Stratonovich form and simply write $X \circ dY$ as XdY to simplify the expressions.

As outlined in the previous section the state vector $|\psi_t\rangle$ in the fixed basis is related to the reduced complexity state vector $|\phi_t\rangle$ via

$$|\psi_t\rangle = U_{\theta_t} |\phi_t\rangle \Leftrightarrow |\phi_t\rangle = U_{\theta_t}^\dagger |\psi_t\rangle \ . \tag{6.25}$$

It then follows that the transformed state evolves according to a modified SSE

$$\begin{aligned} |d\phi_t\rangle &= -idK_\theta |\phi_t\rangle \,, \\ \text{with } \ dK_\theta &:= \underbrace{U_\theta^\dagger dG U_\theta}_{=:dG_\theta} - \sum_j F_j^{>}(\theta_t) d\theta^j . \end{aligned} \tag{6.26}$$

We see that the transformed state has dynamics generated not only by the transformed SSE generator $dG_\theta = U_\theta^\dagger dG U_\theta$ but also by the explicit time dependence of the unitary mapping $-\sum_j F_j^{>}(\theta_t) d\theta^j$. This is similar to the extra terms that arise when transforming a given system to an interaction picture.

6.4. Optimal coordinate dynamics

Assume that at a given time t we are starting at optimal coordinates, i.e., we have already solved for θ_t such that the complexity gradient $y_j(\theta_t) = 0, \ j = 1, 2, \ldots, n$. Then we can determine the coordinate increments $d\theta_t^j$ by requiring $y_j(\theta_{t+dt}) = 0, \ j = 1, 2, \ldots, n$. We may then derive the coordinate dynamics by computing the differential change of the gradient coefficients $dy_j(\theta_t)$ as a function of $d\psi_t$ and $d\theta$ solve for $d\theta$ such that $dy_j(\theta_t) = 0$.

More generally, if we assume that we are not starting exactly at optimal coordinates but close to the optimum, then we can instead choose a gain parameter $\eta > 0$ and solve for $d\theta$ such that

$$dy_j(\theta_t) = -\eta \, y_j(\theta_t) \, dt, \quad j = 1, 2, \ldots, n. \tag{6.27}$$

This reduces to the above case when we are already at the optimum coordinates, but for a good choice of η it leads to robustness to slight deviations as they are exponentially damped over time.

The expectation value of any hermitian observable $X = X^\dagger$ in ϕ_t evolves as

$$d\,\langle X\rangle_{\phi_t} = -2\,\mathrm{Im}\left[\sigma\left(dK_{\theta_t}, X\right)_{\phi_t}\right] \tag{6.28}$$

$$= -2\,\mathrm{Im}\left[\sigma\left(dG'_{\theta_t}, X\right)_{\phi_t}\right] - i\sum_{j=1}^{n}\left\langle\left[F_j^{>}(\theta_t), X\right]\right\rangle_{\phi_t} d\theta^j \tag{6.29}$$

Note that if X explicitly depends on time or θ one needs to add additional terms accordingly. Here we have used Percival's notation [90] for the *quantum correlation*

$$\sigma\left(A, B\right)_{\phi_t} := \left\langle A^\dagger B\right\rangle_{\phi_t} - \langle A\rangle^*_{\phi_t}\,\langle B\rangle_{\phi_t},$$

which defines a semi-definite inner product on the space of operators (cf. Appendix 6.7.2).

Combining Equations (6.27) and (6.28) we find

$$\sum_{k=1}^{n} m_{jk}(\theta_t)d\theta^k = dq_j + \eta\, y_j(\theta_t)\, dt, \quad j = 1, 2, \ldots, n, \tag{6.30}$$

$$\tag{6.31}$$

where we have defined the *bias flow*

$$dq_j := -2\,\mathrm{Im}\left[\sigma\left(dG'_{\theta_t}, Y_j^{>}(\theta_t)\right)_{\phi_t}\right]. \tag{6.32}$$

Assuming a positive definite Hessian $m(\theta_t) = (m_{jk}(\theta_t))_{j,k=1}^n$ this relationship can be solved for the coordinate differentials

$$d\theta^k = \sum_{j=1}^{n}\left[m(\theta_t)^{-1}\right]^{kj}\left[dq_j + \eta\, y_j(\theta_t)\, dt\right], \tag{6.33}$$

for $k = 1, 2, \ldots, n$.

6.4.1. Gradient coupled fiducial state dynamics

When it is known that the system state will remain localized near a semi-classical manifold of generalized coherent states $U_\theta\,|\Omega\rangle$, then it may be advantageous to evaluate the Hessian $(m_{jk}(\theta_t))$ and the drift term (dq_j) not in the actual moving basis state $|\phi_t\rangle$ but in the reference state $|\Omega\rangle$ instead. In this case, Equation (6.33) becomes

$$d\theta^k = \sum_{j=1}^{n}[m^{(\Omega)}(\theta_t)^{-1}]^{kj}\left[dq_j^{(\Omega)} + \eta\, y_j(\theta_t)\, dt\right], \tag{6.34}$$

for $k = 1, 2, \ldots, n$. The advantage of this is that $[m^{(\Omega)}(\theta_t)^{-1}]^{kj}$ and $dq_j^{(\Omega)}$ are purely functions of θ and any other model parameters and they can be understood as semi-classical equations of motion for the coordinates. Their dynamics are still coupled to the true quantum state via the gradient $\eta\, y_j(\theta_t)\, dt$.

Interestingly enough, for oscillator modes and excitation minimization, this approach yields the same equations of motion as the manifold projection method proposed by Hideo Mabuchi in [67]. We have some evidence that this generalizes to other cases in which $|\Omega\rangle$ is a minimum uncertainty state for the generators of U_θ but so far this is only a conjecture.

6.4.2. Computational results

We have created a software package QMANIFOLD [113] that allows simulating in various parametrized bases. It interfaces with QNET [115] and can automate various tedious calculations related to deriving the adjoint group representation and thus the derivation of the right generators $F_j^{>}(\theta_t)$. Thus, given a particular transformation U_θ (cf Section 6.7.1) and complexity functional it can compute the minimum complexity state $|\phi_t\rangle$ and coordinates θ_t for any input state $|\psi_t\rangle$. Furthermore, given a dynamical open system model in form of an $(\mathbf{S}, \mathbf{L}, H)$ object, it can carry out stochastic complex quantum diffusion simulations in the optimal basis or using the fiducial state dynamics coupled to the gradient variables.

In Figure 6.10 we have simulated the gradient coupled fiducial state dynamics for the previously introduced DOPO system above threshold for different values of the coupling gain η. All stochastic simulations were carried out with the same random seed and thus the same realization of the innovation process. Figure 6.10 (c) shows some discrepancies between states reconstructed from simulations with lower η and $\eta_0 = 200$. Specifically we compute $d^2_{\mathrm{FS}}(\psi^{(\eta)}, \psi^{(\eta_0)})$, where the fixed basis representation states are obtained from the respective moving bases representation states via $\left|\psi^{(\eta)}(t)\right\rangle = U_{\theta^{(\eta)}} \left|\phi^{(\eta)}(t)\right\rangle$ and the Fubini-Study distance is defined as:

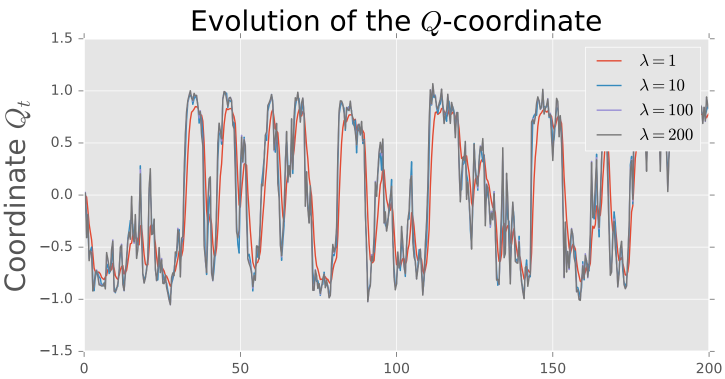
$$d_{\mathrm{FS}}(\phi, \phi') := \arccos\left(|\langle \phi|\phi'\rangle|\right) \tag{6.35}$$

The errors decrease with increasing η suggesting that a strong gradient coupling gain $\eta = O(100)$ is preferable. We intend to investigate this more systematically in [114]. We note that the errors appear to stay constant over time, which is encouraging.

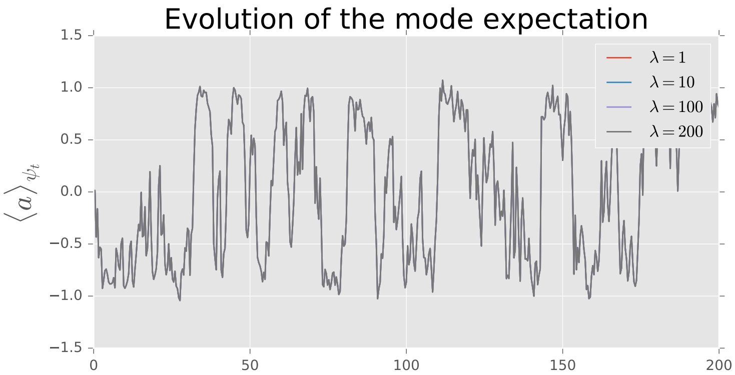
Furthermore, we have simulated our system with different sizes of the basis (cf Figure 6.11) and evaluated the error relative to the most accurate simulation. Surprisingly, we find that the truncation error remains roughly constant, i.e., even for this randomly switching system, the low-dimensional approximations to the system state track the actual system state very well.

6.5. Examples

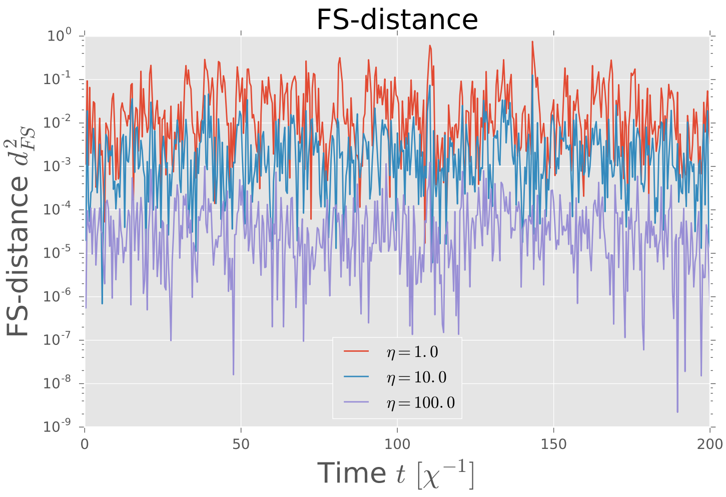
Here we present some examples of transformation groups, counting operators and the corresponding generators and sensitivity variables. A given transformation is characterized by



(a) Manifold coordinate $Q(t)$



(b) Mode expectation $\langle a \rangle_{\psi_t}$



(c) Error distance d^2_{FS}

Figure 6.10.: Gradient coupled fiducial state dynamics (simulated with $K = 30$ basis levels) for varying coupling strength $\eta \in [1, 10, 100, 200]$. In (a) we compare the Q coordinate trajectories and find that they agree very well for large η and appear somewhat 'low-passed' for $\eta = 1$. The mode expectation values (shown in (b)) agree very well for all values of η. Finally, in (c) we present the Fubini-Study distance between each trajectory and the $\eta = 200$ trajectory.

