Technische Universität München

Max-Planck-Institut für Quantenoptik

Entanglement Dynamics in Quantum Information Theory

Toby S. Cubitt

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

Doktors der Naturwissenschaften (Dr. rer. nat.)

genehmigten Dissertation.

Vorsitzender: Univ.-Prof. Dr. habil. R. Gross

Prüfer der Dissertation: 1. Hon.-Prof. I. Cirac, Ph.D.
2. Univ.-Prof. Dr. W. Weise

Die Dissertation wurde am 31.10.2006 bei der
Technischen Universität München eingereicht und
durch die Fakultät für Physik am 29.3.2007 angenommen.

This thesis is dedicated to the memory of my mother, Lynette Cubitt, and to that of our close family friend Gordon Lake. Both sadly died during my doctoral studies. Neither of them were involved in academic research, but they both had a keen curiosity about science. The world needs more people like them, not less. They are sorely missed.

Abstract

This thesis contributes to the theory of entanglement dynamics, that is, the behaviour of entanglement in systems that are evolving with time. Progressively more complex multipartite systems are considered, starting with low-dimensional tripartite systems, whose entanglement dynamics can nonetheless display surprising properties, progressing through larger networks of interacting particles, and finishing with infinitely large lattice models. Firstly, what is perhaps the most basic question in entanglement dynamics is considered: what resources are necessary in order to create entanglement between distant particles? The answer is surprising: sending separable states between the parties is sufficient; entanglement can be created without it being carried by a "messenger" particle. The analogous result also holds in the continuous-time case: two particles interacting indirectly via a common ancilla particle can be entangled without the ancilla ever itself becoming entangled.

The latter result appears to discount any notion of entanglement flow. However, for pure states, this intuitive idea can be recovered, and even made quantitative. A "bottleneck" inequality is derived that relates the entanglement rate of the end particles in a tripartite chain to the entanglement of the middle one. In particular, no entanglement can be created if the middle particle is not entangled. However, although this result can be applied to general interaction networks, it does not capture the full entanglement dynamics of these more complex systems. This is remedied by the derivation of entanglement rate equations, loosely analogous to the rate equations describing a chemical reaction. A complete set of rate equations for a system reflects the full structure of its interaction network, and can be used to prove a lower bound on the scaling with chain length of the time required to entangle the ends of a chain.

Finally, in contrast with these more abstract results, the entanglement and correlation dynamics of a specific spin model is analysed. It is shown that, even when control over the system is limited to a small number of global, external, physical parameters, remarkably precise control over the correlations is still possible. They can be made to propagate in localized wave packets at a well-defined correlation speed, whilst keeping dispersion to a minimum. By varying the external parameters during the evolution, the propagation speed can be adjusted, even to the extent of reducing it to zero. These results are most conveniently derived in the fermionic Gaussian state formalism, and this is described in some detail.

Contents

List of Theorems

List of Figures

Acknowledgements

This the "Oscars speech" section, in which I thank everyone from the Queen to my pet gerbil. Never having met the Queen, and never having owned a pet gerbil, I would nonetheless like to thank a number of people who, whilst they may be less famous (than the gerbil), are more important to me (than the Queen).

Ignacio Cirac has been a wonderful guide and mentor over the course of my doctoral studies. More than anyone else, he has shaped the way I approach physics, and moulded my enthusiasm for science into something productive. And he has done this without lessening that enthusiasm, nor constraining my curiosity. On every single occasion, I have walked out of his office more excited about whatever problem we were discussing than when I walked in, and with a deeper understanding of the problem itself and its relevance to the physics community. I will forever be indebted to him for his teaching, his encouragement, and his friendship.

I would like to thank Frank Verstraete for his collaboration on a number of the results described in this thesis. Even when what he was explaining was "trivial", he was always very happy to take the time and patience to explain it to someone who found it "less trivial", and I have learnt a lot from him. I would also like to thank Michael Wolf for many fruitful discussions, whose style contrasted so greatly with Frank's but from which I learnt just as much.

The Max Planck Institute would never have been as enjoyable a place to work without all the other colleagues and friends I met there. I would especially like to thank Klemens Hammerer for his kind friendship and for being the perfect office mate, Henning Christ for his equally warm friendship and tea, Markus Popp for helping resist the lure of Henning's green tea (though he fell at the final hurdle), Barbara Kraus for her support and for many interesting discussions, Geza Giedke for being happy to discuss both physics and even stranger topics, and Maria Eckholt and Birger Horstmann for taking over from Klemens and filling my office with such cheerful conversation and plants.

I'm sure it can't be good for one's sanity to only find friends within the

physics community, and, though my friends from "outside" may at times be no less crazy than those "inside", they do provide a different brand of craziness, one that at least involves less quantum mechanics. To Emer Duggan and Christian Huber, for allowing me to turn their flat into a hotel on so many occasions, yet unfailingly making me feel that I would be welcome to stay twice as long, I can only say that I hope that I will have the opportunity to stay at Hotel Huber-Duggan for many years to come. I would also like to thank Chris Scollo and Nicole Bator, for making me so welcome every time I turned up for "Videocracy" over an hour late (which was indeed every time).

I also acknowledge the free supply of electricity for my laptop during the writing of this thesis, provided, knowingly or otherwise, by Hugenduble cafe on Marienplatz in Munich, Sunsail club Marverde in Turkey, and English Country Cottages in Cornwall. They gave me a welcome break from the more usual office setting (even if not from the work itself). For supplying free electricity in London, but far more importantly for her support in many other ways, I am deeply grateful to Linda Goldschmidt.

Finally, I can not express enough my warmest thanks to my father Roger, my brother Adam, and my girlfriend Susie Goldschmidt. Without their encouragement throughout my studies, and their support and love during the very difficult period after my mother's death, I would never have got as far as writing this. On top of having the patience to live with me for so many years, Susie has been a wonderful source of support, not to mention translation services, and has made living in Munich so much fun. I hope I'm able to return the favour even half as well.

Introduction

In the modern world of computers, telecommunication and the internet, there can be no doubt that information plays a fundamental role in our daily lives. Information theory, the theory underpinning these technological developments, is founded on classical physics. However, the laws of classical physics are only a crude approximation to the laws of nature; quantum mechanics provides a more accurate description. Quantum information theory marries these two fields, building a theory of information on a foundation of quantum mechanics. Of course, it includes classical information theory as a special case, but it also includes new concepts that have no classical counterpart. And these facilitate tasks that would be impossible if restricted to classical information theory. Currently, the two most compelling applications of quantum information are in the fields of cryptography and computation. Certain computational tasks, such as factoring large numbers [Shor, 1997] or simulating quantum systems [Feynman, 1982], can be carried out exponentially faster using a quantum computer than using the fastest known algorithm running on a classical computer. Communication that is provably secure against any eavesdropping is possible in the quantum world [Gisin et al., 2002], whereas the security of classical cryptographic protocols relies on the assumption that certain mathematical problems are computationally intractable. In fact, the security of the most common public-key crypto-systems relies on the computational difficulty of precisely those tasks that can be solved efficiently on a quantum computer! Fortunately, quantum cryptography, which is immune to such attacks, is technologically far easier to implement than quantum computation.

A *leitmotiv* running through much of quantum information theory is the concept of *entanglement*. Quantum cryptographic protocols can be constructed based on distribution of entangled states [Ekert, 1991], and a number of the security proofs for more practical quantum cryptographic schemes invoke entanglement [Nielsen and Chuang, 2000, Section 12.6]. A quantum computer running Shor's factorization algorithm generates a large amount of entanglement between the different qubits it contains. Indeed, if an algorithm does *not* engender the creation of entanglement, it can be

simulated efficiently on a classical computer, and a quantum computer is of no benefit [Vidal, 2003].

The concept of entanglement, and indeed the coining of the word itself, dates back to the origins of quantum mechanics [Schrödinger, 1935, 1936]. But quantum information theory has provided a new perspective on entanglement: no longer "spooky" quantum phenomenon [Einstein et al., 1935] but quantifiable resource. Much like energy or entropy, entanglement can be exploited in order to carry out information processing tasks, such as quantum teleportation [Bennett et al., 1993]) or cryptography [Ekert, 1991]. There is now a vast literature on entanglement theory. Entanglement of bipartite, pure quantum states is well understood [Nielsen and Chuang, 2000, Section 12.5], and provides a framework for studies of entanglement in more complex systems. However, many basic questions remain unanswered. Even the most basic of all, that of determining whether a given quantum state is entangled or not, is as yet unsolved in the general case.

The majority of the work in entanglement theory has concentrated on its *static* properties; the entanglement properties of fixed quantum states. The study of the *dynamic* behaviour of entanglement is inevitably more difficult; static properties must be understood before any attempt can be made to understand the behaviour of entanglement in quantum systems that are evolving in time. Nonetheless, with the tools of modern entanglement theory, it is now possible to tackle some of these questions. This thesis contributes to the nascent theory of entanglement dynamics. The questions considered in successive chapters involve increasingly complex multipartite systems, starting from low-dimensional tripartite systems and finishing with infinitely large lattice models. In the following, we give an overview of this progression.

We start in Chapter 1 with what is perhaps one of the most basic of entanglement dynamics questions: how can entanglement be created between two spatially separated particles? The scenario can be phrased in terms of two parties who, following a long tradition, we will call Alice and Bob. They each have a particle, and they would like to entangle the two, but the particles are too far apart to interact directly. Since, by definition, local operations and classical communication can not create entanglement, and the particles can not interact directly, it is clearly necessary for Alice and Bob to use quantum communication. One obvious solution is for Alice to entangle her particle with a "messenger" particle and send the latter to Bob, who can then exchange it for his particle. In this way, the messenger particle carries the entanglement from Alice to Bob. Is it *necessary* to send an entangled messenger? Surprisingly, we will show that it is not: Alice and Bob can generate entanglement by sending only *unentangled* messenger particles. To

show this, we start by proving an equivalent result for a system of two particles interacting with a common ancilla. Starting from a fully separable state, the particles can be entangled without the ancilla *ever* becoming entangled. This continuous-time result immediately implies the existence of a discrete protocol in which only unentangled messenger particles are sent. We then give an explicit example of such a protocol, and finally extend this via the Jamiolkowski isomorphism to an analogous result for quantum channels. (The results of this chapter are published in Cubitt et al. [2003].)

The continuous-time result of Chapter 1 forces us to abandon any notion of entanglement flowing through particles. Yet the results rely heavily on the entanglement properties of mixed states. If the system is restricted to pure states, the effect is easily shown to be impossible; if the overall system state is pure and the ancilla is not entangled with the other particles, no entanglement can be generated. This suggests that, for pure states, an unentangled ancilla acts as a bottleneck to entanglement flow. In Chapter 2, we make the concept of entanglement bottlenecks quantitative. Starting with a three-qubit chain, the simplest tripartite system, we derive a bottleneck inequality that relates the rate of entanglement generation between the end qubits to the entanglement of the middle qubit. To generalize this to all tripartite systems, we face the problem that, whilst the overall state may be pure, the reduced state of the two end particles will in general be mixed. And it is not known how to determine whether a mixed state of a general bipartite system is entangled or not, let alone quantify that entanglement (this has only been solved in the special case of two qubits). For that reason, we introduce the *entanglement fidelity*, a quantity that, whilst not an entanglement measure in the strict sense, is physically motivated and gives significant insight into entanglement properties. This allows us to prove a very similar bottleneck inequality for general tripartite chains, but in terms of the entanglement fidelity. (Both these bottleneck results can be found in Cubitt et al. [2005].)

The general bottleneck inequality can be applied to a far wider range of systems than merely tripartite chains. Any multipartite system in which the particles do not all interact directly with each other can be divided into three subsystems, such that one subsystem mediates all interactions between the other too. The bottleneck inequality then implies that the entanglement of this mediating system limits the rate of entanglement generation between the other two. Nonetheless, there are a number of reasons why this is not wholly satisfying. Firstly, as seen from the results of Chapter 1, the argument does not apply to systems in mixed states. And secondly, the internal structure of the subsystems (that is, the interactions between particles grouped within the same subsystem) is ignored.

Chapter 3 addresses these defects, and establishes a quantitative concept

of entanglement flow for completely general systems of interacting particles. The motivation comes from the rate equations that describe complex chemical reactions. Each step in the reaction mechanism can be described quantitatively by a differential equation, relating the concentrations of reactants to the rate at which product compounds are created. The structure of the complete reaction pathway is therefore captured mathematically by a set of coupled differential equations. We derive an analogous *entanglement* rate equation, relating the derivative of the entanglement fidelity of any subsystem to the entanglement fidelity one step back along the interaction network. Thus the complete set of entanglement rate equations for a system captures the full structure of its interaction network. Moreover, the results apply equally well to mixed and pure states. (The results on entanglement flow are also published in Cubitt et al. [2005].)

As a demonstration of the utility of the rate equations, we analyse two very different protocols for generating entanglement between the ends of a spin chain, and see that the entanglement fidelities appearing in the rate equations indeed reflect the different ways in which the two protocols generate entanglement. We then use the rate equations to prove a universal lower bound on the time required to entangle the ends of a spin chain or, more precisely, the scaling of that time with the length of the chain. Subsequent to this work being published, a tighter bound was proven [Bravyi et al., 2006], and it remains an interesting topic for future research to determine whether the inequalities appearing in the rate equations can be improved upon.