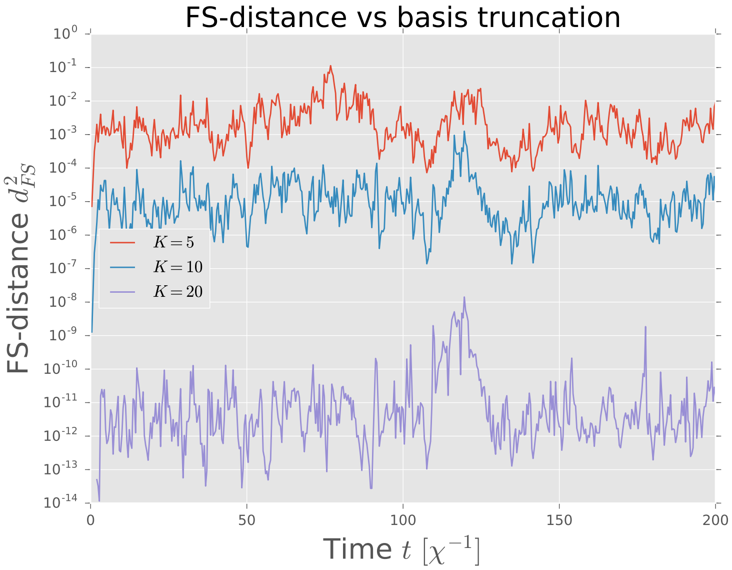


Figure 6.11.: Error due to basis truncation in the moving basis. We compare the Fubini-Study distance between a trajectory simulated with $K = 40$ moving basis levels to trajectories carried out with $K = 5, 10, 20$. Unsurprisingly, a larger basis decreases the error, but the errors do not appear to accumulate over time. All simulations were carried out with $\eta = 10$.

its differential form, its adjoint action onto its Lie-Algebra and the sensitivity operators whose expectations form the gradient and Hessian of a particular excititation minimization problem.

6.5.1. Coherent displacement

The simplest example is that of coherent displacements. Working in a real representation we have

$$U_\theta = \exp(-i\theta^1 p + i\theta^2 q), \qquad U_\theta^\dagger a U_\theta = a + \frac{\theta^1 + i\theta^2}{\sqrt{2}} \tag{6.36}$$

$$\text{where } q = \frac{a + a^\dagger}{\sqrt{2}}, \qquad p = \frac{a - a^\dagger}{\sqrt{2}i}, \tag{6.37}$$

$$F_1^>(\theta_t) = p + \theta^2/2, \qquad F_2^>(\theta_t) = -q - \theta^1/2. \tag{6.38}$$

For the canonical complexity functional $N = a^\dagger a$ the sensitivity operators are

$$Y_1^>(\theta_t) = -i[F_1^>(\theta_t), N] = -q, \qquad Y_2^>(\theta_t) = -i[F_2^>(\theta_t), N] = p \tag{6.39}$$

$$M_{11}^>(\theta_t) = [F_1^>(\theta_t), [N, F_1^>(\theta_t)]] = 1 \qquad M_{12}^>(\theta_t) = [F_1^>(\theta_t), [N, F_2^>(\theta_t)]] = 0 \tag{6.40}$$

$$M_{21}^>(\theta_t) = [F_2^>(\theta_t), [N, F_1^>(\theta_t)]] = 0 \qquad M_{22}^>(\theta_t) = [F_2^>(\theta_t), [N, F_2^>(\theta_t)]] = 1. \tag{6.41}$$

Thus we see that the Hessian is constant and positive definite, ensuring that we always have a unique optimum. In fact, the fiducial state equations of motion will be exactly those of a classical dissipative oscillator.

6.5.2. Squeezing and displacement

Our next example will be a mixture of squeezing and displacing albeit both only parameterized by a single variable, which does not allow to realize the most general pure Gaussian state.

$$U_\theta = \exp(-i\theta^1 p)\exp(i\theta^2 s), \qquad U_\theta^\dagger a U_\theta = \theta^1/\sqrt{2} + \cosh\theta^2 a - \sinh\theta^2 a^\dagger \tag{6.42}$$

$$\text{where } p = \frac{a - a^\dagger}{\sqrt{2}i}, \qquad s = \frac{a^2 - a^{\dagger 2}}{2i}, \tag{6.43}$$

$$F_1^>(\theta_t) = e^{\theta^2} q, \qquad F_2^>(\theta_t) = -s. \tag{6.44}$$

For the canonical complexity functional $N = a^\dagger a$ the sensitivity operators are

$$Y_1^>(\theta_t) = -i[F_1^>(\theta_t), N] = -e^{\theta^2} q, \qquad Y_2^>(\theta_t) = -i[F_2^>(\theta_t), N] = 2r = a^2 + a^{\dagger 2} \tag{6.45}$$

$$M_{11}^>(\theta_t) = e^{2\theta^2} \qquad M_{12}^>(\theta_t) = -e^{\theta^2} q \tag{6.46}$$

$$M_{21}^>(\theta_t) = -e^{\theta^2} q \qquad M_{22}^>(\theta_t) = 2 + 4a^\dagger a. \tag{6.47}$$

This still leads to a positive definite Hessian, but it is now dependent on the coordinates and the state.

6.5.3. Spin coherent states

Our final example here is for a different degree of freedom, namely spin coherent states as introduced by Radcliffe [94]. We work in a single irreducible representation labeled by J such that $\mathbf{J}^2 = J(J+1)$ and the spectrum of J_z is given by $-J, -J+1, \ldots, J-1, J$. The commutator relationships are $[J_z, J_\pm] = \pm J_\pm$ and $[J_+, J_-] = 2J_z$.

$$U_\theta = e^{-\mu J_+} e^{-\log(1+|\mu|^2)J_z} e^{\mu^* J_-}, \tag{6.48}$$

$$U_\theta^\dagger J_- U_\theta = \frac{J_- - \mu^2 J_+ - 2\mu J_z}{1+|\mu|^2}, \qquad U_\theta^\dagger J_z U_\theta = \frac{\mu^* J_- + \mu J_+ + (1-|\mu|^2)J_z}{1+|\mu|^2} \tag{6.49}$$

$$\text{where } \mu = \theta^1 + i\theta^2 \tag{6.50}$$

$$F_1^>(\theta_t) = \frac{-iJ_- + iJ_+ - 2\theta^2 J_z}{1+|\mu|^2} \qquad F_2^>(\theta_t) = \frac{-J_- - J_+ + 2\theta^1 J_z}{1+|\mu|^2}. \tag{6.51}$$

For the canonical spin complexity functional $N = J_z$ the sensitivity operators are

$$Y_1^{>}(\theta_t) = \frac{-J_- - J_+}{1+|\mu|^2}, \qquad Y_2^{>}(\theta_t) = \frac{iJ_- - iJ_+}{1+|\mu|^2} \tag{6.52}$$

$$M_{11}^{>}(\theta_t) = \frac{2\mu J_- + 2\mu^* J_+ - 4J_z}{1+|\mu|^2} \qquad M_{12}^{>}(\theta_t) = \frac{-2i\mu J_- + 2i\mu^* J_+}{1+|\mu|^2} \tag{6.53}$$

$$M_{21}^{>}(\theta_t) = \frac{-2i\mu J_- + 2i\mu^* J_+}{1+|\mu|^2} \qquad M_{22}^{>}(\theta_t) = \frac{-2\mu J_- - 2\mu^* J_+ - 4J_z}{1+|\mu|^2}. \tag{6.54}$$

This does not generally lead to a positive definite Hessian, e.g., for integral J consider the $j_z = 0$ eigenstate of J_z for which the Hessian vanishes.

6.6. Conclusion and Outlook

So far we have only discussed single degree of freedom transformations, but it is possible to apply the formalism to collective transformations that transform to a supermode or quasi-particle picture. There exist a rich variety of semi-simple Lie groups [37] for which these methods could work. For high-dimensional group manifolds is is likely advisable to forego any symbolic solution to the differential transformation and work in a purely numerical representation of the group's adjoint representation. This can lead to considerable complexity (collective mode transformations typically have $O(N^2)$ dimensional Lie-algebras) but it will usually still be polynomial in the number of degrees of freedom N. A drawback of collective transformations is that they can blow up the number of terms appearing in the Hamilton operator (and the collapse operators), making the problem less sparse than typical problems often are. In that case we are literally trading off between overall state space dimension and the density of our numerical representation matrices. It may also be possible to extend the formalism to quantum fields, though this can only make sense when the underlying physical theory leads to localization of those fields in their generalized configuration space.

6.7. Useful additional material

6.7.1. General construction of the coordinate transformation

In this section we outline how to construct complex transformations and derive the right generators. The simplest construction for the unitary transformation is by chaining single parameter transformations

$$U_\eta := V_{\eta^1}^{(1)} V_{\eta^2}^{(2)} \cdots V_{\eta^n}^{(n)}, \tag{6.55}$$

$$\text{where } V_{\eta^j}^{(j)} := \exp(-\eta^j X_j),\ j = 1, 2, \ldots, n \tag{6.56}$$

Thus far we allow for complex coordinates $\{\eta^j\}$ and arbitrary, i.e., not necessarily hermitian, generators $\{X_j\}$. We assume that the generators X_j are elements of a finite dimensional Lie algebra $\mathfrak{g}$. Given a basis $\{Y_1, Y_2, \ldots, Y_q\} \subset \mathfrak{g}$ for the Lie algebra with structure constants $c^l_{jk} \in \mathbb{R}$ implicitly defined via

$$[Y_j, Y_k] = \sum_{l=1}^{q} c^l_{jk} Y_l, \tag{6.57}$$

we represent each transformation generator in this basis as $X_j = \sum_{k=1}^{q} R^k_j Y_k$. Using the structure constants, it is straightforward to compute the conjugation of any basis element by a single parameter transformation [93] to be

$$V^{(j)-1}_{\eta^j} Y_k V^{(j)}_{\eta^j} = \sum_{l=1}^{q} \underbrace{[\exp(A^{(j)}_{\theta^j})]^l_k}_{=:S^{(j)}_{\eta^j}} Y_l \tag{6.58}$$

$$\text{with } [A^{(j)}_{\eta^j}]^l_k = \eta^j \sum_{h=1}^{q} R^h_j c^l_{hk}. \tag{6.59}$$

The matrices $A^{(j)}_{\eta^j}$ are typically very sparse and can be exponentiated symbolically using a tool such as Mathematica [139] the SymPy package [107]. The resulting matrices $S^{(j)}_{\eta^j} \in \mathbb{R}^{n\times n}$ are elements of the adjoint representation of the transformation group and can be used to directly transform the generators. As they each depend on only a single coordinate, they satisfy $S^{(j)}_{-\eta^j} = \left(S^{(j)}_{\eta^j}\right)^{-1}$. With this, it is straightforward to see that

$$dU_\eta = U_\eta \sum_{j=1}^{n} X^{>}_j(\eta) d\eta^j \tag{6.60}$$

$$\text{with } X^{>}_j(\eta) := V^{(n)-1}_{\eta^n} \cdots V^{(j+1)-1}_{\eta^{j+1}} X_j V^{(j+1)}_{\eta^{j+1}} \cdots V^{(n)}_{\eta^n} \tag{6.61}$$

$$= \sum_{l=1}^{n} \sum_{k=1}^{q} R^l_j [S^{(j+1)}_{\eta^{j+1}} \cdots S^{(n)}_{\eta^n}]^k_l Y_k. \tag{6.62}$$

We see that by requiring the differential form of U_η with all generators on the right hand side, each generator needs to be additionally transformed $X_j \to X^{>}_j(\eta)$ by all single parameter transformations that appear to its right.

The differential transformation can be equivalently expressed with the differential generators on the left side of U_{θ_t}:

$$dU_\eta = \left[\sum_{j=1}^{n} X_j^{<}(\eta)d\eta^j\right] U_\eta \tag{6.63}$$

$$\begin{aligned} \text{with } X_j^{<}(\eta) &= U_\eta X_j^{>}(\eta)U_\eta^{-1} && (6.64)\\ &= V_{\eta^1}^{(1)}\cdots V_{\eta^{j-1}}^{(j-1)} X_j V_{\eta^{j-1}}^{(j-1)-1}\cdots V_{\eta^1}^{(1)-1} && (6.65)\\ &= \sum_{k=1}^{n}[S_{-\eta^{j-1}}^{(j-1)}\cdots S_{-\eta^1}^{(1)}]_j^k X_k. && (6.66)\end{aligned}$$

To transform any operator M from the basis associated with $|\psi_t\rangle$ to the moving basis $|\phi_t\rangle$ one first needs to express it exclusively in terms of functions of the generators $M = f(F_1, F_2, \ldots, F_n)$ where $f : \mathbb{C}^n \to \mathbb{C}$ is analytic in all variables, typically f is a polynomial. We then have

$$\begin{aligned} M'_\theta &:= U_\theta^\dagger M U_\theta \\ &= f(F'_1(\theta), F'_2(\theta), \ldots, F'_n(\theta)), && (6.67)\\ \text{with } F'_j(\theta) &:= \sum_k [\underbrace{S_{\theta^1}^{(1)}\cdots S_{\theta^n}^{(n)}}_{=:S_\theta}]_j^k F_k. && (6.68)\end{aligned}$$

We see that having the adjoint representation single parameter transformation matrices $\{S_{\theta^j}^{(j)}, j = 1, 2, \ldots, n\}$ allows us to do all necessary computations. We remark that representing $M = f(F_1, F_2, \ldots, F_n)$ generally does not yield a unique function f as some generators may themselves be polynomials of the other generators.

Finally, note that there exist alternate ways [133] of parametrizing groups and deriving partial derivatives that may come in useful in more complex cases. For very complex parametrizations, analytical/symbolic methods may fail but in that case it should still be possible to work in a purely numerical representation that stores and integrates both θ and elements of the adjoint representation of U_θ.

6.7.2. Properties of the quantum correlation

The quantum correlation is not a strictly positive definite inner product because the quantum self-correlation of an operator A vanishes in any eigenstate

$$A\,|\phi_t\rangle = \lambda_a\,|\phi_t\rangle \Leftrightarrow \sigma\left(A, A\right)_{\phi_t} = 0. \tag{6.69}$$

The sufficiency "⇒" of this condition is obvious, the necessity "⇐" follows from the Cauchy-Schwarz inequality for the regular Hilbert space inner product.

Restricted to hermitian operators $A^\dagger = A, B^\dagger = B$ the quantum correlation can be decomposed into its real and imaginary part as

$$\sigma\left(A,B\right)_{\phi_t} = \operatorname{cov}\left(A,B\right)_{\phi_t} + i\left\langle \frac{[A,B]}{2i}\right\rangle_{\phi_t},$$

with the symmetrized covariance function

$$\operatorname{cov}\left(A,B\right)_{\phi_t} = \frac{1}{2}\left\langle\{A,B\}\right\rangle_{\phi_t} - \left\langle A\right\rangle_{\phi_t}\left\langle B\right\rangle_{\phi_t}.$$

7. Conclusion and Outlook

In this thesis I hope to have shown you a number of tools and projects that enable the computational modeling of quantum feedback networks as well as provide some basic intuition for how useful computational operations could be mapped to (quantum-) photonic hardware.

The thesis summarizes most of the work done during the past 5.5 years though there are also important gaps, ‘failed projects’ or sometimes projects that I was not ready for yet or that the time was not right for yet. In hindsight, however, I was often surprised at how some incomplete earlier work could become very useful later on.

Unfortunately there was not enough time for me to properly describe the various software packages ([115, 112, 113] and more) I developed during my PhD, but they mostly contain documentation and examples. Despite the availability of great open source software such as QuTiP [56] I there is a never-ending need for novel computational research tools and I hope to keep contributing to this field beyond my PhD. It is an exciting time for being a computational researcher as novel languages [7] and frameworks (e.g., [6]) allow programming at an increasingly high level of mathematical abstraction.

If there is a single common underlying thread to my PhD research it is that of attempting to generate tools and intuition for myself and hopefully others to design and model useful quantum systems that may be helped by – but perhaps not totally rely on – some magic material or engineering breakthrough. I believe that very interesting devices ought to be possible with current or very near future capabilities.

Very exciting work is being done on quantum annealers and analog quantum simulators but the question of what – if any – enhancement to the computational capacity is intrinsic to just weakly quantum coherent systems is – I would argue – wide open. Although there is research to analyze what kinds of quantum operations or what quantum communication capacity a given physical process may allow for, these issues still appear quite targeted at conventional unitary gate based quantum computation or ‘mere’ secure communication.

It appears that there is a great potential for insight into how the high intrinsic dimensionality of quantum dynamics could enable even classical signal processing at a very high physical hardware efficiency and I therefore believe that alternative computational paradigms to that of gate based quantum computing need to be explored more.

A. Simplification rules for the circuit algebra

A.1. Simplifying algebraic circuit expressions

A.1.1. Basic algebraic properties

By observing that we can define for a general system $Q = (\mathbf{S}, \mathbf{L}, H)$ its *series inverse* system $Q^{\lhd -1} := (\mathbf{S}^\dagger, -\mathbf{S}^\dagger \mathbf{L}, -H)$

$$(\mathbf{S}, \mathbf{L}, H) \lhd (\mathbf{S}^\dagger, -\mathbf{S}^\dagger \mathbf{L}, -H) = (\mathbf{S}^\dagger, -\mathbf{S}^\dagger \mathbf{L}, -H) \lhd (\mathbf{S}, \mathbf{L}, H) = (\mathbb{1}_n, 0, 0) =: \mathbf{id}n, \tag{A.1}$$

we see that the series product induces a group structure on the set of n-channel circuit components for any $n \geq 1$. It can easily be verified that the series inverse of the basic operations is calculated as follows

$$(Q_1 \lhd Q_2)^{\lhd -1} = Q_2^{\lhd -1} \lhd Q_1^{\lhd -1} \tag{A.2}$$

$$(Q_1 \boxplus Q_2)^{\lhd -1} = Q_1^{\lhd -1} \boxplus Q_2^{\lhd -1} \tag{A.3}$$

$$([Q]_{k \to l})^{\lhd -1} = \left[Q^{\lhd -1}\right]_{l \to k}. \tag{A.4}$$

In the following, we denote the number of channels of any given system $Q = (\mathbf{S}, \mathbf{L}, H)$ by $\mathrm{c}dim\,(Q) := n$. The most obvious expression simplification is the associative expansion of concatenations and series:

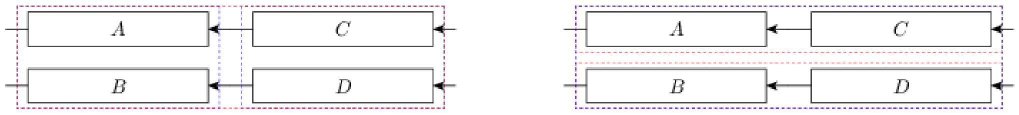
$$(A_1 \lhd A_2) \lhd (B_1 \lhd B_2) = A_1 \lhd A_2 \lhd B_1 \lhd B_2 \tag{A.5}$$

$$(C_1 \boxplus C_2) \boxplus (D_1 \boxplus D_2) = C_1 \boxplus C_2 \boxplus D_1 \boxplus D_2 \tag{A.6}$$

A further interesting property that follows intuitively from the graphical representation (cf. Fig. A.1) is the following tensor decomposition law

$$(A \boxplus B) \lhd (C \boxplus D) = (A \lhd C) \boxplus (B \lhd D), \tag{A.7}$$

which is valid for $\mathrm{c}dim\,(A) = \mathrm{c}dim\,(C)$ and $\mathrm{c}dim\,(B) = \mathrm{c}dim\,(D)$. As mentioned in the caption to Figure A.1, it will most often be preferable to use the RHS expression of Equation (A.7) as this enables us to understand the flow of optical channels more easily from



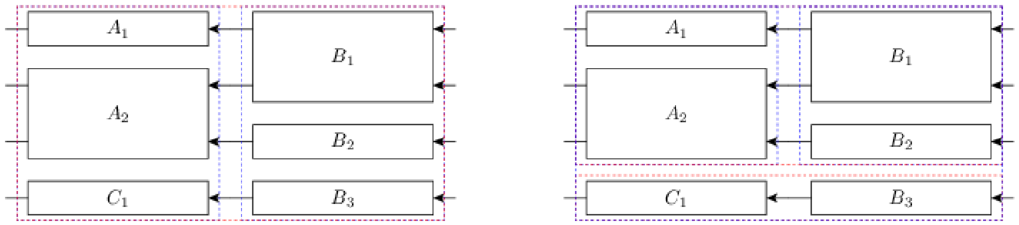
(a) $(A \boxplus B) \lhd (C \boxplus D)$ (b) $(A \lhd C) \boxplus (B \lhd D)$

Figure A.1.: Equivalent circuits, where a red box marks a series product and a blue box marks a concatenation. The second version A.1(b) has the advantage of making more explicit that the overall circuit consists of two channels without direct optical scattering.

the algebraic expression. In [41] Gough and James denote a system that can be expressed as a concatenation as *reducible*. A system that cannot be further decomposed into concatenated subsystems is accordingly called *irreducible*. As follows intuitively from a graphical representation any given complex system $Q = (\mathbf{S}, \mathbf{L}, H)$ admits a decomposition into $1 \leq N \leq cdim\,(Q)$ irreducible subsystems $Q = Q_1 \boxplus Q_2 \boxplus \cdots \boxplus Q_N$, where their channel dimensions satisfy $cdim\,(Q_j) \geq 1$, $j = 1, 2, \dots N$ and $\sum_{j=1}^{N} cdim\,(Q_j) = cdim\,(Q)$. While their individual parameter triplets themselves are not uniquely determined[1], the sequence of their channel dimensions $(cdim\,(Q_1)\,, cdim\,(Q_2)\,, \dots cdim\,(Q_N)) =: bls\,(Q)$ clearly is. We denote this tuple as the block structure of Q. We are now able to generalize the decomposition law in the following way: Given two systems of n channels with the same block structure $bls\,(A) = bls\,(B) = (n_1, ...n_N)$, there exist decompositions of A and B such that

$$A \lhd B = (A_1 \lhd B_1) \boxplus \cdots \boxplus (A_N \lhd B_N) \tag{A.8}$$

with $cdim\,(A_j) = cdim\,(B_j) = n_j$, $j = 1, \dots N$. However, even in the case that the two block structures are not equal, there may still exist non-trivial compatible block decompositions that at least allow a partial application of the decomposition law. Consider the example presented in Figure A.2.



(a) Series "$(1,2,1) \lhd (2,1,1)$" (b) Optimal decomposition into $(3,1)$

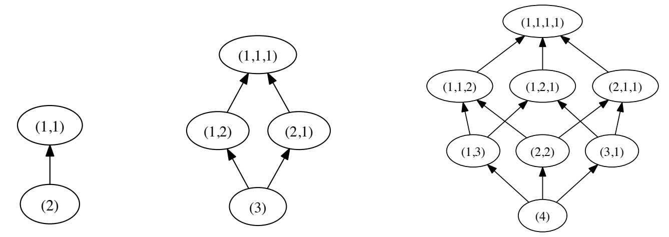
Figure A.2.: Even in the case of a series between systems with unequal block structures, there often exists a non-trivial common block decomposition that simplifies the overall expression.