In contrast to the general and more abstract results of the other chapters, Chapter 4 considers a specific, simple spin model: the XY–model, extensively studied in condensed matter physics. One motivation for studying entanglement dynamics in such systems is that they provide a simple physical model of a quantum repeater [Briegel et al., 1998, Dür et al., 1999a], a device used to distribute entanglement over long distances. Various aspects of the correlation dynamics of the XY–model have been analysed before. But, in contrast to previous work, and motivated by the quantum repeater connection, we investigate to what extent the correlation dynamics can be *controlled and engineered* using only the two simple, physical parameters that appear in the translationally invariant Hamiltonian. By calculating analytically the time-evolution of two-point correlation functions, we find that the correlation dynamics takes the form of wave-packets, propagating according to the dispersion relation given by the system spectrum. This simple wave-packet description allows us to show that correlations can be made to propagate in localized packets, at a well-defined correlation velocity. Indeed, the system parameters allow the correlation velocity to be engineered as desired and, by changing the parameters as the system evolves, increased and decreased as

the correlations propagate. By "quenching" the system, correlations can even be frozen once they have reached a desired location. (These results can be found in Cubitt and Cirac [2006a,b].)

The XY Hamiltonian studied in Chapter 4 is diagonalized by mapping it to a fermionic model, a standard result from condensed matter physics [Lieb et al., 1961]. This fermionic model is quadratic in terms of the creation and annihilation operators. Were it a bosonic model, it would be most conveniently analysed in the Gaussian state formalism, as commonly used in quantum optics. Although it has not been as extensively studied, it is possible to develop a *fermionic* Gaussian state formalism. This is carried out in Appendix A. Although a number of the results are not new and can be found in the literature (especially those concerning Grassmann algebra and fermionic coherent states; see Cahill and Glauber [1999], Simons [2001]), we include them here as they are not widely known, and are needed in order to develop the fermionic Gaussian state formalism used to derive the results of Chapter 4. (Cubitt and Cirac [2006b] includes a more concise treatment of these results.)

Chapter 1

Distributing Entanglement Using Separable States

In this chapter, we will consider what is perhaps one of the most basic questions in the theory of entanglement dynamics: how can entanglement be created between two spatially-separated particles? More concretely, if Alice and Bob each possess a particle, what steps are necessary in order to entangle the two particles if they are too far apart to interact directly?

It is clear they must exchange quantum information. One obvious solution is for Alice to send half of an entangled pair to Bob. This "messenger" particle would then carry the entanglement from Alice to Bob. But is it *necessary* to send an entangled messenger?

A similar question can be posed in a slightly different scenario. Consider two particles that do not interact directly with each other, but interact with the same ancilla particle, so that interactions can be mediated by the ancilla. How entangled must the mediating particle become in order to transmit entanglement between the two particles?

In the following, we will give a rather surprising answer to these questions: entanglement can be created between Alice and Bob by sending a messenger particle that is *not entangled with any other particle*. Or, in the continuous setting, entanglement can be created between two particles interacting with a common ancilla that remains in a separable state *at all times*.

Although the result perhaps appears more surprising in the discrete setting, the continuous setting gives rise to the stronger result. The particle mediating the interactions must remain unentangled throughout the evolution, whereas the messenger particle is only required to be unentangled after each step in the discrete process. Moreover, by rewriting the continuous evolution in terms of the Trotter decomposition, the continuous result immediately implies the corresponding discrete result.

Therefore, we will first consider the continuous setting, giving a physical motivation for the result in Section 1.1, before providing a rigorous proof in Section 1.2. By considering the entanglement properties of the states produced during the evolution, we will gain more insight into this counter-intuitive result, leading in Section 1.3 to an explicit example of a discrete procedure involving only a single transfer of a messenger particle. Finally, in Section 1.4, we will consider an analogous result for quantum channels rather than states.

1.1 Physical Motivation

Before giving rigorous proofs of the results in this chapter, it is useful to gain some physical insight into how this might occur.

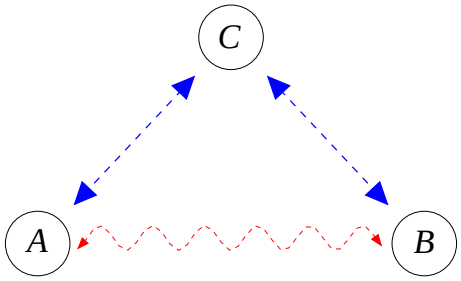

Figure 1.1: Particles A and B interact continuously with the same mediating particle C (blue lines), inducing a virtual interaction between A and B themselves (red line). For certain initial states, C remains separable from A and B *at all times* during the evolution of the system, yet at the end A and B are strongly entangled.

We will model the system by two particles, A and B, both interacting with the same ancilla particle, C, but not directly with each other (Fig. 1.1). If we restrict ourselves to pure states, the claimed effect is impossible: to entangle A and B, C must become entangled with them. This can be seen by looking at the infinitesimal change of a pure, separable state $|a\rangle|b\rangle|c\rangle$ under the action of a Hamiltonian of the form $H_{AC}\otimes\mathbb{1}_B+\mathbb{1}_A\otimes H_{BC}$. The condition that the ancilla remains separable after an infinitesimal time δt is

$$\big(\mathbb{1}+\delta t(H_{AC}+H_{BC})\big)\,|a\rangle|b\rangle|c\rangle = \big(|a\rangle|b\rangle+\delta t\,|\psi_{ab}\rangle\big)\big(|c\rangle+\delta t\,|\psi_c\rangle\big) \tag{1.1}$$

where $|\psi_{ab}\rangle$ and $|\psi_c\rangle$ are unnormalized states. The state $|a\rangle|b\rangle+\delta t\,|\psi_{ab}\rangle$ is entangled to first order in δt if and only if $|\psi_{ab}\rangle$ has a non-vanishing component $\langle a^\perp, b^\perp|\psi_{ab}\rangle$ for some $|a^\perp\rangle, |b^\perp\rangle$ orthogonal to $|a\rangle, |b\rangle$. However, multiplying the above equation by $\langle a^\perp|\,\langle b^\perp|$ and tracing over the ancilla gives $\langle a^\perp, b^\perp|\psi_{ab}\rangle = 0$.

Let us however consider the evolution from a more physical perspective, which will give insight into how the entanglement of the ancilla is related to the entanglement that can be transmitted. Physically, the ancilla has a spectrum of energy levels corresponding to the eigenstates of its Hamiltonian. Imagine that the interactions with particles A and B couple eigenstates of the ancilla, and imagine that the interactions are weak, so that transitions between energy levels are virtual. The most significant processes in the evolution of the system are then second-order virtual transitions of C from an energy level back to itself, accompanied by an effective interaction between A and B. If the ancilla starts in an energy eigenstate, the virtual transition will leave it in the same eigenstate, not entangled with A or B. It *will* become entangled due to less significant higher-order processes, but that entanglement will be weak. Particles A and B, however, can become strongly entangled by the effective interaction if it lasts a long time, and we have a situation close to the the claimed one: A and B can become entangled while the ancilla remains *almost* separable.

But assume now that there is noise in the system, which can be modelled by mixing the quantum states with the maximally mixed state. Mixing a barely entangled state with an amount of noise proportional to the entanglement destroys the entanglement ([Braunstein et al., 1999, Vidal and Tarrach, 1999]). We can therefore add a sufficiently large weight of the maximally mixed state to remove the weak entanglement with the ancilla due to higher-order processes, but not so large that it destroys the strong entanglement between A and B built up by the leading-order processes. Thus, if *mixed* states are considered, the desired effect seems to be possible.

1.2 Continuous Case

1.2.1 System and Hamiltonian

In a number of the physical systems used for quantum information processing, the qubits do not interact directly with each other. Rather, all qubits interact with a common bosonic mode that is used as a "data bus" to transmit information between them. For example, in the original ion trap quantum computation scheme ([Cirac and Zoller, 1995]), internal states of the ions play the role of the qubits, and the data bus consists of the collective vibrational modes of the trapped ions. Or, in cavity-QED experiments, atoms interact via the cavity mode.

Motivated by these, we consider a simplified system in which the bosonic mode has been truncated to a qutrit C, and only two qubits A and B are

present, each interacting weakly with the ancilla qutrit. The Hamiltonian is then given by

$$H = c^\dagger c \otimes \mathbb{1}_{AB} + \frac{\epsilon}{2}\left(c + c^\dagger\right) \otimes \left(\sigma_A^x + \sigma_B^x\right), \tag{1.2}$$

where

$$c = |0\rangle\langle 1|_C + \sqrt{2}\,|1\rangle\langle 2|_C \tag{1.3}$$

is the "annihilation" operator in the truncated Hilbert space of the ancilla, so that

$$c^\dagger c = |1\rangle\langle 1|_C + 2\,|2\rangle\langle 2|_C\,, \tag{1.4}$$

and ϵ characterizes the strength of the interaction. The interactions couple energy levels of the ancilla, as described in the previous section, and are weak if ϵ is small.

Although we could in principle diagonalize the Hamiltonian and calculate the evolution operator exactly, this would give little insight into the physics of the system, and would in any case be of little use. The analysis in the previous section shows that we must consider mixed states of the system, and it is not known how to even determine whether such a state is entangled or not in Hilbert spaces of dimension larger than six. Motivated by the previous discussion, we are interested in the leading-order evolution (and how much the true evolution deviates from it), so we turn instead to perturbation theory.

The Hamiltonian operates in four invariant subspaces, given by projecting on the eigenstates $|\pm\rangle = (|0\rangle + |1\rangle)/\sqrt{2}$ of the σ^x operators:

$$H = H_C^+ \otimes \mathcal{P}_{AB}^{++} + H_C^0 \otimes \left(\mathcal{P}_{AB}^{+-} + \mathcal{P}_{AB}^{-+}\right) + H_C^- \otimes \mathcal{P}_{AB}^{--}, \tag{1.5}$$

where

$$H^0 = c^\dagger c, \quad H^\pm = c^\dagger c \pm \epsilon(c + c^\dagger), \tag{1.6}$$

and the $\mathcal{P}$'s denote projectors on the two-qubit states indicated by their superscripts. The evolution operator, then, can be written as a direct sum of operators on these spaces:

$$U = e^{-iH^+ t} \oplus e^{-iH^0 t} \oplus e^{-iH^0} \oplus e^{-iH^- t}. \tag{1.7}$$

Expanding $H^\pm$ using standard perturbation theory, we have

$$e^{-iH^\pm t} = e^{-iDt} + O(\epsilon^4 t) + O(\epsilon), \tag{1.8}$$

where

$$D = \mathrm{diag}(-\epsilon^2, 1 - \epsilon^2, 2 + 2\epsilon^2) \tag{1.9}$$

is the matrix of eigenvalues of $H^{\pm}$ approximated to leading-order (in fact accurate to third order since all odd-order terms vanish in this particular case). The $O(\epsilon^4 t)$ term arises from higher-order perturbations of the eigenvalues, the $O(\epsilon)$ term from perturbation of the eigenstates. Note that perturbations to the eigenvalues of the Hamiltonian are exponentiated in the evolution operator and accumulate over time, whereas perturbations to the eigenvectors do not.

From Eq. (1.8), and using Eq. (1.7) for the evolution operator, we can approximate the evolution operator by

$$U = U_{\text{eff}} + O(\epsilon^4 t) + O(\epsilon), \tag{1.10}$$

where

$$U_{\text{eff}} = e^{-iDt} \oplus e^{-iH_0 t} \oplus e^{-iH_0 t} \oplus e^{-iDt}. \tag{1.11}$$

How do states evolve under U_{eff}? If we start the ancilla C in an eigenstate of U_{eff}, its state remains unchanged up to a global phase, and can not become entangled with A and B. If A and B start in a superposition of eigenstates of U_{eff}, the $|++\rangle$ and $|--\rangle$ portions acquire a phase difference relative to the $|+-\rangle$ and $|-+\rangle$ portions ($e^{-2i\epsilon^2 t}$ if the ancilla is in the $|2\rangle$ state, $e^{i\epsilon^2 t}$ otherwise, recalling the expressions for D and H^0). Thus the effective interaction generates a controlled phase gate between qubits A and B, so that these can become highly entangled for times of order $1/\epsilon^2$, even though the ancilla remains separable at all times. In the limit of small ϵ, the evolution of the system will be "close" to evolution under U_{eff}. In the next section, we give a rigorous analysis of what is meant here by "close".

1.2.2 Bound on the Approximate Evolution

The true evolution under U deviates from that under U_{eff} due to the higher-order terms in Eq. (1.10), and *will* entangle the ancilla if the system starts in a pure state. To show that a small amount of noise is sufficient to remove this entanglement whilst leaving A and B highly entangled, we must bound this deviation at times $t = O(1/\epsilon^2)$ to $O(\epsilon)$. That is, we must bound the deviation between what we *would* get if the system evolved under U_{eff}, and what we *do* get when the system evolves under U. Comparing U in Eq. (1.7) with U_{eff} in Eq. (1.11), we will therefore need to bound the matrix norm $\|e^{-iH^{\pm}t} - e^{-iDt}\|$.

Define X to be the matrix of eigenvectors of $H^{\pm}$ approximated to third order in ϵ, and recall that D was defined to be the matrix of eigenvalues approximated to second order (accurate up to third order), so that $H^{\pm}X =$

$XD + O(\epsilon^4)$. By standard perturbation theory,

$$D = \begin{pmatrix} -\epsilon^2 & 0 & 0 \\ 0 & 1-\epsilon^2 & 0 \\ 0 & 0 & 2+2\epsilon^2 \end{pmatrix} \tag{1.12a}$$

$$X = \begin{pmatrix} 1 & \epsilon+\epsilon^3 & \epsilon^2/\sqrt{2} \\ -\epsilon & 1 & (2\epsilon-3\epsilon^3)/\sqrt{2} \\ \epsilon^2/\sqrt{2} & -\sqrt{2}\epsilon+\sqrt{2}\epsilon^3 & 1 \end{pmatrix}. \tag{1.12b}$$

Using the triangle inequality, we have

$$\|e^{-iH^{\pm}t} - e^{-iDt}\| \leq \|e^{-iH^{\pm}t} - Xe^{-iDt}X^{-1}\| + \left\|Xe^{-iDt}X^{-1} - e^{-iDt}\right\|. \tag{1.13}$$

Since both X and D are accurate to third order in ϵ, the first term will be of order $\epsilon^4 t$, whereas, as eigenvalues are invariant under similarity transformations and $X = \mathbb{1} + O(\epsilon)$, the second term will be of order ϵ. Equation (1.13) therefore separates the deviation into a term that is dominated by perturbations of the eigenvalues, and one that is dominated by perturbations of the eigenvectors.

To bound the first term of Eq. (1.13), we will use two standard theorems from linear algebra.

Theorem 1.1 (Matrix Inverse Norm Bound)
For an invertible matrix A,

$$\left\|A^{-1}\right\| = \left\|(\mathbb{1} - (\mathbb{1} - A))^{-1}\right\| \leq \frac{1}{1 - \|\mathbb{1} - X\|} \tag{1.14}$$

Proof See Golub and van Loan [1996, Section 2.3.4]. □

Theorem 1.2 (Matrix Exponential Integral Identity)
For matrices A and B, the difference of their exponentials can be written

$$e^{-iAt} - e^{-iBt} = \int_0^t \mathrm{d}s\, e^{-iB(t-s)}\,(A-B)\,e^{-iAs}. \tag{1.15}$$

Proof See Golub and van Loan [1996, Section 11.3.2]. □

Applying these to the first term in Eq. (1.13), we can obtain an upper bound.

Lemma 1.3 (Matrix Norm 1)
For any Hermitian matrices $H^{\pm}$ and D, an arbitrary matrix X, and any unitarily invariant norm,

$$\left\| e^{-iH^{\pm}t} - Xe^{-iDt}X^{-1} \right\| \leq \frac{\|X\|}{(1 - \|\mathbb{1} - X\|)^2} \left\| H^{\pm}X - XD \right\| t. \tag{1.16}$$

With $H^{\pm}$, D and X given by Eq. (1.6)*, Eq.* (1.12a) *and Eq.* (1.12b)*, this bound is of order $\epsilon^4 t$.*

Proof Using Theorems 1.1 and 1.2,

$$\begin{aligned}
&\left\| e^{-iH^{\pm}t} - Xe^{-iDt}X^{-1} \right\| \\
&\quad \leq \int_0^t \mathrm{d}s \left\| Xe^{-iD(t-s)}X^{-1} \left(H^{\pm} - XDX^{-1} \right) e^{-iH^{\pm}s} \right\| && (1.17a) \\
&\quad \leq \int_0^t \mathrm{d}s \, \|X\| \left\| X^{-1} \right\| \left\| H^{\pm} - XDX^{-1} \right\| && (1.17b) \\
&\quad \leq \left\| X^{-1} \right\|^2 \|X\| \left\| H^{\pm}X - XD \right\| t && (1.17c) \\
&\quad \leq \frac{\|X\|}{(1 - \|\mathbb{1} - X\|)^2} \left\| H^{\pm}X - XD \right\| t. && (1.17d)
\end{aligned}$$

Note that we have used the unitary invariance of the norm in Eq. (1.17b). Since $H^{\pm}X = XD + O(\epsilon^4)$, this bound is of order $\epsilon^4 t$, as claimed. □

We also need to bound the second term of Eq. (1.13).

Lemma 1.4 (Matrix Norm 2)
For 3×3 matrices D and X, with D diagonal,

$$\begin{aligned}
&\left\| Xe^{-iDt}X^{-1} - e^{-iDt} \right\| \\
&\qquad \leq \frac{2\left(d_3^2 - d_1^2\right) \|XD - DX\| + 2\left(d_3 - d_1\right) \|XD^2 - D^2X\|}{|\Delta| \left(1 - \|\mathbb{1} - X\|\right)^2}.
\end{aligned} \tag{1.18}$$

With D and X given by Eq. (1.12a) *and Eq.* (1.12b)*, this bound is of order ϵ independent of t.*

Proof We begin by noting that a matrix satisfies its own characteristic equation, giving an equation for the n^{th} power of the matrix in terms of lower powers (where n is the dimension of the matrix). By repeated substitution of this relation, we can reduce the (infinite-order) series expansion for the matrix exponential to a polynomial of order $n - 1$.

For the 3×3 matrix D, we have

$$e^{-iDt} = c_2(t)D^2 + c_1(t)D + c_0(t)\mathbb{1} \tag{1.19}$$

which, since D is diagonal, we can rewrite as

$$\begin{pmatrix} e^{-id_1 t} \\ e^{-id_2 t} \\ e^{-id_3 t} \end{pmatrix} = \begin{pmatrix} 1 & d_1 & d_1^2 \\ 1 & d_2 & d_2^2 \\ 1 & d_3 & d_3^2 \end{pmatrix} \begin{pmatrix} c_0(t) \\ c_1(t) \\ c_2(t) \end{pmatrix} \tag{1.20}$$

where d_i are the entries of D in increasing order. Inverting the Vandermonde matrix appearing in this relation gives

$$\begin{pmatrix} c_0(t) \\ c_1(t) \\ c_2(t) \end{pmatrix} = \frac{1}{\Delta} \begin{pmatrix} d_2^2 d_3 - d_3^2 d_2 & d_3^2 d_1 - d_1^2 d_3 & d_1^2 d_2 - d_2^2 d_1 \\ d_3^2 - d_2^2 & d_1^2 - d_3^2 & d_2^2 - d_1^2 \\ d_2 - d_1 & d_3 - d_1 & d_1 - d_2 \end{pmatrix} \begin{pmatrix} e^{-id_1 t} \\ e^{-id_2 t} \\ e^{-id_3 t} \end{pmatrix} \tag{1.21}$$

where $\Delta = d_1^2(d_2 - d_3) + d_2^2(d_3 - d_1) + d_3^2(d_1 - d_2)$ is the determinant of the Vandermonde matrix.

We bound the c_1 and c_2 by assuming the exponentials sum in phase, giving

$$\sup_t |c_1(t)| \leq \frac{2}{|\Delta|}\left(d_3^2 - d_1^2\right) \tag{1.22a}$$

$$\sup_t |c_2(t)| \leq \frac{2}{|\Delta|}\left(d_3 - d_1\right). \tag{1.22b}$$

(These bounds are tight, since in general there is no reason to assume the d_i will be commensurate.)

Finally, using Eqs. (1.19) and (1.22) and Theorem 1.1, we obtain the following bound on the second term of Eq. (1.13):

$$\begin{aligned} &\left\| X e^{-iDt} X^{-1} - e^{-iDt} \right\| \\ &\leq \left\| X^{-1} \right\| \left\| X e^{-iDt} - e^{-iDt} X \right\| && \text{(1.23a)} \\ &= \left\| X^{-1} \right\| \left\| X \left(c_2(t)D^2 + c_1(t)D\right) - \left(c_2(t)D^2 + c_1(t)D\right) X \right\| && \text{(1.23b)} \\ &\leq \frac{2\left(d_3^2 - d_1^2\right)\|XD - DX\| + 2\left(d_3 - d_1\right)\|XD^2 - D^2X\|}{|\Delta|\left(1 - \|\mathbb{1} - X\|\right)}. && \text{(1.23c)} \end{aligned}$$

Since $X = \mathbb{1} + O(\epsilon)$, this is of order ϵ at all times, as claimed. □

We are now in a position to give a bound on the divergence of the approximated evolution of a state ρ under U_{eff} from the exact evolution under

U. We can express the time-evolution of the system as

$$U\rho U^\dagger = U_{\text{eff}}\rho U_{\text{eff}}^\dagger + \kappa\epsilon E, \tag{1.24a}$$

$$\kappa = \frac{1}{\epsilon}\big\|U\rho U^\dagger - U_{\text{eff}}\rho U_{\text{eff}}^\dagger\big\|, \tag{1.24b}$$

where the operator E is Hermitian and has $\text{Tr}[E] = 0$ to ensure that the right hand side is a valid density matrix, and we have absorbed its norm into the factor κ, so that $\|E\| = 1$. (We have included the factor of $1/\epsilon$ in the definition of κ for later convenience.)

Now,

$$\begin{aligned}&\big\|U\rho U^\dagger - U_{\text{eff}}\rho U_{\text{eff}}^\dagger\big\| \\ &\quad\le \Big(\big\|U - U_{\text{eff}}\big\| + \big\|U^\dagger - U_{\text{eff}}^\dagger\big\| + \big\|U - U_{\text{eff}}\big\|\,\big\|U^\dagger - U_{\text{eff}}^\dagger\big\|\Big)\,\|\rho\|\,.\end{aligned} \tag{1.25}$$

Recalling the forms of U and U_{eff} from Eq. (1.7) and eqrefeq:messenger:Ueff, and using Eq. (1.13), we have

$$\big\|U - U_{\text{eff}}\big\| = \big\|U^\dagger - U_{\text{eff}}^\dagger\big\| \tag{1.26a}$$

$$\le \big\|e^{-iH^+t} - e^{-iDt}\big\| + \big\|e^{-iH^-t} - e^{-iDt}\big\| \tag{1.26b}$$

$$\le 2\Big(\big\|e^{-iH^\pm t} - Xe^{-iDt}X^{-1}\big\| + \big\|Xe^{-iDt}X^{-1} - e^{-iDt}\big\|\Big) \tag{1.26c}$$

$$= 2Q, \tag{1.26d}$$

defining, for notational convenience,

$$Q = \big\|e^{-iH^\pm t} - Xe^{-iDt}X^{-1}\big\| + \big\|Xe^{-iDt}X^{-1} - e^{-iDt}\big\|. \tag{1.27}$$

The norms on the right hand side are already bounded by Lemma 1.3 and Lemma 1.4, respectively. Using Eq. (1.25) in Eq. (1.24b) defining κ (the error in approximating evolution of a state under U by evolution under U_{eff}), we have

$$\kappa \le \frac{1}{\epsilon}\Big(\big\|U - U_{\text{eff}}\big\| + \big\|U^\dagger - U_{\text{eff}}^\dagger\big\| + \big\|U - U_{\text{eff}}\big\|\,\big\|U^\dagger - U_{\text{eff}}^\dagger\big\|\Big) \tag{1.28a}$$

$$= \frac{1}{\epsilon}\big\|U - U_{\text{eff}}\big\|\Big(2 + \big\|U - U_{\text{eff}}\big\|\Big) \tag{1.28b}$$

$$\le \frac{4Q}{\epsilon}\Big(1 + Q\Big). \tag{1.28c}$$

1.2.3 Analysis of the Exact Evolution

We now consider the exact evolution of the initial state

$$\rho = \frac{1}{(1+\alpha)^2}\left(|0\rangle\langle 0|_A + \alpha\,\frac{\mathbb{1}_A}{2}\right) \otimes \left(|0\rangle\langle 0|_B + \alpha\,\frac{\mathbb{1}_B}{2}\right) \otimes \frac{\mathbb{1}_C}{3} \tag{1.29a}$$

$$= \frac{1}{(1+\alpha)^2}\left(\rho_0 + \alpha\rho_a + \alpha\rho_b + \alpha^2\,\frac{\mathbb{1}_{ABC}}{12}\right), \tag{1.29b}$$

with

$$\rho_0 = |00\rangle\langle 00|_{AB} \otimes \frac{\mathbb{1}_C}{3}, \tag{1.30a}$$

$$\rho_a = |0\rangle\langle 0|_A \otimes \frac{\mathbb{1}_{BC}}{6}, \tag{1.30b}$$

$$\rho_b = |0\rangle\langle 0|_B \otimes \frac{\mathbb{1}_{AC}}{6}. \tag{1.30c}$$

Note that ρ is not only completely separable, it also contains no classical correlations.