---

[1] Actually the scattering matrices $\{\mathbf{S}_j\}$ and the coupling vectors $\{\mathbf{L}_j\}$ *are* uniquely determined, but the Hamiltonian parameters $\{H_j\}$ must only obey the constraint $\sum_{j=1}^{N} H_j = H$.

An irrelevant mathematical excursion

This can be proven rigorously by defining a relation $\preceq$ on the set of all possible block structures with a total number of n channels: We write $(n_1, n_2, \dots n_N) \preceq (m_1, m_2, \dots m_M)$, $\sum n_j = \sum m_k = n$ iff it is possible to group the m_j: $((m_1, \dots m_{k_1-1}), (m_{k_1}, \dots m_{k_2-1}), \dots (m_{k_N}, \dots m_M))$, such that $\sum_{k=k_l}^{k_{l+1}-1} m_k = n_l$, where $k_1 := 1$ and $k_{N+1} := M+1$. It can then be shown that this relation induces a partial order (i.e., it is transitive, reflexive and anti-symmetric) on the set of all block structures of n channels. Since there exist a minimal element (n) and a maximal element $(\overbrace{1, 1, \dots, 1}^{n})$ it also follows from this that for any two block structures $(q_1, \dots q_{N_1})$ and $(p_1, \dots p_{N_2})$ there exists a unique *maximal common block structure* $(n_1, \dots n_N) \preceq (q_1, \dots q_{N_1}), (p_1, \dots p_{N_2})$. Hence, given a system with block structure $(m_1, m_2, \dots m_M)$ we can also express it is as a concatenation of N subsystems with channel dimensions $n_1, n_2, \dots, n_N$ as long as $(n_1, n_2, \dots n_N) \preceq (m_1, m_2, \dots m_M)$. This implies that there always exists a maximal common block decomposition of two complex systems such that their series product can be accordingly simplified by the decomposition law. This can be visualized as in Figure A.3, where we have illustrated the relation $\preceq$ between all block structures for the case of $n = 4$. As as curious side-note, Figure A.3 seems to suggest that the graphs for arbitrary n follow a simple geometrical pattern: For n total channels the graph assumes the form of a hypercube in $n-1$ dimensions. This hypothesis is further strengthened by the coincidence of the number of different block structures for n with the number of vertices of a hypercube: $2^{(n-1)}$. However, these considerations are completely irrelevant to our original topic.



(a) $n = 2$ (b) $n = 3$ (c) $n = 4$

Figure A.3.: A directed vertex between two nodes $a \to b$ represents the relation $a \preceq b$. Note that we have omitted all edges that would follow from the transitivity of the relation, e.g. $(4) \to (1,1,2)$ is implied by the (alternative) paths via the intermediate nodes $(1,3)$ or $(2,2)$.

Even without the above considerations it is straightforward to create a recursive algorithm to find the maximal common block structure for two given, differing block structures. A sample implementation in Python can be found in Listing A.1.

Listing A.1: Recursively obtain the maximal common block structure

```python
def common_block_structure ( bs1 , bs2 ):
   if len ( bs1 ) == len ( bs2 ) == 0:
      return () # empty tuple
   i = j = 1
   # the strategy is to find the minimal first i elements of bs1
   # and j elements of bs2 such that their sums are equal
   while( True ):
      lsum = sum ( bs1 [: i ]) # this is sort of inefficient , but more pedagogical
      rsum = sum ( bs2 [: j ]) # and these sums are usually very small
      if( lsum < rsum ):
         i +=1
      elif ( rsum < lsum ):
         j += 1
      else: # lsum == rsum
         break
   # upon success the first element of the common block structure
   # is given by that equal sum , proceed to calculate the rest recursively
   return ( lsum , ) + common_block_structure ( bs1 [i :] , bs2 [j :])
```

A.1.2. Permutation objects

The algebraic representation of complex circuits often requires systems that only permute channels without actual scattering. The group of permutation matrices is simply a subgroup of the unitary (operator) matrices. For any permutation matrix $\mathbf{P}$, the system described by $(\mathbf{P}, \mathbf{0}, 0)$ represents a pure permutation of the optical fields (cf. Fig. A.4). A permutation σ of n elements ($\sigma \in \Sigma_n$) is often represented in the following form $\begin{pmatrix} 1 & 2 & \dots & n \\ \sigma(1) & \sigma(2) & \dots & \sigma(n) \end{pmatrix}$, but obviously it is also sufficient to specify the tuple of images $(\sigma(1), \sigma(2), \dots, \sigma(n))$. We

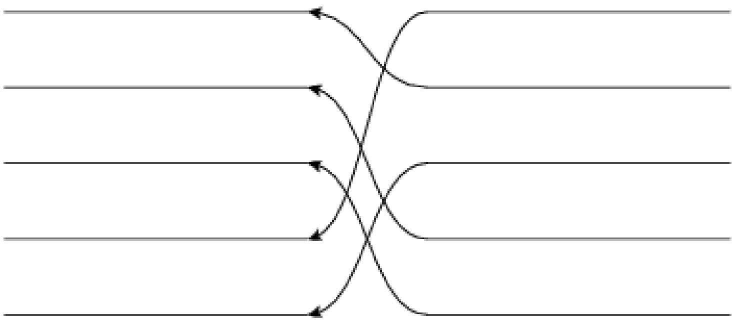


Figure A.4.: A graphical representation of $\mathbf{P}_\sigma$ where $\sigma \equiv (4, 1, 5, 2, 3)$ in image tuple notation.

now define the permutation matrix via its matrix elements

$$(\mathbf{P}_\sigma)_{kl} = \delta_{k\sigma(l)} = \delta_{\sigma^{-1}(k)l}. \tag{A.9}$$

Such a matrix then maps the j-th unit vector onto the $\sigma(j)$-th unit vector or equivalently the j-th incoming optical channel is mapped to the $\sigma(j)$-th outgoing channel. In contrast to a definition often found in mathematical literature this definition ensures that the representation matrix for a composition of permutations $\sigma_2 \circ \sigma_1$ results from a product of the individual representation matrices in the same order $\mathbf{P}_{\sigma_2\circ\sigma_1} = \mathbf{P}_{\sigma_2}\mathbf{P}_{\sigma_1}$. This can be shown directly on the order of the matrix elements

$$(\mathbf{P}_{\sigma_2\circ\sigma_1})_{kl} = \delta_{k(\sigma_2\circ\sigma_1)(l)} = \sum_j \delta_{kj}\delta_{j(\sigma_2\circ\sigma_1)(l)} = \sum_j \delta_{k\sigma_2(j)}\delta_{\sigma_2(j)(\sigma_2\circ\sigma_1)(l)} \tag{A.10}$$

$$= \sum_j \delta_{k\sigma_2(j)}\delta_{\sigma_2(j)\sigma_2(\sigma_1(l))} = \sum_j \delta_{k\sigma_2(j)}\delta_{j\sigma_1(l)} = \sum_j (\mathbf{P}_{\sigma_2})_{kj}(\mathbf{P}_{\sigma_1})_{jl}, \tag{A.11}$$

where the third equality corresponds simply to a reordering of the summands and the fifth equality follows from the bijectivity of σ_2. In the following we will often write P_σ as a shorthand for $(\mathbf{P}_\sigma, \mathbf{0}, 0)$. Thus, our definition ensures that we may simplify any series of permutation systems in the most intuitive way: $P_{\sigma_2} \lhd P_{\sigma_1} = P_{\sigma_2\circ\sigma_1}$. Obviously the set of permutation systems of n channels and the series product are a subgroup of the full system series group of n channels. Specifically, it includes the identity $\mathbf{id}n = P_{\sigma_{\mathrm{id}_n}}$.

From the orthogonality of the representation matrices it directly follows that $\mathbf{P}_\sigma^T = \mathbf{P}_{\sigma^{-1}}$ For future use we also define a concatenation between permutations

$$\sigma_1 \boxplus \sigma_2 := \begin{pmatrix} 1 & 2 & \dots & n & n+1 & n+2 & \dots & n+m \\ \sigma_1(1) & \sigma_1(2) & \dots & \sigma_1(n) & n+\sigma_2(1) & n+\sigma_2(2) & \dots & n+\sigma_2(m) \end{pmatrix}, \tag{A.12}$$

which satisfies $P_{\sigma_1} \boxplus P_{\sigma_2} = P_{\sigma_1 \boxplus \sigma_2}$ by definition. Another helpful definition is to introduce a special set of permutations that map specific ports into each other but leave the relative order of all other ports intact:

$$\omega_{l\leftarrow k}^{(n)} := \begin{cases} \begin{pmatrix} 1 & \dots & k-1 & k & k+1 & \dots & l-1 & l & l+1 & \dots & n \\ 1 & \dots & k-1 & l & k & \dots & l-2 & l-1 & l+1 & \dots & n \end{pmatrix} & \text{for } k < l \\ \begin{pmatrix} 1 & \dots & l-1 & l & l+1 & \dots & k-1 & k & k+1 & \dots & n \\ 1 & \dots & l-1 & l+1 & l+2 & \dots & k & l & k+1 & \dots & n \end{pmatrix} & \text{for } k > l \end{cases} \tag{A.13}$$

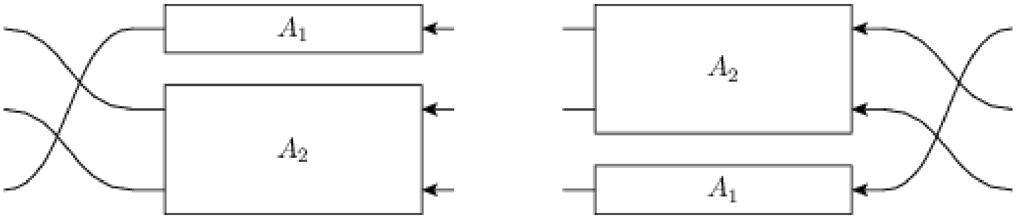
We define the corresponding system objects as $W_{l\leftarrow k}^{(n)} := P_{\omega_{l\leftarrow k}^{(n)}}$.

A.1.3. Permutations and Concatenations

Given a series $P_\sigma \lhd (Q_1 \boxplus Q_2 \boxplus \cdots \boxplus Q_N)$ where the Q_j are irreducible systems, we analyze in which cases it is possible to (partially) "move the permutation through" the concatenated expression. Obviously we could just as well investigate the opposite scenario $(Q_1 \boxplus Q_2 \boxplus \cdots \boxplus Q_N) \lhd P_\sigma$, but this second scenario is closely related[2].