We now fix the matrix norm we have been using throughout to be the trace norm, defined as the sum over the singular values, so that $\|\rho\|_{\mathrm{Tr}} = 1$ for any density matrix ρ. As it only depends on the singular values, the trace norm is unitarily invariant, as assumed in Lemma 1.3. Under the exact evolution, using Eq. (1.24) for the evolution of a state under U in terms of U_{eff} and an approximation error, we then have

$$\begin{aligned} U\rho U^\dagger = \frac{1}{(1+\alpha)^2}\Big(&U_{\mathrm{eff}}\rho_0 U_{\mathrm{eff}}^\dagger + \kappa_0\epsilon E_0 + \alpha\big(U_{\mathrm{eff}}\rho_a U_{\mathrm{eff}}^\dagger + \kappa_a\epsilon E_a\big) \\ &+ \alpha\big(U_{\mathrm{eff}}\rho_b U_{\mathrm{eff}}^\dagger + \kappa_b\epsilon E_b\big) + \alpha^2\frac{\mathbb{1}^{ABC}}{12}\Big). \end{aligned} \tag{1.31}$$

The ancilla C is in a mixture of eigenstates of H^0 (Eq. (1.6)) so, from the discussion in Section 1.1, we expect it to remain separable under evolution by U_{eff}, whereas the qubits A and B may become entangled. Indeed, at time $t = 2\pi/3\epsilon^2$, the reduced state of the qubits

$$\mathrm{Tr}_C\left[U_{\mathrm{eff}}\rho U_{\mathrm{eff}}^\dagger\right] = \frac{1}{4}\Big(|++\rangle - \frac{1+\sqrt{3}i}{2}\big(|+-\rangle + |-+\rangle\big) + |--\rangle\Big) \tag{1.32}$$

is a *pure* entangled state.

To proceed, we will need to turn the terms containing E's in Eq. (1.31) into valid density operators, which we will do by combining them with parts of the other terms. As eigenvalues are invariant under unitary transformations, both

$U\rho_0 U^\dagger$ and $U_{\text{eff}}\rho_0 U_{\text{eff}}^\dagger$ are rank 3, with the same three non-zero eigenvalues $+\frac{1}{3}$. From Lemmas 1.3 and 1.4, κ is of order 1 for times of order $1/\epsilon^2$. By the following perturbation theorem on eigenvalues of Hermitian matrices, we can choose ϵ small enough that adding $\kappa\epsilon E_0$ to $U_{\text{eff}}\rho_0 U_{\text{eff}}^\dagger$ can not change its range.

Theorem 1.5 (Hermitian Eigenvalue Perturbation)
If A and B are Hermitian, and if $\lambda_1 \leq \lambda_2 \leq \cdots \leq \lambda_n$ and $\lambda'_1 \leq \lambda'_2 \leq \cdots \leq \lambda'_n$ are the ordered eigenvalues of A and $A+B$ respectively, then $|\lambda'_k - \lambda_k| \leq \rho(B)$ where $\rho(B)$ is the spectral radius of B.

Proof See Horn and Johnson [1985, Section 6.3]. □

This implies that the range of E_0 must be contained in the range of $U_{\text{eff}}\rho_0 U_{\text{eff}}^\dagger$. The most negative E_0 consistent with $\text{Tr}[E_0] = 0$ and $\|E_0\|_{\text{Tr}} = 1$ has two non-zero eigenvalues $\pm\frac{1}{2}$. Therefore, the Hermitian operator

$$\rho_{E_0} = U_{\text{eff}}\rho_0 U_{\text{eff}}^\dagger + \frac{2}{3}E_0 \tag{1.33}$$

is guaranteed to be positive and has unit trace, i.e. it is a density operator.

By similar arguments, the operators

$$\rho_{E_{a/b}} = U_{\text{eff}}\rho_{a/b} U_{\text{eff}} + \frac{1}{3}E_{a/b} \tag{1.34}$$

are also density operators. We can then rewrite Eq. (1.31) describing the evolution of our initial state as

$$\begin{aligned} U\rho U^\dagger = \frac{1}{(1+\alpha)^2}\Bigg(&\left(1 - \frac{3}{2}\kappa\epsilon\right) U_{\text{eff}}\rho_0 U_{\text{eff}}^\dagger + \alpha(1-3\kappa\epsilon) U_{\text{eff}}(\rho_a + \rho_b) U_{\text{eff}}^\dagger \\ &+ \left(\frac{3}{2} + 6\alpha\right)\kappa\epsilon\rho_E + \alpha^2 \frac{\mathbb{1}^{ABC}}{12}\Bigg), \end{aligned} \tag{1.35}$$

where we have incorporated ρ_{E_0}, ρ_{E_a} and ρ_{E_b} into a single density operator ρ_E.

We must ensure that the ancilla C is separable at all times. As the terms of Eq. (1.35) involving ρ_0, ρ_a and ρ_b are already separable in the $C|AB$ partition, all that is required is to ensure that the term involving ρ_E is similarly separable. We make use of the following theorem on the robustness of entanglement:

Theorem 1.6 (Entanglement Robustness Bound)
Given a state ρ on an N-partite Hilbert space of total dimension d, the state

$$\frac{1}{1+s}\left(\rho + s\frac{\mathbb{1}}{d}\right) \tag{1.36a}$$

with

$$s \geq \left(1+\frac{d}{2}\right)^{N-1} - 1 \tag{1.36b}$$

is completely separable.

Proof See Vidal and Tarrach [1999, Eq. (C8)]. □

This allows us to derive a sufficient condition for the state ρ_E of Eq. (1.35) to be made separable in the $C|AB$ partition when it is mixed with the identity. We have $N = 2$ and $d = 12$, so from Theorem 1.6 we require $s \geq 6$. From Eq. (1.35), we have

$$s = \frac{\alpha^2}{\left(\frac{3}{2} + 6\alpha\right)\kappa\epsilon}. \tag{1.37}$$

This implies the following result:

Result 1.7 (Ancilla Separability Condition)
The inequality

$$\epsilon \leq \frac{\alpha^2}{(1+4\alpha)9\kappa} \tag{1.38}$$

is a sufficient condition for the time-evolved state $U\rho U^\dagger$ of Eq. (1.35) *to be separable in the $C|AB$ partition.*

Finally, we must ensure that the reduced state of the qubits A and B, namely $\mathrm{Tr}_C[U\rho U^\dagger]$, contains a finite amount of entanglement at time $t = 2\pi/3\epsilon^2$. The following theorem relates the entanglement robustness of a two-qubit state to its entanglement fidelity.

Theorem 1.8 (Entanglement Robustness and Fidelity)
Given a two-qubit entangled state ρ and any other two-qubit state ρ_s, the state

$$(1-s)\rho + s\rho_s \tag{1.39}$$

is entangled provided

$$s < 1 - \tfrac{1}{2F(\rho)}, \tag{1.40}$$

where the entanglement fidelity $F(\rho) = \max_\phi \langle\phi|\,\rho\,|\phi\rangle$, with the maximization taken over all maximally entangled states $|\phi\rangle$.

Proof See Verstraete and Verschelde [2003]. □

Define ρ_{AB} to be the reduced state of the qubits excluding the term involving ρ_E in Eq. (1.35). With ρ_{AB} playing the role of ρ and ρ_E playing that of ρ_s in Theorem 1.8, we have $s = \frac{3/2+6\alpha}{(1+\alpha^2)}$ from Eq. (1.35), and the sufficient condition for ρ_{AB} to be entangled becomes an inequality relating ϵ and α:

$$\frac{\left(\frac{3}{2}+6\alpha\right)\kappa\epsilon}{(1+\alpha)^2} < 1 - \frac{1}{2F(\rho_{AB})}. \tag{1.41}$$

If we substitute any lower bound on the entanglement fidelity F in Eq. (1.41), we obtain a tighter bound on α that is certainly still sufficient. Calculating the fidelity with the particular maximally entangled state $|\phi\rangle = \frac{1}{4}(|++\rangle + i\,|+-\rangle + i\,|-+\rangle + |--\rangle)$, we can bound

$$F(\rho_{AB}) \geq \frac{1}{N}\left(\left(1-\frac{3}{2}\kappa\epsilon\right)\frac{1+\frac{\sqrt{3}}{2}}{2} + 2\alpha\left(1-3\kappa\epsilon\right)\frac{1+\frac{\sqrt{3}}{2}}{4} + \frac{\alpha^2}{4}\right), \tag{1.42a}$$

$$N = 1 - \frac{3}{2}\kappa\epsilon + 2(1-3\kappa\epsilon)\alpha + \alpha^2. \tag{1.42b}$$

Equation (1.41), together with Eq. (1.42), then gives the following result:

Result 1.9 (Qubit Entanglement Condition)
The inequality

$$\frac{\left(\frac{3}{2}+6\alpha\right)\kappa\epsilon}{(1+\alpha)^2} < 1 - \frac{1}{2F}, \tag{1.43}$$

with

$$F = \frac{1}{N}\left(\left(1-\frac{3}{2}\kappa\epsilon\right)\frac{1+\frac{\sqrt{3}}{2}}{2} + 2\alpha\left(1-3\kappa\epsilon\right)\frac{1+\frac{\sqrt{3}}{2}}{4} + \frac{\alpha^2}{4}\right), \tag{1.44a}$$

$$N = 1 - \frac{3}{2}\kappa\epsilon + 2(1-3\kappa\epsilon)\alpha + \alpha^2 \tag{1.44b}$$

is a sufficient condition for the state $\mathrm{Tr}_C[U\rho U^\dagger]$ *of qubits A and B to be entangled.*

1.2.4 Creating Entanglement Without Entangling

Results 1.7 and 1.9, together with the bounds in Lemmas 1.3 and 1.4, give sufficient conditions for the ancilla to remain separable and the qubits to become entangled. Thus, if both these conditions can be satisfied, we can achieve the desired result.

The trace norm involved in these results is inconvenient to calculate, so we will instead bound it by the Hilbert-Schmidt norm.

Lemma 1.10 (Trace and Hilbert-Schmidt Norms)
For any $d \times d$ matrix M,

$$\|M\|_{\mathrm{tr}} \leq \sqrt{d}\, \|M\|_{\mathrm{HS}} \tag{1.45}$$

where $\|M\|_{\mathrm{tr}}$ and $\|M\|_{\mathrm{HS}}$ denote the trace and Hilbert-Schmidt norms of M, which can be expressed in terms of the singular values σ_i of M:

$$\|M\|_{\mathrm{tr}} = \sum_i \sigma_i \tag{1.46}$$

$$\|M\|_{\mathrm{HS}} = \sum_{i,j} M_{ij} = \sqrt{\sum_i \sigma_i^2}. \tag{1.47}$$

Proof Ordering the singular values $\sigma_1 \geq \sigma_i \geq \cdots \geq \sigma_d$,

$$\|M\|_{\mathrm{tr}}^2 = \sum_{i,j} \sigma_i \sigma_j = d \sum_i \sigma_i^2 - (d-1) \sum_i \sigma_i^2 + 2 \sum_{i<j} \sigma_i \sigma_j \tag{1.48a}$$

$$= d \sum_i \sigma_i^2 - \sum_{i<j} (\sigma_i - \sigma_j)^2 \leq d \sum_i \sigma_i^2 \tag{1.48b}$$

$$= d\, \|M\|_{\mathrm{HS}}^2 . \tag{1.48c}$$

□

To summarize, gathering all the results together, and using Lemma 1.10 to bound trace norms by Hilbert-Schmidt norms, we require both that

$$\epsilon \leq \frac{\alpha^2}{(1+4\alpha)9\kappa} \tag{1.49a}$$

$$\kappa = \frac{4}{\epsilon}(\mathcal{N}_1 + \mathcal{N}_2)(1 + \mathcal{N}_1 + \mathcal{N}_2) \tag{1.49b}$$

$$\mathcal{N}_1 = \frac{\sqrt{3}\, \|X\|_{\mathrm{HS}}}{(1 - \|\mathbb{1} - X\|_{\mathrm{HS}})^2} \left\| H^{\pm} X - X D \right\|_{\mathrm{HS}} t \tag{1.49c}$$

$$\mathcal{N}_2 = \frac{2\sqrt{3}\,(d_3^2 - d_1^2)\, \|XD - DX\|_{\mathrm{HS}} + 2\sqrt{3}\,(d_3 - d_1)\, \|XD^2 - D^2 X\|_{\mathrm{HS}}}{|\Delta|\,(1 - \|\mathbb{1} - X\|_{\mathrm{HS}})^2} \tag{1.49d}$$

$$\Delta = d_1^2(d_2 - d_3) + d_2^2(d_3 - d_1) + d_3^2(d_1 - d_2) \tag{1.49e}$$

to ensure that the ancilla C remains separable at all times (the factors of $\sqrt{3}$

arise from Lemma 1.10), and that at the same time

$$\frac{\left(\frac{3}{2}+6\alpha\right)\kappa\epsilon}{(1+\alpha)^2} < 1 - \frac{1}{2F} \tag{1.50a}$$

$$F = \frac{1}{N}\left(\left(1-\frac{3}{2}\kappa\epsilon\right)\frac{1+\frac{\sqrt{3}}{2}}{2} + 2\alpha\left(1-3\kappa\epsilon\right)\frac{1+\frac{\sqrt{3}}{2}}{4} + \frac{\alpha^2}{4}\right) \tag{1.50b}$$

$$N = 1 - \frac{3}{2}\kappa\epsilon + 2(1-3\kappa\epsilon)\alpha + \alpha^2 \tag{1.50c}$$

to ensure that qubits A and B become strongly entangled by time $t = 2\pi/\epsilon^2$. The matrices X and D are given by Eq. (1.12b) and Eq. (1.12a), so the matrix norms can be calculated explicitly.

We could solve the above inequalities simultaneously. But we are not particularly interested in calculating the optimum value of ϵ and α (the bounds are in any case not tight), only in showing that it is possible to satisfy both conditions simultaneously for *some* ϵ and α. To do this, we can start by assuming an upper bound on ϵ, use this to obtain a lower bound on α from Eq. (1.49a), and verify that these are consistent with Eq. (1.50a).

For example, assuming $\epsilon \leq 10^{-5}$ gives $\alpha \geq 1.54 \times 10^{-2}$ from Eq. (1.49a). Substituting these values in Eq. (1.50a) gives $3.02 \times 10^{-2} \leq 1.99 \times 10^{-1}$, so these ϵ and α are clearly consistent with both conditions. This demonstrates that the sufficient conditions for separability of the ancilla (Result 1.7) and entanglement of the qubits (Result 1.9) can be satisfied simultaneously, which implies that it is possible to create entanglement between two particles interacting via a common ancilla, without ever entangling that ancilla.

This effect forces us to abandon the notion of entanglement *flowing* from A to B via C. Instead, we have shown entanglement can "tunnel" through the particle mediating the interactions. (However, in Chapter 3, we will show that there is a sense in which the intuitively appealing idea of entanglement flow can be recovered.)

1.3 Discrete Case

We now turn to the discrete case, in which Alice allows her particle to interact with a messenger particle without the two becoming entangled, sends that unentangled messenger to Bob, who allows it to interact with his particle, thereby entangling his particle with that of Alice (Fig. 1.2).

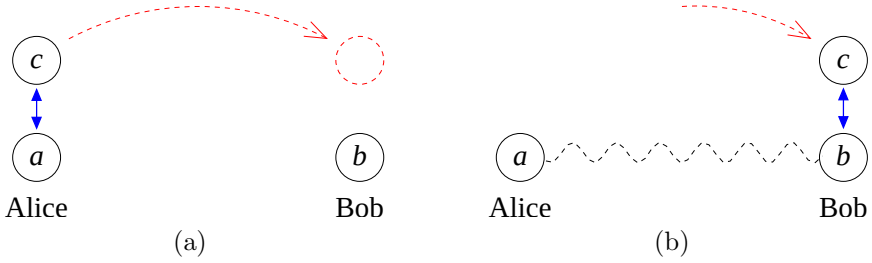

Figure 1.2: Alice and Bob each have a particle they wish to entangle with the other. (a) Alice interacts a "messenger" particle C with her particle A, sends C to Bob, (b) who interacts C with his particle B. At the end they share some entanglement. Surprisingly, C does not have to become entangled with A and B.

1.3.1 Entanglement Properties During the Evolution

As noted in the introduction, the continuous evolution described in the previous section can be approximated arbitrarily accurately by a sequence of discrete quantum operations, for example via the Trotter decomposition:

$$e^{-it(H_{AC}+H_{CB})} = \lim_{n\to\infty} (e^{-itH_{AC}/n} e^{-itH_{CB}/n})^n, \tag{1.51}$$

and these can be implemented by sending a messenger particle back and forth between Alice and Bob. As the results of Section 1.2 guarantee that the state remains at least some distance $\delta > 0$ away from any state in which the messenger particle C is entangled, it is sufficient to approximate the evolution to within an accuracy δ, which requires sending the messenger particle between Alice and Bob a finite (though large) number of times (of order $1/\delta$). This leads directly to a discrete scheme, albeit one in which the messenger particle must bounce back and forth many times.

What are the entanglement properties of the sequence of states produced during this discretized evolution? Let us assume without loss of generality that the first operation that changes the entanglement properties of the tripartite system is one that operates on particles A and C. The system starts in a completely separable state (Fig. 1.3a), so such an operation can not entangle A with B. However, we have shown in Section 1.2 that particle C *never* becomes entangled with A or B during the evolution. Therefore, A can not become entangled with either B or C.

Given that A and B do eventually become entangled, the only remaining possibility is that the operation on A and C entangles A with BC when the latter is considered as a single system, even though A remains separable from both B and C individually. Thus the state created by this operation must have the strange property that it is separable in the $C|AB$ and $B|AC$

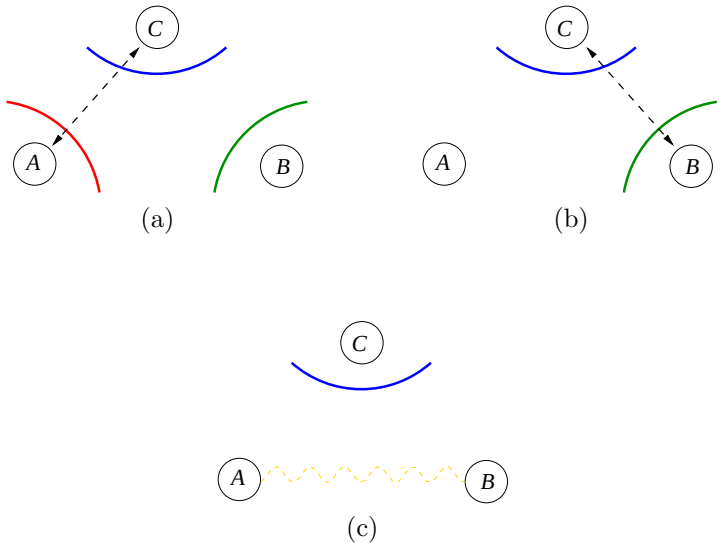

Figure 1.3: Entanglement properties of the sequence of states produced by the discretized version of the evolution described in Section 1.2. (a) The system starts in a completely separable state, so is separable across every bipartite partition. (a)–(b) An operation on particles A and C entangles A with BC, where BC is temporarily considered as a single system. (b) The state is no longer separable across the $A|BC$ partition (red line). Note that the messenger remains unentangled by separability across the $C|AB$ partition (blue line). (b)–(c) An operation on particles B and C transfers the entanglement created across the $A|BC$ partition in the previous step to B. (c) Particles A and B are now entangled.

partitions, yet is *not* separable in the third partition $A|BC$ (Fig. 1.3b).

Do such states exist? For pure states, we can easily demonstrate that they do not. Separability in the $C|AB$ partition implies it is possible to write the state as $|\psi\rangle = |\psi_{AB}\rangle\,|c\rangle$. But separability in the $B|AC$ partition implies separability when tracing over C, thus $|\psi_{AB}\rangle = |a\rangle\,|b\rangle$, and the overall state must be completely separable: $|\psi\rangle = |a\rangle\,|b\rangle\,|c\rangle$. This provides another proof that the messenger particle necessarily becomes entangled when distributing entanglement between Alice and Bob if the overall state is pure (see Section 1.1).

However, the overall state of the system described in Section 1.2 is mixed. Dür et al. [1999b] have shown that the the entanglement properties across bipartite partitions of mixed, tripartite quantum states are independent; that is, *mixed* states with the properties we require *do* exist. Following their terminology, we will call such states *biseparable*.

The final change in entanglement properties during the evolution must involve an operation on particles B and C, which transfers the entanglement created across the $A|BC$ partition in the previous step to particle B, without entangling the messenger C (Fig. 1.3c).

It is worth noting that there are two surprising aspects to this evolution: firstly, that biseparable mixed states exist at all and, secondly, that we can create a biseparable state from a separable state using an operation that acts on only two of the particles. Although we know from the discretization of the continuous-time result that this entanglement evolution is possible if the messenger particle is allowed to bounce back and forth many times, in the next section we give an explicit example of a discrete protocol that only requires Alice to send the messenger to Bob *once*.

1.3.2 Explicit Example

We consider an example in which all particles are qubits. In the first step, Alice and Bob prepare the following completely separable (though classically correlated) state, where Alice has the messenger particle in her laboratory:

$$\rho_{abc} = \frac{1}{6}\sum_{k=0}^{3} |\Psi_k, \Psi_{-k}, 0\rangle\langle\Psi_k, \Psi_{-k}, 0| + \sum_{i=0}^{1} \frac{1}{6}\,|i, i, 1\rangle\langle i, i, 1| \tag{1.52}$$

where $|\Psi_k\rangle = \frac{1}{\sqrt{2}}\left(|0\rangle + e^{ik\pi/2}\,|1\rangle\right)$. This is manifestly separable as it is a convex combination of separable pure states, so it can be prepared by local operations and classical communication.

In the second step, Alice applies a CNOT operation on particles A and C

with A the control qubit, producing the state

$$\sigma_{abc} = \frac{1}{3}\left|\Psi_{GHZ}\right\rangle\left\langle\Psi_{GHZ}\right| + \sum_{i,j,k=0}^{1} \beta_{ijk}\mathcal{P}_{ijk} \tag{1.53}$$

where $|\Psi_{GHZ}\rangle = \frac{1}{\sqrt{2}}(|000\rangle + |111\rangle)$, $\mathcal{P}_{ijk} = |ijk\rangle\langle ijk|$, and all β's are 0 apart from $\beta_{001} = \beta_{010} = \beta_{101} = \beta_{110} = \frac{1}{6}$. An operation on A and C can not change the entanglement across the $AC|B$ partition, so B remains separable. However, the state is symmetric under interchange of particles B and C, so must also be separable across the $C|AB$ partition, i.e. the messenger particle has not become entangled.

What about entanglement across the $A|BC$ partition? The state σ_{ABC} has precisely the form of the family of states considered in Dür et al. [1999b] and, by the results described therein, is a biseparable state with entanglement across the $A|BC$ partition.

In the third step, Alice sends the messenger particle to Bob, who applies another CNOT operation to it and his particle B, with B the control qubit, producing the state

$$\tau_{ABC} = \frac{1}{3}\left|\phi^+\right\rangle\left\langle\phi^+\right|_{AB} \otimes |0\rangle\langle 0|_C + \frac{2}{3}\mathbb{1}_{AB} \otimes |1\rangle\langle 1|_C\,, \tag{1.54}$$

where $|\phi^+\rangle = \frac{1}{\sqrt{2}}(|00\rangle + |11\rangle)$ is maximally entangled. This is manifestly still separable across the $C|AB$ partition, but the state now contains entanglement between A and B.

Bob can extract this entanglement in a number of ways. Measuring C in the computational basis, he can extract a maximally entangled state of AB with probability $\frac{1}{3}$. Alternatively, if a deterministic effect is required, he can apply a local completely positive map to particles B and C, defined by

$$\mathcal{E}_{BC}(\rho) = \sum_j E^{(k)}_{BC}\rho\, E^{(k)\dagger}_{BC} \tag{1.55}$$

with Kraus operators

$$E^{(1)}_{bc} = \mathbb{1}_B \otimes |0\rangle\langle 0|_C\,, \tag{1.56a}$$

$$E^{(2)}_{bc} = |0\rangle\langle 0|_B \otimes |1\rangle\langle 1|_C\,, \tag{1.56b}$$

$$E^{(3)}_{bc} = |0\rangle\langle 1|_B \otimes |1\rangle\langle 1|_C\,. \tag{1.56c}$$

These satisfy

$$\sum_j E^{(k)\dagger}E^{(k)} = \mathbb{1}. \tag{1.57}$$

so the map is trace-preserving. Throwing away the ancilla leaves the state

$$\rho_{AB} = \mathrm{Tr}_C \left[\mathcal{E}_{BC}(\tau_{ABC})\right] = \frac{1}{3}\left|\phi^+\right\rangle\!\left\langle\phi^+\right| + \frac{1}{3}\left|00\right\rangle\!\left\langle 00\right| + \frac{1}{3}\left|10\right\rangle\!\left\langle 10\right|. \tag{1.58}$$

The partial transpose of this density matrix is negative, so the state must be (distillable) entangled [Horodecki et al., 1997, Peres, 1996]. Thus Alice and Bob have entangled their particles by a single exchange of an unentangled messenger particle; entanglement can be transferred by a particle that does not carry the entanglement with it. We are forced to abandon the notion of entanglement being sent *through* a quantum channel. For example, the Ekert protocol for quantum cryptography could be implemented by sending separable states [Ekert, 1991].

1.4 Quantum Channels

As a final result in the same spirit as those presented in the previous sections, we can use the Jamiolkowski isomorphism [Cirac et al., 2001, Jamiołkowski, 1972] between states and channels (trace-preserving completely positive maps) to derive an analogous result for quantum channels.