Block-permuting permutations

The simples case is realized when the permutation simply permutes whole blocks intactly (cf. Fig. A.5) Given a block structure $\mathbf{n} := (n_1, n_2, \dots n_N)$ a permutation $\sigma \in \Sigma_n$ is said to *block permute* $\mathbf{n}$ iff there exists a permutation $\tilde{\sigma} \in \Sigma_N$ such that



(a) $P_\sigma \lhd (A_1 \boxplus A_2)$ (b) $(A_2 \boxplus A_1) \lhd P_\sigma$

Figure A.5.: A block permuting series.

$$\begin{aligned} P_\sigma \lhd (Q_1 \boxplus Q_2 \boxplus \cdots \boxplus Q_N) &= (P_\sigma \lhd (Q_1 \boxplus Q_2 \boxplus \cdots \boxplus Q_N) \lhd P_{\sigma^{-1}}) \lhd P_\sigma \\ &= (Q_{\tilde{\sigma}(1)} \boxplus Q_{\tilde{\sigma}(2)} \boxplus \cdots \boxplus Q_{\tilde{\sigma}(N)}) \lhd P_\sigma \end{aligned}$$

Hence, the permutation σ, given in image tuple notation, block permutes $\mathbf{n}$ iff for all $1 \leq j \leq N$ and for all $0 \leq k < n_j$ we have $\sigma(o_j + k) = \sigma(o_j) + k$, where we have introduced the block offsets $o_j := 1 + \sum_{j' < j} n_j$. When these conditions are satisfied, $\tilde{\sigma}$ may be obtained by demanding that $\tilde{\sigma}(a) > \tilde{\sigma}(b) \Leftrightarrow \sigma(o_a) > \sigma(o_b)$. This equivalence reduces the computation of $\tilde{\sigma}$ to sorting a list in a specific way.

Block-factorizing permutations

The next-to-simplest case is realized when a permutation σ can be decomposed $\sigma = \sigma_{\mathrm{b}} \circ \sigma_{\mathrm{i}}$ into a permutation σ_{b} that block permutes the block structure $\mathbf{n}$ and an internal permutation σ_{i} that only permutes within each block, i.e. $\sigma_{\mathrm{i}} = \sigma_1 \boxplus \sigma_2 \boxplus \cdots \boxplus \sigma_N$. In this case we can perform the following simplifications

$$P_\sigma \lhd (Q_1 \boxplus Q_2 \boxplus \cdots \boxplus Q_N) = P_{\sigma_b} \lhd [(P_{\sigma_1} \lhd Q_1) \boxplus (P_{\sigma_2} \lhd Q_2) \boxplus \cdots \boxplus (P_{\sigma_N} \lhd Q_N)]. \tag{A.14}$$

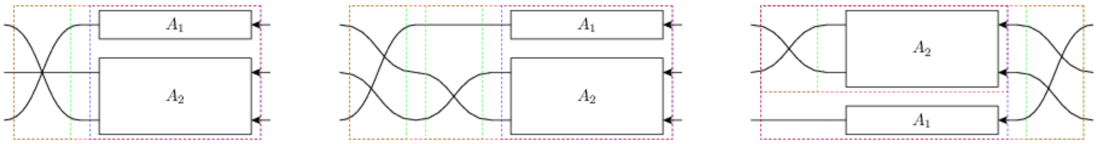
We see that we have reduced the problem to the above discussed case. The result is now

$$P_\sigma \lhd (Q_1 \boxplus \cdots \boxplus Q_N) = \left[(P_{\sigma_{\tilde{\sigma_{\mathrm{b}}}(1)}} \lhd Q_{\tilde{\sigma_{\mathrm{b}}}(1)}) \boxplus \cdots \boxplus (P_{\sigma_{\tilde{\sigma_{\mathrm{b}}}(N)}} \lhd Q_{\tilde{\sigma_{\mathrm{b}}}(N)})\right] \lhd P_{\sigma_{\mathrm{b}}}. \tag{A.15}$$

---

[2] Series-Inverting a series product expression also results in an inverted order of the operand inverses $(Q_1 \lhd Q_2)^{\lhd -1} = Q_2^{\lhd -1} \lhd Q_1^{\lhd -1}$. Since the inverse of a permutation (concatenation) is again a permutation (concatenation), the cases are in a way "dual" to each other.

In this case we say that σ *block factorizes* according to the block structure $\mathbf{n}$. Figure illustrates an example of this case. A permutation σ block factorizes according to the block structure $\mathbf{n}$ iff for all $1 \leq j \leq N$ we have $\max_{0\leq k<n_j} \sigma(o_j+k) - \min_{0\leq k'<n_j} \sigma(o_j+k') = n_j - 1$, with the block offsets defined as above. In other words, the image of a single block is coherent in the sense that no other numbers from outside the block are mapped into the integer range spanned by the minimal and maximal points in the block's image. The equivalence follows from our previous result and the bijectivity of σ.



(a) $P_\sigma \lhd (A_1 \boxplus A_2)$ (b) $P_{\sigma_b} \lhd P_{\sigma_i} \lhd (A_1 \boxplus A_2)$ (c) $((P_{\sigma_2} \lhd A_2) \boxplus A_1) \lhd P_{\sigma_\mathrm{b}}$

Figure A.6.: A block factorizable series.

The general case

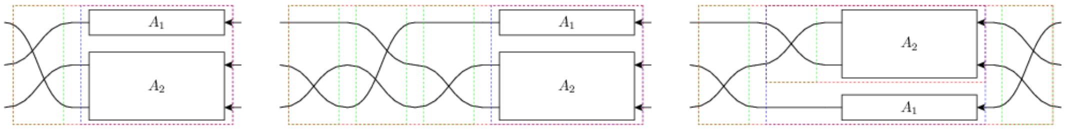
In general there exists no unique way how to split apart the action of a permutation on a block structure. However, it is possible to define a some rules that allow us to "move as much of the permutation" as possible to the RHS of the series. This involves the factorization $\sigma = \sigma_\mathrm{x} \circ \sigma_\mathrm{b} \circ \sigma_\mathrm{i}$ defining a specific way of constructing both σ_b and σ_i from σ. The remainder σ_x can then be calculated through

$$\sigma_\mathrm{x} := \sigma \circ \sigma_\mathrm{i}^{-1} \circ \sigma_\mathrm{b}^{-1}. \tag{A.16}$$

Hence, by construction, $\sigma_\mathrm{b} \circ \sigma_\mathrm{i}$ factorizes according to $\mathbf{n}$ so only σ_x remains on the exterior LHS of the expression.

So what then are the rules according to which we construct the block permuting σ_b and the decomposable σ_i? We wish to define σ_i such that the remainder $\sigma \circ \sigma_\mathrm{i}^{-1} = \sigma_\mathrm{x} \circ \sigma_\mathrm{b}$ does not cross any two signals that are emitted from the same block. Since by construction σ_b only permutes full blocks anyway this means that σ_x also does not cross any two signals emitted from the same block. This completely determines σ_i and we can therefore calculate $\sigma \circ \sigma_\mathrm{i}^{-1} = \sigma_\mathrm{x} \circ \sigma_\mathrm{b}$ as well. To construct σ_b it is sufficient to define an total order relation on the blocks that only depends on the block structure $\mathbf{n}$ and on $\sigma \circ \sigma_\mathrm{i}^{-1}$. We define the order on the blocks such that they are ordered according to their minimal image point under σ. Since $\sigma \circ \sigma_\mathrm{i}^{-1}$ does not let any block-internal lines cross, we can thus order the blocks according to the order of the images of the first signal $\sigma \circ \sigma_\mathrm{i}^{-1}(o_j)$. In Figure A.7 we have illustrated this with an example.

Finally, it is a whole different question, why we would want move part of a permutation through the concatenated expression in this first place as the expressions usually appear to



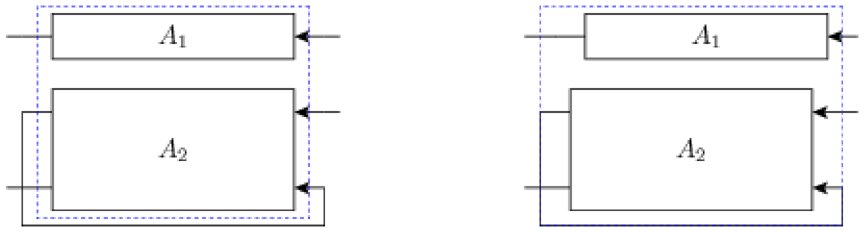
(a) $P_\sigma \lhd (A_1 \boxplus A_2)$ (b) $P_{\sigma_{\rm x}} \lhd P_{\sigma_{\rm b}} \lhd P_{\sigma_{\rm i}} \lhd (A_1 \boxplus A_2)$ (c) $(P_{\sigma_{\rm x}} \lhd (P_{\sigma_2} \lhd A_2) \boxplus A_1) \lhd P_{\sigma_{\rm b}}$

Figure A.7.: A general series with a non-factorizable permutation. In the intermediate step A.7(b) we have explicitly separated $\sigma = \sigma_{\rm x} \circ \sigma_{\rm b} \circ \sigma_{\rm i}$.

become more complicated rather than simpler. This is, because we are currently focussing only on single series products between two systems. In a realistic case we have many systems in series and among these there might be quite a few permutations. Here, it would seem advantageous to reduce the total number of permutations within the series by consolidating them where possible: $P_{\sigma_2} \lhd P_{\sigma_1} = P_{\sigma_2 \circ \sigma_1}$. To do this, however, we need to try to move the permutations through the full series and collect them on one side (in our case the RHS) where they can be combined to a single permutation. Since it is not always possible to move a permutation through a concatenation (as we have seen above), it makes sense to at some point in the simplification process reverse the direction in which we move the permutations and instead collect them on the LHS. Together these two strategies achieve a near perfect permutation simplification.

A.1.4. Feedback of a concatenation

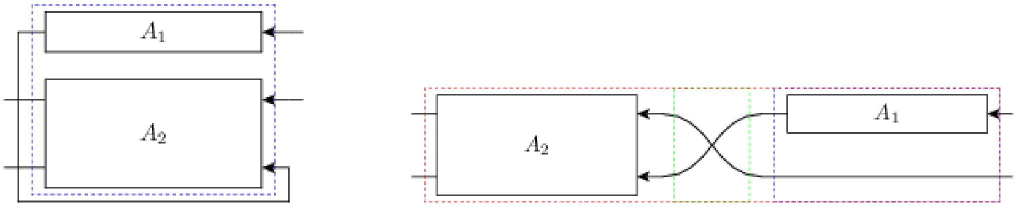
A feedback operation on a concatenation can always be simplified in one of two ways: If the outgoing and incoming feedback ports belong to the same irreducible subblock of the concatenation, then the feedback can be directly applied only to that single block. For an illustrative example see Figure A.8.



(a) $[A_1 \boxplus A_2]_{2\to3}$ (b) $A_1 \boxplus [A_2]_{1\to2}$

Figure A.8.: Reduction to feedback of subblock.

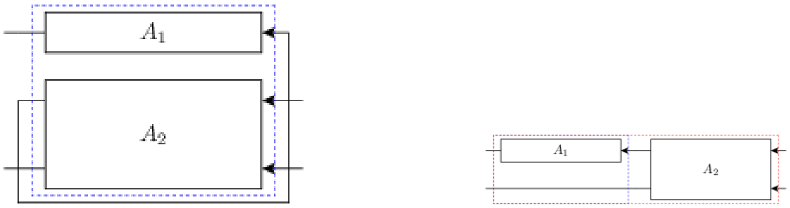
If, on the other, the outgoing feedback port is on a different subblock than the incoming, the resulting circuit actually does not contain any real feedback and we can find a way to reexpress it algebraically by means of a series product (cf. Figures A.9 and A.10).