Consider a map $\mathcal{E}$ acting on states in the Hilbert space $\mathcal{H}_A$. If we act on one half of an (unnormalized) maximally entangled state in a larger space $\mathcal{H}_A \otimes \mathcal{H}_B$, we obtain a state

$$\sigma_{AB} = (\mathcal{E} \otimes \mathcal{I}) \left(\sum_{ij} \left|ii\right\rangle\!\left\langle jj\right|_{AB} \right) = \sum_{ij} \mathcal{E}(\left|i\right\rangle\!\left\langle j\right|_A) \otimes \left|i\right\rangle\!\left\langle j\right|_B, \tag{1.59}$$

which is dual to $\mathcal{E}$, in the sense that the original map can be recovered from it. If

$$\rho_B = \sum_{kl} \rho_{kl} \left|k\right\rangle_{A_2} \left|l\right\rangle_{B_2} \tag{1.60}$$

is an arbitrary state in $\mathcal{H}_B$, then

$$\mathrm{Tr}_B \left[\sigma_{AB}(\mathbb{1}_A \otimes \rho_B^T)\right] = \sum_{ij} \left(\mathcal{E}(\left|i\right\rangle\!\left\langle j\right|_A)\ \mathrm{Tr}\Big[\sum_{kl} \left|i\right\rangle\!\left\langle j\right|_B \rho_{lk} \left|k\right\rangle\!\left\langle l\right|_B\Big] \right) \tag{1.61a}$$

$$= \sum_{ij} \mathcal{E}(\rho_{ij} \left|i\right\rangle\!\left\langle j\right|_A) \tag{1.61b}$$

$$= \mathcal{E}(\rho). \tag{1.61c}$$

This is the Jamiolkowski isomorphism.

A *separable* map on multipartite states is defined to be one whose Kraus operators are tensor products, e.g. in the bipartite case

$$\mathcal{E}(\rho_{AB}) = \sum_k (A^{(k)} \otimes B^{(k)}) \rho (A^{(k)} \otimes B^{(k)})^\dagger. \tag{1.62}$$

The definition is motivated by the fact that these maps can not create entanglement when they act on a separable state. The Jamiolkowski isomorphism can easily be extended to maps acting on multipartite states [Cirac et al., 2001], with the dual state obtained by acting on each half of a number of maximally entangled bipartite states (one for each party). A map that is separable in some partition gives rise to a dual state that is itself separable across the corresponding partition.

Motivated by the state of Eq. (1.53) created in the second step of the discrete protocol (Section 1.3), consider the map $\mathcal{E}_1$ acting on three qubits A, B and C, with Kraus operators

$$\begin{aligned} A_1 &= |000\rangle\langle 000| + |111\rangle\langle 111|\,, \\ A_2 = |001\rangle\langle 001|\,, &\quad A_3 = |010\rangle\langle 010|\,, \quad A_4 = |101\rangle\langle 101|\,, \\ A_5 = |110\rangle\langle 110|\,, &\quad A_6 = |000\rangle\langle 011|\,, \quad A_7 = |111\rangle\langle 100|\,. \end{aligned} \tag{1.63}$$

Its dual state, defined on a six qubit space, is given by

$$\sigma_{A_1A_2B_1B_2C_1C_2} = |\Phi_{GHZ}\rangle\langle\Phi_{GHZ}| + \sum_{i,j,k=0}^{1} \beta_{ijk} \mathcal{P}_{iijjkk} + \mathcal{P}_{000101} + \mathcal{P}_{111010}, \tag{1.64}$$

where $|\Phi_{GHZ}\rangle = |000000\rangle + |111111\rangle$, $\mathcal{P}_{iijjkk} = |iijjkk\rangle\langle iijjkk|$, and all β's are 0 apart from $\beta_{001} = \beta_{010} = \beta_{101} = \beta_{110} = 1$.

Without the last two terms, this state would be equivalent to the three qubit state of Eq. (1.53), where the three qubits of Eq. (1.53) are identified with the $\{|00\rangle, |11\rangle\}$ subspaces of the three pairs of qubits in Eq. (1.64). Mixing with the separable states of the last two terms can not increase the entanglement of the state across any partition. Thus, like the state of Eq. (1.53), this dual state is separable in the $B_1B_2|A_1A_2C_1C_2$ and $C_1C_2|A_1A_2B_1B_2$ partitions. Since the separability properties of the dual state are equivalent to the separability properties of the map, $\mathcal{E}_1$ can not create entanglement across the $B|AC$ or $C|AB$ partitions, so it clearly can not entangle qubits A and B.

Define a second map $\mathcal{E}_2$ with the same Kraus operators Eq. (1.63), which we label $B^{(k)}$, but with the roles of A and B reversed. Clearly, $\mathcal{E}_2$ is separable across the $A|BC$ and $C|AB$ partitions, thus like $\mathcal{E}_1$ can not entangle A and B. The map defined by the composition of the two, $\mathcal{E} = \mathcal{E}_2 \circ \mathcal{E}_1$, has Kraus

operators $C^{(k)} = A^{(i)}B^{(j)}$:

$$\begin{aligned} C_1 &= |000\rangle\langle 000| + |111\rangle\langle 111|\,, \\ C_2 = |001\rangle\langle 001|\,, &\quad C_3 = |000\rangle\langle 010|\,, \quad C_4 = |000\rangle\langle 011|\,, \\ C_5 = |111\rangle\langle 100|\,, &\quad C_6 = |111\rangle\langle 101|\,, \quad C_7 = |110\rangle\langle 110|\,. \end{aligned} \tag{1.65}$$

Its dual state is given by

$$\begin{aligned} \tau_{A_1A_2B_1B_2C_1C_2} = &|\Phi_{GHZ}\rangle\langle\Phi_{GHZ}| + \mathcal{P}_{000011} + \mathcal{P}_{000100} + \mathcal{P}_{000101} \\ &+ \mathcal{P}_{111010} + \mathcal{P}_{111011} + \mathcal{P}_{111100}. \end{aligned} \tag{1.66}$$

As both $\mathcal{E}_1$ and $\mathcal{E}_2$ are separable in the $C|AB$ partition, $\mathcal{E}$ is also separable across that partition. However, the dual state τ is entangled in the $A_1A_2|B_1B_2C_1C_2$ and $B_1B_2|A_1A_2C_1C_2$ partitions, as can easily be verified using the partial transposition criterion [Horodecki et al., 1997, Peres, 1996]. Thus $\mathcal{E}$ *is* entangling across the $A|BC$ and $B|AC$ partitions, which implies that it can entangle qubits A and B when it operates on a separable state. For example, applying $\mathcal{E}$ to the product state $|++\!+\rangle$ produces

$$\mathcal{E}(|+\!+\!+\rangle) = \frac{1}{4}\,|\Psi_{GHZ}\rangle\langle\Psi_{GHZ}| + \frac{1}{4}(\mathcal{P}_{000} + \mathcal{P}_{111}) + \frac{1}{8}(\mathcal{P}_{001} + \mathcal{P}_{110}). \tag{1.67}$$

Performing a local measurement on C in the $|+\rangle$, $|-\rangle$ basis produces the following states on the qubits A and B conditioned on the outcome ($+$ or $-$) of the measurement:

$$\frac{1}{4}(|00\rangle \pm |11\rangle)(\langle 00| \pm \langle 11|) + \frac{3}{8}(|00\rangle\langle 00| + |11\rangle\langle 11|), \tag{1.68}$$

both of which are entangled. (Again, this can be verified using the partial transposition criterion.)

Thus we have demonstrated that the composition of two separable maps can be entangling. More precisely, two non-local maps, each of which can not create any entanglement between two qubits (or between those qubits and an ancilla) when acting individually, *can* entangle those qubits when composed.

Chapter 2

Entanglement Bottlenecks

When two particles are interacting indirectly via a third, mediating particle, we might expect the mediating particle to act as a "bottleneck" for entanglement generation: if the middle particle is only slightly entangled, the rate of entanglement transfer to the end particles should be slow. Of course, we have seen in Chapter 1 that this is not the case: entanglement can be created without entangling the mediating particle at all!

Does this force us to abandon any concept of an "entanglement bottleneck"? The results of Chapter 1 relied on the entanglement properties of mixed states. On the other hand, we showed in Section 1.1 of the same chapter that the effect is impossible if the overall state of the system is pure. In the pure-state case, unless the middle particle is entangled, no entanglement can be created between the other two. This suggests that, for *pure states*, the entanglement of the mediating particle may indeed limit the rate of entanglement generation.

In this chapter, we put the concept of an entanglement bottleneck on a rigorous footing. In Section 2.1, we will derive a quantitative relationship between the rate of entanglement generation between the end qubits of a three-qubit chain and the entanglement of the middle qubit. Then, in Section 2.2, we will derive a corresponding result for the more difficult case of tripartite systems of arbitrary dimension.

2.1 Three-Qubit Systems

2.1.1 Motivation

The simplest conceivable system with non-trivial entanglement dynamics consists of two interacting qubits that start (and remain) in a pure state.

This system was investigated in Dür et al. [2001], leading to an equation for the entanglement rate:

$$\frac{\mathrm{d}E}{\mathrm{d}t} = 2\lambda_1\lambda_2 E'(\lambda_1) \langle\varphi_1, \chi_1| H |\varphi_2, \chi_2\rangle \tag{2.1}$$

where E is any entanglement measure, $|\psi\rangle_{AB} = \sum_i \sqrt{\lambda_i} |\varphi_i\rangle_A |\chi_i\rangle_B$ is the Schmidt decomposition of the system state, the prime indicates the derivative with respect to a Schmidt coefficient, and H is the system Hamiltonian.

What is interesting about this equation is that it separates cleanly into two independent parts: the first, $2\lambda_1\lambda_2 E'(\lambda_1)$, is a function of the entanglement of the system, and the second, $\langle\varphi_1, \chi_1| H |\varphi_2, \chi_2\rangle$, is a function of the interaction Hamiltonian. These provide answers to two kinds of question about the entanglement dynamics: how does the entanglement rate depend on the entanglement already in the system? And what is the entangling power of different Hamiltonians?

In moving to the more complex three-qubit chain, we will concentrate on the first of these questions. Of course, there are many more ways to measure entanglement in a tripartite system: entanglement across all three bipartite partitions, entanglement of any of the bipartite reduced states, genuine tripartite entanglement. We will be interested specifically in the relationship between two of them: the entanglement rate of the two end qubits and the entanglement of the middle one, as this will lead to the concept of an entanglement bottleneck.

2.1.2 Bottleneck Inequality

Although we will restrict the overall state of our three-qubit system to be a pure state, the reduced state of any two of the qubits can, and typically will, be mixed. It is not known in general how to quantify the entanglement of mixed states. However, in the special case of two qubits, it can be quantified by the *concurrence* [Wootters, 1998]. This is an entanglement measure in its own right, but its significance lies in its equivalence to one of the important, operationally defined entanglement measures: the *entanglement of formation*, which measures how many maximally entangled states are required to create the given state using only local operations and classical communication.

Definition 2.1 (Concurrence)

$$C(\rho_{AB}) = \max\{0, \mu_1 - \mu_2 - \mu_3 - \mu_4\} \tag{2.2}$$

where the μ's are the square-roots of the eigenvalues of $\rho_{AB}\tilde{\rho}_{AB}$, arranged in decreasing order: $\mu_1 \geq \mu_2 \geq \mu_3 \geq \mu_4$, and the tilde denotes the spin-flip

operation:

$$\tilde{\rho}_{AB} = \Sigma \rho_{AB}^{*} \Sigma, \quad \Sigma = \sigma^{y} \otimes \sigma^{y} \tag{2.3}$$

with the complex conjugation taken in the basis in which the Pauli matrix σ^{y} is defined.

Note that for pure states, the concurrence reduces to

$$C(|\psi\rangle_{AB}) = \lambda_1 \lambda_2 \tag{2.4}$$

where the λ's are the state's Schmidt coefficients.

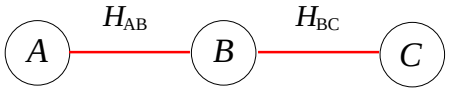

Figure 2.1: A tripartite chain, showing the labels and interaction Hamiltonians referred to in the main text.

Labelling the qubits A, B and C with B the middle qubit, the interactions of the three-qubit chain are described by a Hamiltonian of the form $H = H_{AB} + H_{BC}$ (Fig. 2.1). We write the state of the system in its Schmidt decomposition with respect to the partition $B|AC$:

$$|\psi\rangle_{ABC} = \sum_{i=1}^{2} \lambda_i \, |\varphi_i\rangle_{AC} \, |\chi_i\rangle_{B} \, . \tag{2.5}$$

It is useful to reshape the state vector $|\psi\rangle_{ABC}$ so that the state is represented by the 4×2 matrix

$$X = \sum_{i=1}^{2} \lambda_i \, |\varphi_i\rangle \langle\chi_i| \equiv \begin{pmatrix} \vdots & \vdots \\ \lambda_1 \, |\varphi_1\rangle & \lambda_2 \, |\varphi_2\rangle \\ \vdots & \vdots \end{pmatrix}, \tag{2.6}$$

whose columns are given by the Schmidt vectors of AC weighted by the corresponding Schmidt coefficients. The reduced state of AC is then given by $\rho_{AC} = XX^{\dagger}$.

We can therefore write

$$\rho_{AC} \tilde{\rho}_{AC} = XX^{\dagger} \Sigma X^{*} X^{T} \Sigma = X (X^{T} \Sigma X)^{\dagger} (X^{T} \Sigma X) X^{-1}. \tag{2.7}$$

Since eigenvalues are invariant under similarity transformations, the concurrence of ρ_{AB} can be obtained from the eigenvalues of $(X^T\Sigma X)^\dagger(X^T\Sigma X)$ or, equivalently, from the singular values ς_i of the 2×2 matrix $A = X^T\Sigma X$: $C_{AC} = \varsigma_1 - \varsigma_2$, with the singular values ordered such that $\varsigma_1 \geq \varsigma_2$ [Audenaert et al., 2001]. Since $\operatorname{Tr} A^\dagger A = \varsigma_1^2 + \varsigma_2^2$ and $|\det A| = \varsigma_1\varsigma_2$, we can also write this as

$$C_{AC} = \sqrt{\operatorname{Tr} A^\dagger A - 2|\det A|}. \tag{2.8}$$

To calculate the time-derivative of the concurrence, we must therefore calculate the time-derivatives of $\operatorname{Tr} A^\dagger A$ and $|\det A|$. From the definition of A,

$$\dot{A} = \dot{X}^T\Sigma X + X^T\Sigma\,\dot{X}, \tag{2.9}$$

so that

$$\frac{\mathrm{d}(\operatorname{Tr} A^\dagger A)}{\mathrm{d}t} = \operatorname{Tr}[A^\dagger\dot{A} + \dot{A}^\dagger A] = 4\operatorname{Re}\left(\operatorname{Tr}\left[\Sigma\rho^*\Sigma\,\dot{X}X^\dagger\right]\right), \tag{2.10}$$

where we have substituted Eq. (2.9) and used the cyclic property of the trace. Since A is a 2×2 matrix, we have

$$\det A = \frac{1}{2}\operatorname{Tr}[A\sigma^y A^T\sigma^y]. \tag{2.11}$$

(This can easily be verified by writing out the matrices in component form.) Thus

$$\frac{\mathrm{d}(\det A)}{\mathrm{d}t} = \frac{1}{2}\operatorname{Tr}[\dot{A}\sigma^y A^T\sigma^y + A\sigma^y\dot{A}^T\sigma^y] = 4\operatorname{Tr}\left[X\sigma^y X^T\Sigma X\sigma^y\dot{X}^T\Sigma\right], \tag{2.12}$$

again using Eq. (2.9) and cyclically permuting inside the trace.

The two-qubit Hamiltonian H_{AB} has a product decomposition

$$H_{AB} = k\mathbb{1}_{AB} + \sum_{i=x,y,z}(s_i\sigma_A^i\otimes\mathbb{1}_B + t_i\mathbb{1}_A\otimes\sigma_B^i) + \sum_{i,j=x,y,z} a_{ij}\,\sigma_A^i\otimes\sigma_B^j \tag{2.13}$$

in terms of Pauli matrices defined in the $|\chi_1\rangle, |\chi_2\rangle$ basis given by the states in the Schmidt decomposition of Eq. (2.5), where the coefficients k, s_i, t_i and a_{ij} are real. However, the terms with coefficients k, s_i and t_i are local, and can not change any entanglement properties of the system, so we will neglect them from now on. H_{BC} can be similarly decomposed, with real coefficients c_{ij}, so that the non-local part of the overall Hamiltonian is given by

$$H = \sum_{i,j=x,y,z}(a_{ij}\sigma_A^i\otimes\sigma_B^j\otimes\mathbb{1}_C + c_{ij}\mathbb{1}_A\otimes\sigma_B^i\otimes\sigma_C^j). \tag{2.14}$$

As X is a reshaped version of the state vector, the Schrödinger equation describing the evolution of the system state $|\psi\rangle_{ABC}$ translates into an equation for the evolution of X. For any Hamiltonian of the form $H_{AC} \otimes H_B$, we have

$$\frac{\mathrm{d}X}{\mathrm{d}t} = -iH_{AC}XH_B^T, \tag{2.15}$$

Since the system Hamiltonian written in the product decomposition of Eq. (2.14) is a sum of terms of this form, we obtain

$$\dot{X} = -i \sum_{i,j=x,y,z} (a_{ij}\sigma^i \otimes \mathbb{1} + c_{ij}\mathbb{1} \otimes \sigma^i)X{\sigma^j}^T. \tag{2.16}$$

We can use this in Eq. (2.8) for the concurrence, along with Eq. (2.10) and Eq. (2.12) for the derivatives of the trace and determinant, to obtain an expression for its time-derivative:

$$\frac{\mathrm{d}C_{AC}^2}{\mathrm{d}t} = h(H, |\psi\rangle)\lambda_1\lambda_2. \tag{2.17}$$

The factor $h(H, |\psi\rangle)$ depends on both the interactions and the system state, and is a rather complicated sum over terms involving a_{ij} and c_{ij}. Define

$$s_{ij}^k = \langle\tilde{\varphi}_i| \, \sigma^k \otimes \mathbb{1} \, |\varphi_j\rangle \tag{2.18a}$$
$$t_{ij}^k = \langle\tilde{\varphi}_i| \, \mathbb{1} \otimes \sigma^k \, |\varphi_j\rangle \tag{2.18b}$$
$$o_{ij} = \langle\varphi_i|\tilde{\varphi}_j\rangle \tag{2.18c}$$
$$h_{ix}^a = -i(\lambda_1^2 s_{12}^i o_{11} + \lambda_2^2 s_{21}^i o_{22}) \tag{2.18d}$$
$$h_{iy}^a = \lambda_2^2 s_{21}^i o_{22} - \lambda_1^2 s_{12}^i o_{11} \tag{2.18e}$$
$$h_{iz}^a = -i\lambda_1\lambda_2(s_{21}^i o_{12} - s_{12}^i o_{21}), \tag{2.18f}$$

and define h_{ij}^c similarly to h_{ij}^a, but with the s_{ij}'s replaced by t_{ij}'s. Note that $s_{ii}^k = t_{ii}^k = 0$. The tildes again denote the spin-flip operation: $|\tilde{\varphi}\rangle = \sigma^y \otimes \sigma^y \, |\varphi^*\rangle$, and the states $|\varphi_i\rangle$ are those appearing in the Schmidt decomposition of Eq. (2.5). Then

$$h(H, |\psi\rangle) = 4 \, \mathrm{Re}\left(\sum_{i,j=x,y,z} a_{ij}h_{ij}^a + c_{ij}h_{ij}^c \right) + 4 \left| \sum_{i,j=x,y,z} a_{ij}h_{ij}^a + c_{ij}h_{ij}^c \right|. \tag{2.19}$$

However, as we are primarily interested in the dependence on entanglement (i.e. the dependence on the Schmidt coefficients), we can bound the magnitudes of s_{ij}^k, t_{ij}^k and o_{ij} by 1, and assume all terms sum in phase, giving

$$h(H, |\psi\rangle) \leq \, 8 \sum_{ij} (|a_{ij}| + |c_{ij}|) = 8 \left(\|a\|_1 + \|c\|_1 \right), \tag{2.20}$$

where $\| \bullet \|_1$ denotes the l_1 norm. This is independent of the system state, depending only on the interaction strengths.

As B is a qubit, the overall pure state of the system has Schmidt rank 2 in the partition $B|AC$. Therefore, the entanglement across that partition is equivalent to that of a pure, two-qubit state, and can also be measured by the concurrence. But we saw in Eq. (2.4) that the concurrence of a pure state is given by the product of its Schmidt coefficients, so that $\lambda_1 \lambda_2 = C_{B|AC}$.

Using Eq. (2.20) in Eq. (2.17), we obtain a bound on the time-derivative of the concurrence of AC whose dependence on the system state is entirely determined by the concurrence $C_{B|AC}$ of the $B|AC$ partition:

$$\frac{\mathrm{d}C_{AC}^2}{\mathrm{d}t} \leq 8 \left(\|a\|_1 + \|c\|_1 \right) C_{B|AC}. \tag{2.21}$$

The interpretation of this equation is that, for pure states of a system of two qubits interacting indirectly via a third, the entanglement of the qubit mediating the interactions acts as a bottleneck for entanglement generation between the other two. For example, if the mediating qubit is only slightly entangled, entanglement can only be generated slowly. (In particular, we recover the result of Section 1.1 of Chapter 1, that it is impossible to generate any entanglement at all if the mediating particle is not itself entangled.) This is true independent of the form of the interactions.

Note that the entanglement rate must go to zero as the state of AC approaches a maximally entangled state. Although this is not explicit in Eq. (2.21), the bound on the rate does implicitly go to zero as systems A and C become maximally entangled since they can not then be entangled with B, so that $C_{B|AC}$ goes to zero.

2.2 General Tripartite Chains

We have demonstrated that the entanglement of a mediating particle acts as a bottleneck for entanglement generation when all the particles are qubits (and the system is in a pure state). In this section, we extend the result to tripartite systems of arbitrary dimension.

2.2.1 Entanglement and Fidelity

As the state of a subsystem will in general be mixed even if the overall system is in a pure state, we must confront the issue of how to quantify bipartite entanglement of mixed states in arbitrary dimensions. This remains a difficult

open problem; even the question of determining whether such states are entangled or not is unsolved in general.

It is instructive to consider the physical motivation behind our investigation of entanglement dynamics. Entanglement measures are defined in the LOCC paradigm: local operations and classical communication can not increase the entanglement of a state. This is the natural paradigm when thinking about entanglement from an information-theoretic point of view, in which entangled states are shared between different parties who are free to act locally on their part of the state.

But, throughout this chapter, we have had in mind a physical system of interacting particles. From this standpoint, it is not clear what classical communication means. Any transfer of classical information between particles would still have to take place via the (quantum) interactions. In this context, it may be reasonable to weaken the condition, and require only that evolution under a local Hamiltonian leaves the quantity used to measure entanglement invariant (that is, it should not change under local unitary operations). Of course, as this is less restrictive than invariance under LOCC, it is automatically satisfied by any entanglement measure.

We will turn to a quantity that, whilst not an entanglement measure per se (it can increase under LOCC [Verstraete and Verschelde, 2003]), *does* satisfy this weaker condition, and also has a useful operational interpretation: the *entanglement fidelity*, which is related to the usual fidelity between two states.

Definition 2.2 (Fidelity)
The fidelity between two states ρ and σ in the same Hilbert space is defined as

$$F(\rho,\sigma) = \left(\mathrm{Tr}\,\sqrt{\sigma^{1/2}\rho\,\sigma^{1/2}}\right)^2. \tag{2.22}$$

Though it is not obvious from the definition, the fidelity is symmetric: $F(\rho,\sigma) = F(\sigma,\rho)$. If $\sigma = |\varphi\rangle\langle\varphi|$ is a pure state, the fidelity reduces to $F(\rho,|\varphi\rangle) = \langle\varphi|\,\rho\,|\varphi\rangle$ (the form used in the entanglement fidelity). If $\rho = |\psi\rangle\langle\psi|$ is also pure, it reduces to the square of the overlap, $F(|\psi\rangle,|\varphi\rangle) = |\langle\psi|\varphi\rangle|^2$.

Definition 2.3 (Entanglement Fidelity)
For a bipartite state ρ_{AB}, the entanglement fidelity is given by

$$F(\rho_{AB}) = \max_{|\phi\rangle} \langle\phi|\,\rho_{AB}\,|\phi\rangle\,, \tag{2.23}$$

where the maximization is taken over all pure, maximally entangled states $|\phi\rangle_{AB}$ in the bipartite Hilbert space of ρ_{AB}.

The fidelity is an experimentally relevant quantity. When trying to engineer an evolution to produce a particular state (a highly entangled one, for instance), we want to know how close the actual state is to the desired one: precisely what the fidelity measures. For example, in teleportation experiments, it is the entanglement fidelity of the entangled pair that determines how close the teleported state is to the original [Horodecki et al., 1999].

If a state is separable, its entanglement fidelity is less than or equal to $1/n$ (with n the dimension of the smaller of the two Hilbert spaces making up the bipartite space), whereas a state is maximally entangled if and only if the entanglement fidelity is 1. It can also be used to give upper and lower bounds on entanglement measures such as the concurrence (and hence entanglement of formation) of two-qubit states [Verstraete and Verschelde, 2003].

As any maximally entangled state can be obtained from some fixed maximally entangled state by local unitaries, we can give an equivalent definition of the entanglement fidelity in terms of a maximization over unitaries instead of states.

Definition 2.4 (Entanglement Fidelity (generalization))
For a bipartite state ρ_{AB}, the entanglement fidelity is defined by

$$F(\rho_{AB}) = \max_{U_A, U_B} \langle\phi| U_A \otimes U_B \rho_{AB} U_A^\dagger \otimes U_B^\dagger |\phi\rangle . \tag{2.24}$$

From this definition, it is clear that entanglement fidelity has an operational interpretation: it is the maximum overlap with a maximally entangled state that is achievable using only local unitary operations.

A very useful theorem due to Uhlmann, which we will use repeatedly in the next section and in Chapter 3, re-expresses the fidelity in terms of the overlap of purifications of the states.

Theorem 2.5 (Uhlmann)
If ρ and σ are two states in the same Hilbert space $\mathcal{H}$, let $|\psi\rangle$ and $|\varphi\rangle$ be purifications of ρ and σ into a (in general larger) Hilbert space $\mathcal{H}'$. Then

$$F(\rho, \sigma) = \max_{|\psi\rangle, |\varphi\rangle} |\langle\psi|\varphi\rangle|^2 \tag{2.25}$$

where the maximization is over all purifications.