(a) $[A_1 \boxplus A_2]_{1\to 3}$

(b) $A_2 \lhd W^{(2)}_{2\leftarrow 1} \lhd (A_2 \boxplus \mathbf{id}1)$

Figure A.9.: Reduction of feedback to series, first example



(a) $[A_1 \boxplus A_2]_{2\to 1}$

(b) $(A_1 \boxplus \mathbf{id}1) \lhd A_2$

Figure A.10.: Reduction of feedback to series, second example

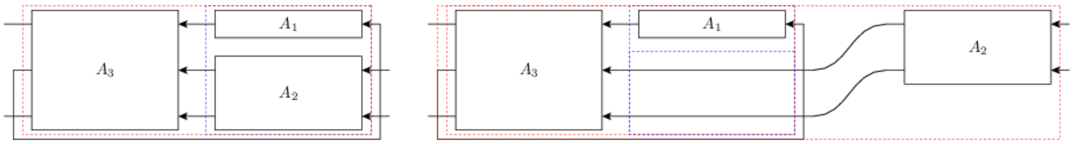
To discuss the case in full generality consider the feedback expression $[A \boxplus B]_{k\to l}$ with $\mathrm{c}\mathit{dim}\,(A) = n_A$ and $\mathrm{c}\mathit{dim}\,(B) = n_B$ and where A and B are not necessarily irreducible. There are four different cases to consider.

1. $k, l \le n_A$: In this case the simplified expression should be $[A]_{k\to l} \boxplus B$
2. $k, l > n_A$: Similarly as before but now the feedback is restricted to the second operand $A \boxplus [B]_{(k-n_A)\to(l-n_A)}$, cf. Fig. A.8.
3. $k \le n_A < l$: This corresponds to a situation that is actually a series and can be re-expressed as $(\mathbf{id}n_A - 1 \boxplus B) \lhd W^{(n)}_{(l-1)\leftarrow k} \lhd (A + \mathbf{id}n_B - 1)$, cf. Fig. A.9.
4. $l \le n_A < k$: Again, this corresponds a series but with a reversed order compared to above $(A + \mathbf{id}n_B - 1) \lhd W^{(n)}_{l\leftarrow(k-1)} \lhd (\mathbf{id}n_A - 1 \boxplus B)$, cf. Fig. A.10.

A.1.5. Feedback of a series

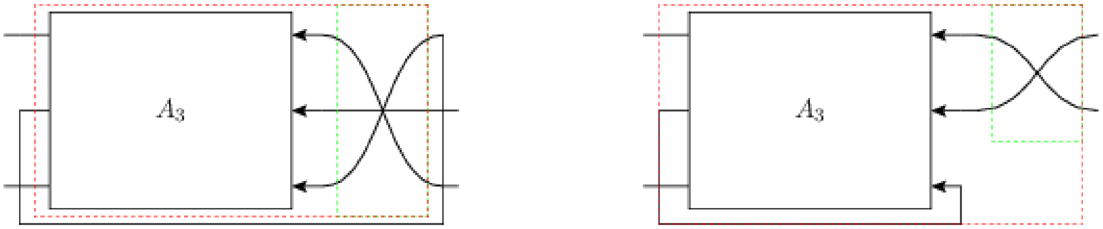
There are two important cases to consider for the kind of expression at either end of the series: A series starting or ending with a permutation system or a series starting or ending with a concatenation.

1. $[A \lhd (C \boxplus D)]_{k\to l}$: We define $n_C = \mathrm{c}\mathit{dim}\,(C)$ and $n_A = \mathrm{c}\mathit{dim}\,(A)$. Without too much loss of generality, let's assume that $l \le n_C$ (the other case is quite similar). We can then pull D out of the feedback loop: $[A \lhd (C \boxplus D)]_{k\to l} \longrightarrow [A \lhd (C \boxplus \mathbf{id}n_D)]_{k\to l} \lhd$



(a) $[A_3 \lhd (A_1 \boxplus A_2)]_{2\to 1}$ (b) $(A_3 \lhd (A_1 \boxplus \mathbf{id}2)) \lhd A_2$

Figure A.11.: Reduction of series feedback with a concatenation at the RHS.



(a) $[A_3 \lhd P_\sigma]_{2\to 1}$ (b) $[A_3]_{2\to 3} \lhd P_{\tilde\sigma}$

Figure A.12.: Reduction of series feedback with a permutation at the RHS.

$(\mathbf{id}n_C - 1 \boxplus D)$. Obviously, this operation only makes sense if $D \neq \mathbf{id}n_D$. The case $l > n_C$ is quite similar, except that we pull C out of the feedback. See Figure A.11 for an example.

2. We now consider $[(C \boxplus D) \lhd E]_{k\to l}$ and we assume $k \leq n_C$ analogous to above. Provided that $D \neq \mathbf{id}n_D$, we can pull it out of the feedback and get $(\mathbf{id}n_C - 1 \boxplus D) \lhd [(C \boxplus \mathbf{id}n_D) \lhd E]_{k\to l}$.

3. $[A \lhd P_\sigma]_{k\to l}$: The case of a permutation within a feedback loop is a lot more intuitive to understand graphically (e.g., cf. Figure A.12). Here, however we give a thorough derivation of how a permutation can be reduced to one involving one less channel and moved outside of the feedback. First, consider the equality $[A \lhd W^{(n)}_{j\leftarrow l}]_{k\to l} = [A]_{k\to j}$ which follows from the fact that $W^{(n)}_{j\leftarrow l}$ preserves the order of all incoming signals except the l-th. Now, rewrite

$$\begin{aligned} [A \lhd P_\sigma]_{k\to l} &= [A \lhd P_\sigma \lhd W^{(n)}_{l\leftarrow n} \lhd W^{(n)}_{n\leftarrow l}]_{k\to l} && \text{(A.17)} \\ &= [A \lhd P_\sigma \lhd W^{(n)}_{l\leftarrow n}]_{k\to n} && \text{(A.18)} \\ &= [A \lhd W^{(n)}_{\sigma(l)\leftarrow n} \lhd (W^{(n)}_{n\leftarrow \sigma(l)} \lhd P_\sigma \lhd W_{l\leftarrow n})]_{k\to n} && \text{(A.19)} \end{aligned}$$

Turning our attention to the bracketed expression within the feedback, we clearly see that it must be a permutation system $P_{\sigma'} = W^{(n)}_{n\leftarrow\sigma(l)} \lhd P_\sigma \lhd W^{(n)}_{l\leftarrow n}$ that maps $n \to l \to \sigma(l) \to n$. We can therefore write $\sigma' = \tilde\sigma \boxplus \sigma_{\mathrm{id}_1}$ or equivalently $P_{\sigma'} = P_{\tilde\sigma} \boxplus \mathbf{id}1$ But this means, that the series within the feedback ends with a concatenation and from our above rules we know how to handle this:

$$\begin{aligned}[A \lhd P_\sigma]_{k\to l} &= [A \lhd W^{(n)}_{\sigma(l)\leftarrow n} \lhd (P_{\tilde{\sigma}} \boxplus \mathbf{id}1)]_{k\to n} \qquad &\text{(A.20)}\\ &= [A \lhd W^{(n)}_{\sigma(l)\leftarrow n}]_{k\to n} \lhd P_{\tilde{\sigma}} \qquad &\text{(A.21)}\\ &= [A]_{k\to\sigma(l)} \lhd P_{\tilde{\sigma}}, \qquad &\text{(A.22)}\end{aligned}$$

where we know that the reduced permutation is the well-defined restriction to $n-1$ elements of $\sigma' = \left(\omega^{(n)}_{n\leftarrow\sigma l} \circ \sigma \circ \omega^{(n)}_{l\leftarrow n}\right)$.

4. The last case is analogous to the previous one and we will only state the results without a derivation:

$$[P_\sigma \lhd A]_{k\to l} = P_{\tilde{\sigma}} \lhd [A]_{\sigma^{-1}(k)\to l}, \tag{A.23}$$

where the reduced permutation is given by the (again well-defined) restriction of $\omega^{(n)}_{n\leftarrow k} \circ \sigma \circ \omega^{(n)}_{\sigma^{-1}(k)\leftarrow n}$ to $n-1$ elements.

Bibliography

[1] Abraham, E., and Smith, S. D. Optical bistability and related devices. *Reports on Progress in Physics 45* (1982).

[2] Ananthanarayanan, R., Esser, S. K., Simon, H. D., and Modha, D. S. The cat is out of the bag. In *Proceedings of the Conference on High Performance Computing Networking, Storage and Analysis - SC '09* (New York, New York, USA, 2009), no. c, ACM Press, p. 1.

[3] Ascher, D., Dubois, P. F., Hinsen, K., Hugunin, J., Jones, E., Oliphant, T., and Peterson, P. SciPy and Numpy: Open source scientific tools for python, 2001. `http://www.scipy.org/`.

[4] Barchielli, A., and Lupieri, G. Quantum stochastic calculus, operation valued stochastic processes, and continual measurements in quantum mechanics. *Journal of mathematical physics 26*, September (1985), 2222–2230.

[5] Beazley, D. PLY - python lex-yacc, 2001. `http://www.dabeaz.com/ply`.

[6] Bergstra, J., Bastien, F., Breuleux, O., Lamblin, P., Pascanu, R., Delalleau, O., Desjardins, G., Warde-Farley, D., Goodfellow, I., Bergeron, A., and Bengio, Y. Theano: Deep learning on gpus with python. *Journal of Machine Learning Research 1* (2011), 1–48.

[7] Bezanson, J., Edelman, A., Karpinski, S., and Shah, V. B. Julia: A fresh approach to numerical computing. 41.

[8] Boixo, S., Rønnow, T. F., Isakov, S. V., Wang, Z., Wecker, D., Lidar, D. a., Martinis, J. M., and Troyer, M. Evidence for quantum annealing with more than one hundred qubits. *Nature Physics 10*, 3 (2 2014), 218–224.

[9] Boixo, S., Smelyanskiy, V. N., Shabani, A., Isakov, S. V., Dykman, M., Denchev, V. S., Amin, M., Smirnov, A., Mohseni, M., and Neven, H. Computational role of collective tunneling in a quantum annealer. 1–31.

[10] Bouten, L. Filtering and control in quantum optics. *Arxiv preprint quantph0410080* (2004).

[11] Bouten, L., Handel, R. V., James, M., and Van Handel, R. An introduction to quantum filtering. *SIAM Journal on Control and Optimization 46*, 6 (2007), 2199–2241.

[12] Bouten, L., and Silberfarb, A. Adiabatic elimination in quantum stochastic models. *Communications in Mathematical Physics 283*, 2 (5 2008), 491–505.

[13] Bouten, L., van Handel, R., and Silberfarb, A. Approximation and limit theorems for quantum stochastic models with unbounded coefficients. *Journal of Functional Analysis 254*, 12 (6 2008), 3123–3147.

[14] Caldeira, AO, and Leggett, A. Quantum tunneling in a dissipative system. *Annals of Physics 149* (1983), 374–456.

[15] Carmichael, H. Quantum trajectory theory for cascaded open systems. *Physical Review Letters 70*, 15 (8 1993), 2273–2276.

[16] Carter, S. Quantum theory of nonlinear fiber optics: Phase-space representations. *Physical Review A 51*, 4 (4 1995), 3274–3301.

[17] Cassidy, A. S., Alvarez-Icaza, R., Akopyan, F., Sawada, J., Arthur, J. V., Merolla, P. A., Datta, P., Tallada, M. G., Taba, B., Andreopoulos, A., Amir, A., Esser, S. K., Kusnitz, J., Appuswamy, R., Haymes, C., Brezzo, B., Moussalli, R., Bellofatto, R., Baks, C., Mastro, M., Schleupen, K., Cox, C. E., Inoue, K., Millman, S., Imam, N., Mcquinn, E., Nakamura, Y. Y., Vo, I., Guok, C., Nguyen, D., Lekuch, S., Asaad, S., Friedman, D., Jackson, B. L., Flickner, M. D., Risk, W. P., Manohar, R., and Modha, D. S. Real-time scalable cortical computing at 46 giga-synaptic ops/watt with extasciitilde100x speedup in time-to-solution and extasciitilde100,000x reduction in energy-to-solution. In *Proceedings of the International Conference for High Performance Computing, Networking, Storage and Analysis* (Piscataway, NJ, USA, 2014), SC ’14, IEEE Press, pp. 27–38.

[18] Caves, C. M. Quantum-mechanical noise in an interferometer. *Physical Review D 28*, 8 (1981).

[19] Choudhary, S., Sloan, S., Fok, S., Neckar, A., Trautmann, E., Gao, P., Stewart, T., Eliasmith, C., and Boahen, K. Silicon neurons that compute. In

Artificial Neural Networks and Machine Learning–ICANN 2012, vol. 7552 of *Lecture Notes in Computer Science*. Springer, Berlin, Heidelberg, 2012, pp. 121–128.

[20] Cohen, J. D., Meenehan, S. M., MacCabe, G. S., Gr oblacher, S., Safavi-Naeini, A. H., Marsili, F., Shaw, M. D., and Painter, O. Phonon counting and intensity interferometry of a nanomechanical resonator. *arXiv preprint arXiv:1410.1047* (10 2014), 1–10.

[21] Cortes, C., and Vapnik, V. Support-vector networks. *Machine Learning 20*, 3 (9 1995), 273–297.

[22] Cover, T. M. Geometrical and statistical properties of systems of linear inequalities with applications in pattern recognition. *IEEE Transactions on Electronic Computers EC-14*, 3 (6 1965), 326–334.

[23] Crisafulli, O., Tezak, N., Soh, D. B. S., Armen, M. a., and Mabuchi, H. Squeezed light in an optical parametric oscillator network with coherent feedback quantum control. *Optics Express 21*, 15 (7 2013), 18371.

[24] Dejonckheere, A., Duport, F., Smerieri, A., Fang, L., Oudar, J.-L., Haelterman, M., and Massar, S. All-optical reservoir computer based on saturation of absorption. *Optics Express 22*, 9 (2014), 10868.

[25] Del'Haye, P., Schliesser, A., Arcizet, O., Wilken, T., Holzwarth, R., and Kippenberg, T. J. Optical frequency comb generation from a monolithic microresonator. *Nature 450*, 7173 (12 2007), 1214–1217.

[26] Duport, F., Schneider, B., Smerieri, A., Haelterman, M., and Massar, S. All-optical reservoir computing. *Optics Express 20*, 20 (9 2012), 22783.

[27] Englund, D., Faraon, A., Fushman, I., Stoltz, N., Petroff, P., and Vucković, J. Controlling cavity reflectivity with a single quantum dot. *Nature 450*, 7171 (12 2007), 857–861.

[28] Fisher, R. A. The use of multiple measurements in taxonomic problems. *Annals of Eugenics 7*, 2 (9 1936), 179–188.

[29] Fok, M. P., Tian, Y., Rosenbluth, D., and Prucnal, P. R. Asynchronous spiking photonic neuron for lightwave neuromorphic signal processing. *Optics letters 37*, 16 (8 2012), 3309–11.

[30] Fok, M. P., Tian, Y., Rosenbluth, D., and Prucnal, P. R. Pulse lead/lag timing detection for adaptive feedback and control based on optical spike-timing-dependent plasticity. *Optics letters 38*, 4 (2 2013), 419–21.

[31] Fredkin, E., and Toffoli, T. Conservative logic. *International Journal of Theoretical Physics 21*, 3/4 (1982), 219–253.

[32] Fushman, I., Englund, D., Faraon, A., Stoltz, N., Petroff, P., and Vuckovic, J. Controlled phase shifts with a single quantum dot. *Science (New York, N.Y.) 320*, 5877 (5 2008), 769–772.

[33] Gardiner, C. Driving a quantum system with the output field from another driven quantum system. *Physical review letters 70*, 15 (1993), 2269–2272.

[34] Gardiner, C., Parkins, A., and Zoller, P. Wave-function quantum stochastic differential equations and quantum-jump simulation methods. *Physical Rev. A 46*, 7 (1992), 4363. extlessm:linebreak extgreater extless/m:linebreak extgreater.

[35] Gardiner, C. W., and Collett, M. J. Input and output in damped quantum systems: Quantum stochastic differential equations and the master equation. *Physical Review A 31*, 6 (1985), 3761–3774.

[36] Gardiner, C. W., and Zoller, P. *Quantum Noise: A Handbook of Markovian and Non-Markovian Quantum Stochastic Methods with Applications to Quantum Optics*, 2nd ed. Springer Series in Synergetics, 2000.

[37] Gazeau, J.-P. *Coherent States in Quantum Physics.* Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim, Germany, 9 2009.

[38] Gisin, N., and Percival, I. The quantum-state diffusion model applied to open systems. *Journal of Physics A: Mathematical and General 25* (1992), 5677.

[39] Gough, J. Quantum stratonovich calculus and the quantum wong-zakai theorem. *Journal of Mathematical Physics 47*, 11 (2006), 1–24.

[40] Gough, J., and James, M. R. Quantum feedback networks: Hamiltonian formulation. *Communications in Mathematical Physics 287*, 3 (12 2008), 1109–1132.

[41] Gough, J., and James, M. R. The series product and its application to quantum feedforward and feedback networks. *IEEE Transactions on Automatic Control 54*, 11 (11 2009), 2530–2544.

[42] Gough, J., and Wildfeuer, S. Enhancement of field squeezing using coherent feedback. *Physical Review A 80*, 4 (10 2009), 1–8.

[43] Gough, J. E., James, M. R., and Nurdin, H. I. Squeezing components in linear quantum feedback networks. *Physical Review A 81*, 2 (2 2010).

[44] Gough, J. E., Nurdin, H. I., and Wildfeuer, S. Commutativity of the adiabatic elimination limit of fast oscillatory components and the instantaneous feedback limit in quantum feedback networks. *Journal of Mathematical Physics 51*, 12 (2010), 123518.

[45] Graham, R., and Haken, H. The quantum-fluctuations of the optical parametric oscillator. i. *Zeitschrift für Physik 210*, 3 (6 1968), 276–302.

[46] Hamerly, R. *COHERENT LQG CONTROL, FREE-CARRIER OSCILLATIONS, OPTICAL ISING MACHINES AND PULSED OPO DYNAMICS*. PhD thesis, Stanford University, 2016.

[47] Hamerly, R., and Mabuchi, H. Optical devices based on limit cycles and amplification in semiconductor optical cavities.

[48] Hamerly, R., and Mabuchi, H. Quantum noise of free-carrier dispersion in semiconductor optical cavities.

[49] Haribara, Y., Utsunomiya, S., Kawarabayashi, K.-i., and Yamamoto, Y. A coherent ising machine for maximum cut problems : Performance evaluation against semidefinite programming relaxation and simulated annealing. 1–10.

[50] Haroche, S., and Raimond, J.-M. *Exploring the quantum : atoms, cavities and photons*. Oxford University Press, Oxford; New York, 2006.

[51] Hudson, R. L., and Parthasarathy, K. R. Quantum ito's formula and stochastic evolutions. *Communications in Mathematical Physics 93*, 3 (9 1984), 301–323.

[52] Husimi, K. Some formal properties of the density matrix. *Proceedings of the Physico-Mathematical Society of Japan. 3rd Series 22*, 4 (1940), 264–314.

[53] Inagaki, T., Inaba, K., Hamerly, R., Inoue, K., Yamamoto, Y., and Takesue, H. Large-scale ising spin network based on degenerate optical parametric oscillators. *Nature Photonics 2*, 6 (2016), 1–3.

[54] Intel Corporation. MKL - math kernel library, 2003. `http://software.intel.com/en-us/intel-mkl`.

[55] Jacobs, K., Nurdin, H., Strauch, F., and James, M. Frequency conversion: Side-band cooling, state-swapping, and coherent control of mechanical resonators. *Arxiv preprint arXiv:1003.2653* (2010), 3–6.

[56] Johansson, J., Nation, P., and Nori, F. Qutip: An open-source python framework for the dynamics of open quantum systems. *Computer Physics Communications 183*, 8 (8 2012), 1760–1772.

[57] Kerckhoff, J., Nurdin, H., Pavlichin, D., and Mabuchi, H. Designing quantum memories with embedded control: Photonic circuits for autonomous quantum error correction. *Physical Review Letters 105*, 4 (7 2010), 040502.

[58] Kerckhoff, J., Pavlichin, D. S., Chalabi, H., and Mabuchi, H. Design of nanophotonic circuits for autonomous subsystem quantum error correction. *New Journal of Physics 13*, 5 (5 2011), 055022.

[59] Kinsler, P., and Drummond, P. D. Quantum dynamics of the parametric oscillator. *Physical Review A 43*, 11 (1991), 6194–6208.

[60] Kippenberg, T., Spillane, S., and Vahala, K. Kerr-nonlinearity optical parametric oscillation in an ultrahigh-q toroid microcavity. *Physical Review Letters 93*, 8 (8 2004), 83904.

[61] Kushner, H. J., and Dupuis, P. *Numerical Methods for Stochastic Control Problems in Continuous Time*, vol. 24 of *Stochastic Modelling and Applied Probability.* Springer New York, New York, NY, 2001.

[62] Larger, L., Soriano, M. C., Brunner, D., Appeltant, L., Gutierrez, J. M., Pesquera, L., Mirasso, C. R., and Fischer, I. Photonic information processing beyond turing: an optoelectronic implementation of reservoir computing. *Optics express 20*, 3 (1 2012), 3241–3249.

[63] Lattner, C., and Adve, V. Llvm: A compilation framework for lifelong program analysis and transformation. In *International Symposium on Code Generation and Optimization, 2004. CGO 2004.* (2004), no. c, IEEE, pp. 75–86.

[64] Levy, J. S., Gondarenko, A., Foster, M. a., Turner-Foster, A. C., Gaeta, A. L., and Lipson, M. Cmos-compatible multiple-wavelength oscillator for on-chip optical interconnects. *Nature Photonics 4*, 1 (12 2009), 37–40.

[65] Levy, M. Nanomagnetic route to bias-magnet-free, on-chip faraday rotators. *JOSA B 22*, 1 (2005), 254–260.

[66] Loudon, R., and Knight, P. Squeezed light. *Journal of Modern Optics 34*, 6-7 (6 1987), 709–759.

[67] Mabuchi, H. Derivation of maxwell-bloch-type equations by projection of quantum models. *Physical Review A 78*, 1 (7 2008), 15–18.

[68] Mabuchi, H. Coherent-feedback control strategy to suppress spontaneous switching in ultralow power optical bistability. *Applied Physics Letters 98*, 19 (2011), 193109.

[69] Mabuchi, H. Nonlinear interferometry approach to photonic sequential logic. *Applied Physics Letters 99*, 15 (2011), 153103.

[70] Mabuchi, H. Qubit limit of cavity nonlinear optics. *Physical Review A 85*, 1 (1 2012), 015806.

[71] Marandi, A., Wang, Z., Takata, K., Byer, R. L., and Yamamoto, Y. Network of time-multiplexed optical parametric oscillators as a coherent ising machine. *Nature Photonics 8*, 12 (10 2014), 937–942.

[72] MathWorks Inc. MATLAB - the language of technical computing, 1984. `http://www.mathworks.com/products/matlab`.

[73] McLachlan, G. J. *Discriminant Analysis and Statistical Pattern Recognition.* Wiley Series in Probability and Statistics. John Wiley & Sons, Inc., Hoboken, NJ, USA, 3 1992.

[74] Milburn, G. Quantum optical fredkin gate. *Physical Review Letters 62*, 18 (5 1989), 2124–2127.

[75] Miller, D. A. B. Novel analog self-electrooptic-effect devices. *Quantum Electronics, IEEE Journal of 29*, 2 (1993).

[76] Miller, D. A. B. Physical reasons for optical interconnection. *International Journal of Optoelectronics 11*, 3 (1997), 155–168.

[77] Miller, D. A. B. Are optical transistors the logical next step? *Nature Photonics 4*, 1 (1 2010), 3–5.

[78] Mirrahimi, M., Leghtas, Z., Albert, V. V., Touzard, S., Schoelkopf, R. J., Jiang, L., and Devoret, M. H. Dynamically protected cat-qubits: a new paradigm for universal quantum computation. *New Journal of Physics 16*, 4 (4 2014), 045014.

[79] Mølmer, K., Castin, Y., and Dalibard, J. Monte carlo wave-function method in quantum optics. *Journal of the Optical Society of America B 10*, 3 (3 1993), 524.

[80] Neven, H. Hardware initiative at quantum artificial intelligence lab, 2014. `http://googleresearch.blogspot.com/2014/09/ucsb-partners-with-google-on-hardware.html`.

[81] Niederberger, A. C. R., Fattal, D. a., Gauger, N. R., Fan, S., and Beausoleil, R. G. Sensitivity analysis and optimization of sub-wavelength optical gratings using adjoints. *Optics Express 22*, 11 (2014), 12971.

[82] Nielsen, A. E. B., Hopkins, A. S., and Mabuchi, H. Quantum filter reduction for measurement-feedback control via unsupervised manifold learning. *New Journal of Physics 11*, 10 (10 2009), 105043.

[83] Nozaki, K., Tanabe, T., Shinya, A., and Matsuo, S. Sub-femtojoule all-optical switching using a photonic-crystal nanocavity. *Nature Photonics 4*, July (2010).

[84] Nurdin, H. I. Synthesis of linear quantum stochastic systems via quantum feedback networks. In *Proceedings of the 48h IEEE Conference on Decision and Control (CDC) held jointly with 2009 28th Chinese Control Conference* (12 2009), IEEE, pp. 2492–2497.

[85] Nurdin, H. I., James, M. R., and Doherty, A. C. Network synthesis of linear dynamical quantum stochastic systems. *SIAM Journal on Control and Optimization 48*, 4 (12 2009), 2686.

[86] Pavlichin. *PHOTONIC CIRCUITS AND PROBABILISTIC COMPUTING.* PhD thesis, Stanford University, 2014.

[87] Pavlichin, D. S., and Mabuchi, H. Photonic circuits for iterative decoding of a class of low-density parity-check codes.

[88] Pavlichin, D. S., and Mabuchi, H. Photonic circuits for iterative decoding of a class of low-density parity-check codes. *New Journal of Physics 16*, 10 (10 2014), 105017.

[89] Pedroni, V. A. *Circuit Design and Simulation with VHDL.* MIT Press, Cambridge, 2004.

[90] Percival, I. C. Localization of wide-open quantum systems. *Journal of Physics A: Mathematical and General 27*, 3 (1999), 1003–1020.

[91] Perelomov, A. Generalized coherent states and some of their applications, 1977.

[92] Poustie, A. J., and Blow, K. J. Demonstration of an all-optical fredkin gate. *Optics Communications 174*, 1-4 (1 2000), 317–320.

[93] Puri, R. R. *Mathematical methods of quantum optics.* 2001.

[94] Radcliffe, J. M. Some properties of coherent spin states. *Journal of Physics A: General Physics 4*, 3 (2001), 313–323.

[95] Razzari, L., Duchesne, D., Ferrera, M., Morandotti, R., Chu, S., Little, B. E., and Moss, D. J. Cmos-compatible integrated optical hyper-parametric oscillator. *Nature Photonics 4*, 1 (12 2009), 41–45.

[96] Rosenblatt, F. The perceptron-a perceiving and recognizing automaton. *Report 85, Cornell Aeronautical Laboratory* (1957).

[97] Rosenblatt, F. The perceptron: a probabilistic model for information storage and organization in the brain. *Psychological review 65* (1958), 386–408.

[98] Santori, C., Pelc, J. S., Beausoleil, R. G., Tezak, N., Hamerly, R., and Mabuchi, H. Quantum noise in large-scale coherent nonlinear photonic circuits. *Physical Review Applied 1* (2014), 054005.

[99] Savchenkov, A., Matsko, A., Strekalov, D., Mohageg, M., Ilchenko, V., and Maleki, L. Low threshold optical oscillations in a whispering gallery mode caf2 resonator. *Physical Review Letters 93*, 24 (12 2004), 243905.

[100] Schack, R., Brun, T. a., and Percival, I. C. Quantum state diffusion, localization and computation. *J. Phys. A: Math. Gen. 28* (1995), 16.

[101] Shastri, B. J., Tait, A. N., Nahmias, M. A., and Prucnal, P. R. Photonic spike processing: ultrafast laser neurons and an integrated photonic network. *IEEE Photonics Society Newsletter 28*, 3 (7 2014), 4–11.

[102] Smith, S. D. Optical bistability: Towards the optical computer. *Nature 307*, 5949 (1 1984), 315–316.

[103] Steimle, T., Alber, G., and Percival, I. C. Mixed classical-quantal representation for open quantum systems. *Journal of Physics A: Mathematical and General 28*, 18 (1999), L491–L496.

[104] Strukov, D. B., Snider, G. S., Stewart, D. R., and Williams, R. S. The missing memristor found. *Nature 453*, 7191 (5 2008), 80–83.

[105] Sudarshan, E. C. G. Equivalence of semiclassical and quantum mechanical descriptions of statistical light beams. *Physical Review Letters 10*, 7 (1963), 277–279.

[106] Sussillo, D., and Barak, O. Opening the black box: low-dimensional dynamics in high-dimensional recurrent neural networks. *Neural computation 25*, 3 (3 2013), 626–49.

[107] SymPy Development Team. SymPy: Python library for symbolic mathematics, 2014.

[108] Tait, A. N., Shastri, B. J., Fok, M. P., Nahmias, M. A., and Prucnal, P. R. The dream: An integrated photonic thresholder. *Journal of Lightwave Technology 31*, 8 (4 2013), 1263–1272.

[109] Takata, K., Marandi, A., Hamerly, R., Haribara, Y., Maruo, D., Tamate, S., Sakaguchi, H., Utsunomiya, S., and Yamamoto, Y. A 16-bit coherent ising machine for one-dimensional ring and cubic graph problems. 1–39.

[110] Tan, S. M. A computational toolbox for quantum and atomic optics. *Journal of Optics B: Quantum and Semiclassical Optics 1*, 4 (8 1999), 424–432.

[111] Tezak, N. Perceptron-files, 2014. `https://github.com/ntezak/perceptron-files`.

[112] Tezak, N. QHDLJ, 2014. `https://bitbucket.org/ntezak/qhdlj.jl`.

[113] Tezak, N. QMANIFOLD, 2016. `https://github.com/ntezak/quantum_manifold`.

[114] Tezak, N., Amini, N. H., and Mabuchi, H. Low dimensional manifolds for exact representation of open quantum systems. *in preparation* (2016).

[115] Tezak, N., and Goerz, M. Qnet, 2012.

[116] Tezak, N., and Mabuchi, H. A coherent perceptron for all-optical learning. *EPJ Quantum Technology 2*, 1 (12 2015), 10.

[117] Tezak, N., Niederberger, A., Pavlichin, D. S., Sarma, G., and Mabuchi, H. Specification of photonic circuits using quantum hardware description language. *Philosophical transactions. Series A, Mathematical, physical, and engineering sciences 370*, 1979 (11 2012), 5270–90. 10.1098/rsta.2011.0526.

[118] Tezak, N., Van Vaerenbergh, T., Pelc, J. S., Mendoza, G., Kielpinski, D., and Beausoleil, R. G. All-optical ising machines based on degenerate dispersive non-linearities. *in preparation* (2016).

[119] The gEDA project. gEDA - GPL electronic design automation suite, 1998. `http://www.gpleda.org`.

[120] The Modelica Association. Modelica, 1997. `http://modelica.org`.

[121] Utsunomiya, S., Takata, K., and Yamamoto, Y. Mapping of ising models onto injection-locked laser systems. *Optics Express 19*, 19 (2011), 18091.

[122] Van Rossum, G. The python language, 1991. `http://www.python.org/`.

[123] Van Vaerenbergh, T. *All-Optical Spiking Neurons Integrated on a Photonic Chip.* 2014.

[124] Van Vaerenbergh, T., Fiers, M., Mechet, P., Spuesens, T., Kumar, R., Morthier, G., Schrauwen, B., Dambre, J., and Bienstman, P. Cascadable excitability in microrings. *Optics Express 20*, 18 (8 2012), 20292.

[125] Vandoorne, K., Dambre, J., Verstraeten, D., Schrauwen, B., and Bienstman, P. Parallel reservoir computing using optical amplifiers. *IEEE transactions on neural networks / a publication of the IEEE Neural Networks Council 22*, 9 (9 2011), 1469–81.

[126] Vandoorne, K., Dierckx, W., Schrauwen, B., Verstraeten, D., Baets, R., Bienstman, P., and Van Campenhout, J. Toward optical signal processing using photonic reservoir computing. *Optics express 16*, 15 (7 2008), 11182–92.

[127] Vandoorne, K., Mechet, P., Van Vaerenbergh, T., Fiers, M., Morthier, G., Verstraeten, D., Schrauwen, B., Dambre, J., and Bienstman, P. Experimental demonstration of reservoir computing on a silicon photonics chip. *Nature Communications 5* (3 2014), 1–6.

[128] Verstraete, F., Cirac, J. I., and Murg, V. Matrix product states, projected entangled pair states, and variational renormalization group methods for quantum spin systems. 1–99.

[129] Verstraeten, D. *Reservoir Computing : computation with dynamical systems.* PhD thesis, 2010.

[130] Walls, D. F. Squeezed states of light. *Nature 306*, 5939 (11 1983), 141–146.

[131] Wang, Z., Marandi, A., Wen, K., Byer, R. L., and Yamamoto, Y. Coherent ising machine based on degenerate optical parametric oscillators. *Physical Review A 88*, 6 (12 2013), 063853.

[132] Wigner, E. On the quantum correction for thermodynamic equilibrium. *Physical Review 40*, 5 (6 1932), 749–759.

[133] Wilcox, R. M. Exponential operators and parameter differentiation in quantum physics. *Journal of Mathematical Physics 8*, 4 (1967), 962–982.

[134] Wiseman, H. Quantum theory of continuous feedback. *Physical Review A 49*, 3 (3 1994), 2133–2150.

[135] Wiseman, H., and Milburn, G. Quantum theory of optical feedback via homodyne detection. *Physical Review Letters 70*, 5 (2 1993), 548–551.

[136] Wiseman, H. M., and Milburn, G. J. Interpretation of quantum jump and diffusion processes illustrated on the bloch sphere. *Physical Review A 47*, 3 (3 1993), 1652–1666.

[137] Wiseman, H. M., and Milburn, G. J. Quantum theory of field-quadrature measurements. *Physical Review A 47*, 1 (1 1993), 642–662.

[138] Wiseman, H. M., and Milburn, G. J. *Quantum Measurement and Control*. Cambridge University Press, Cambridge, 3 2009.

[139] Wolfram Research, I. Mathematica, 1988. `http://www.wolfram.com/mathematica`.

[140] Zhang, G., and James, M. M. R. Quantum feedback networks and control: A brief survey. *Chinese Science Bulletin 57*, 18 (5 2012), 2200–2214.