Proof See Jozsa [1994], Uhlmann [1976]. □

Since any purification can be transformed into another by a unitary acting on $\mathcal{H}'$, we can fix one of the purifications and only maximize over the other one. Also, global phases can always be chosen to ensure the overlap $\langle\varphi|\psi\rangle$ is real and positive, allowing the absolute value to be dropped.

2.2.2 Bottleneck Inequality

Returning to our general tripartite chain, we are interested in how the entanglement fidelity rate $\dot{F}_{AC}$ of particles A and C depends on the entanglement fidelity $F_{B|AC}$ of the overall system state $|\psi\rangle_{ABC}$ in the $B|AC$ partition (see Fig. 2.1). Applying Uhlmann's theorem (Theorem 2.5) to F_{AC}, we can write F_{AC} in terms of purifications. Given that $|\psi\rangle_{ABC}$ is always a purification of the reduced state ρ_{AC} of AC, we can use the freedom in fixing one of the purifications to choose the purification of ρ_{AC} to be $|\psi\rangle_{ABC}$. Thus the entanglement fidelity at some time t is

$$F_{AC}(t) = \max_{|\varphi\rangle} |\langle\varphi|\psi\rangle|^2 = \langle\bar{\varphi}|\psi\rangle^2 , \tag{2.26}$$

where $|\varphi\rangle_{ABC}$ is an extension of a maximally entangled state $|\phi\rangle_{AC}$ to $\mathcal{H}_{ABC}$: $|\varphi\rangle_{ABC} = |\phi\rangle_{AC} |\xi\rangle_B$. We denote the particular state achieving the maximum by $|\bar{\varphi}\rangle_{ABC} = |\bar{\phi}\rangle_{AC} |\bar{\xi}\rangle_B$, whose global phase is chosen to ensure the overlap is real and positive. Thus

$$\langle\bar{\varphi}|\psi\rangle \geq \langle\varphi|\psi\rangle \tag{2.27}$$

for all extensions $|\varphi\rangle_{ABC}$ of any maximally entangled state $|\phi\rangle_{AC}$.

Allowing the system to evolve for an infinitesimal time δt, the entanglement fidelity becomes

$$F_{AC}(t+\delta t) = \max_{|\chi\rangle} \left\langle\chi\right| e^{-iH\delta t} \left|\psi\right\rangle^2 = \langle\bar{\chi}| e^{-iH\delta t} |\psi\rangle^2 \tag{2.28}$$

where $H = H_{AB} + H_{BC}$ is the interaction Hamiltonian, we have again chosen one purification to be the (time-evolved) system state, and $|\chi\rangle_{ABC}$ is another extension of a maximally entangled state on AC.

Now, Eq. (2.27) is true for all $|\varphi\rangle$, so in particular it must be true for infinitesimal changes $|\varphi\rangle = |\bar{\varphi}\rangle + \delta t\, |\bar{\varphi}^\perp\rangle$, where $|\bar{\varphi}^\perp\rangle$ can be taken to be orthogonal to $|\bar{\varphi}\rangle$ without loss of generality. Thus $\left\langle\bar{\varphi}^\perp\right|\left.\psi\right\rangle \leq 0$. However, if there were some $|\bar{\varphi}^\perp\rangle$ for which the left hand side were strictly negative, the state $-|\bar{\varphi}^\perp\rangle$ would make it strictly positive. Therefore,

$$\left\langle\bar{\varphi}^\perp\right|\left.\psi\right\rangle = 0 \tag{2.29}$$

for all states $|\bar{\varphi}^\perp\rangle$ orthogonal to $|\bar{\varphi}\rangle$.

Equation (2.28) for the entanglement fidelity at time $t+\delta t$ must tend to Eq. (2.26) as $t \to 0$, so the state maximising Eq. (2.28) must have the form $|\bar{\chi}\rangle = |\bar{\varphi}\rangle + \delta t\, |\bar{\varphi}^\perp\rangle$. Substituting this in Eq. (2.28), expanding the exponential to first order in δt, and using Eq. (2.29) gives

$$F_{AC}(t+\delta t) = \langle\bar{\varphi}| \left(\mathbb{1} - i\delta t H\right) |\psi\rangle^2 + O\left(\delta t^2\right) . \tag{2.30}$$

That is, the state maximizing Eq. (2.26) also maximizes Eq. (2.28), to first order in δt.

Because we have assumed global phases were chosen to make it real and positive, the overlap $\langle\bar{\varphi}|\psi\rangle = \sqrt{F_{AC}(t)}$. Expanding the square in Eq. (2.30) and retaining terms to first order in δt, we obtain an expression for the time-derivative of the entanglement fidelity of particles A and C:

$$\dot{F}_{AC}(t) = -i\sqrt{F_{AC}}\big(\langle\bar{\varphi}|\, H\,|\psi\rangle - \langle\psi|\, H\,|\bar{\varphi}\rangle\big). \tag{2.31}$$

The Schmidt decomposition of the system state in the partition $B|AC$ is given by

$$|\psi\rangle_{ABC} = \sum_i \lambda_i\,|\psi_i\rangle_{AC}\,|i\rangle_B\,, \tag{2.32}$$

where we sort the Schmidt coefficients in descending order: $\lambda_0 \geq \lambda_1 \geq \cdots \geq \lambda_{n-1}$. Extending the set $\{|i\rangle_B\}$ to form a complete basis for $\mathcal{H}_B$, and recalling that $|\bar{\varphi}\rangle$ is the extension of some maximally entangled state on $\mathcal{H}_A \otimes \mathcal{H}_B$, $|\bar{\varphi}\rangle$ can be written in a product decomposition, where the coefficients α_i are in general complex:

$$|\bar{\varphi}\rangle = |\bar{\phi}\rangle_{AC} \sum_i \alpha_i\,|i\rangle_B\,. \tag{2.33}$$

We know that $|\bar{\varphi}\rangle = |\bar{\phi}\rangle_{AC}\,|\bar{\xi}\rangle_B$ maximizes $\langle\varphi|\psi\rangle$, so that $\langle\bar{\varphi}|\psi\rangle = \max_{\alpha_i} \sum_i \alpha_i^* \lambda_i \langle\bar{\phi}|\psi_i\rangle$. This is just an inner product between a vector $\boldsymbol{\alpha}$ with components α_i and normalization $|\boldsymbol{\alpha}| = 1$, and a vector $\boldsymbol{x}$ with components $\lambda_i \langle\bar{\phi}|\psi_i\rangle$. Therefore, $\boldsymbol{\alpha}$ is proportional to $\boldsymbol{x}$, yielding

$$\alpha_i = \frac{\lambda_i \langle\bar{\phi}|\psi_i\rangle}{\sqrt{\sum_k \lambda_k^2 \left|\langle\bar{\phi}|\psi_k\rangle\right|^2}} = \frac{\lambda_i \langle\bar{\phi}|\psi_i\rangle}{\sqrt{F_{AC}}}. \tag{2.34}$$

Note that the denominator is just the square-root of the entanglement fidelity F_{AC}, since using the Schmidt decomposition of $|\psi\rangle_{ABC}$ from Eq. (2.32),

$$\sum_k \lambda_k^2 \left|\langle\bar{\phi}|\psi_k\rangle\right|^2 = \langle\bar{\phi}|\left(\sum_k \lambda_k^2\,|\psi_k\rangle\langle\psi_k|\right)|\bar{\phi}\rangle = \langle\bar{\phi}|\,\mathrm{Tr}_B\big[|\psi\rangle\langle\psi|\big]\,|\bar{\phi}\rangle = F_{AC}. \tag{2.35}$$

Writing $|\psi\rangle$ and $|\bar{\varphi}\rangle$ in their product decompositions of Eq. (2.32) and Eq. (2.33), we have

$$\langle\bar{\varphi}|\, H\,|\psi\rangle = \sum_{i,j} \alpha_j^* \lambda_i \langle\bar{\phi}, j|\, H\,|\psi_i, i\rangle\,. \tag{2.36}$$

Using an argument similar to that leading to Eq. (2.29), we will derive a relationship between the diagonal and non-diagonal terms of this expression.

Since a local unitary acting on half of a maximally entangled state produces another maximally entangled state, Eq. (2.27) gives

$$\langle\bar{\varphi}|\psi\rangle \geq \langle\bar{\varphi}|\, U_A\, |\psi\rangle\,, \tag{2.37}$$

similarly for any local unitary U_C acting only on C. In particular this is true for infinitesimal unitaries $U_A = \exp(-iH_A\delta t) = \mathbb{1} - iH_A\delta t + O(\delta t^2)$. Inserting this in Eq. (2.37) and taking the limit $\delta t \to 0$, we obtain $\langle\bar{\varphi}|\, H_A\, |\psi\rangle = \langle\bar{\varphi}|\, H_C\, |\psi\rangle = 0$, or, using Eq. (2.36),

$$\alpha_0^*\lambda_0 \left\langle\bar{\phi}\right|\, H_A\, |\psi_0\rangle = -\sum_{i\neq 0} \alpha_i^*\lambda_i \left\langle\bar{\phi}\right|\, H_A\, |\psi_i\rangle\,, \tag{2.38}$$

with a similar expression for a Hamiltonian H_C acting only on C.

Now, the system Hamiltonian $H = H_{AB} + H_{AC}$, and

$$\langle\bar{\varphi}|\, H_{AB}\, |\psi\rangle = \sum_{i,j} \alpha_j^*\lambda_i\, \langle\bar{\phi}|\, \langle j|_B\, H_{AB}\, |i\rangle_B\, |\psi_i\rangle\,. \tag{2.39}$$

For the $i = j = 0$ terms, $\langle 0|_B\, H_{AB}\, |0\rangle_B$ is just some Hamiltonian acting only on A. Similarly for $\langle 0|_B\, H_{BC}\, |0\rangle_B$. Therefore, using Eq. (2.38) to replace the $i = j = 0$ term, we obtain

$$\langle\bar{\varphi}|\, H\, |\psi\rangle = \sum_{i\neq 0} \alpha_i^*\lambda_i\big(\left\langle\bar{\phi}, i\right|\, H\, |\psi_i, i\rangle - \left\langle\bar{\phi}, 0\right|\, H\, |\psi_i, 0\rangle\big) + \sum_{i\neq j} \alpha_j^*\lambda_i \left\langle\bar{\phi}, j\right|\, H\, |\psi_i, i\rangle\,. \tag{2.40}$$

Substituting for the α_i's from Eq. (2.34), this becomes

$$\langle\bar{\varphi}|\, H\, |\psi\rangle = \sum_{i\neq 0} \lambda_i^2 h_{ii} + \sum_{i\neq j} \lambda_i\lambda_j h_{ij}, \tag{2.41}$$

where

$$h_{ii} = \frac{\left\langle\psi_i\middle|\bar{\phi}\right\rangle}{\sqrt{F_{AC}}} \left(\left\langle\bar{\phi}, i\right|\, H\, |\psi_i, i\rangle - \left\langle\bar{\phi}, 0\right|\, H\, |\psi_i, 0\rangle\right), \tag{2.42a}$$

$$h_{ij} = \frac{\left\langle\psi_i\middle|\bar{\phi}\right\rangle}{\sqrt{F_{AC}}} \left\langle\bar{\phi}, i\right|\, H\, |\psi_i, j\rangle \quad \text{for } i \neq j. \tag{2.42b}$$

Substituting Eq. (2.41) in Eq. (2.31) for the time-derivative of F_{AC} gives

$$\dot{F}_{AC}(t) = -i\sqrt{F_{AC}}\Big(\sum_i \lambda_i^2(h_{ii} - h_{ii}^\dagger) + \sum_{i,j} \lambda_i\lambda_j(h_{ij} - h_{ij}^\dagger)\Big). \tag{2.43}$$

Finally, as we are interested in the dependence of the entanglement fidelity rate of AC on the entanglement across the $AC|B$ partition (which depends

only on the Schmidt coefficients λ_i), any bound on the h_{ij} that depends only on the interaction strengths is sufficient for our purposes. For any i and j, we have

$$\frac{h_{ij} - h_{ij}^\dagger}{i} \leq 2\,|h_{ij}| \leq 2\frac{\left|\langle\bar{\phi}, i|\, H\,|\psi_i, j\rangle\right| + \left|\langle\bar{\phi}, 0|\, H\,|\psi_i, 0\rangle\right|}{\sqrt{F_{AC}}} \leq \frac{4\rho(H)}{\sqrt{F_{AC}}} \leq \frac{4\,\|H\|}{\sqrt{F_{AC}}} \tag{2.44}$$

where we have used the bound $\left|\langle\bar{\phi}|\psi_k\rangle\right| \leq 1$ in the second inequality, $\rho(H)$ is the spectral radius of the system Hamiltonian, and we have used the fact that the spectral radius is a lower bound on any matrix norm in the final inequality.

Putting Eq. (2.43) and Eq. (2.44) together and rearranging the sums, we arrive at our final result:

$$\dot{F}_{AC} \leq 4\,\|H\|\left(\sum_{i,j}\lambda_i\lambda_j - \lambda_0^2\right). \tag{2.45}$$

The quantity in brackets is closely related to the entanglement fidelity of the pure state $|\psi\rangle_{ABC}$ in the $B|AC$ partition:

$$F_{B|AC}(|\psi\rangle\langle\psi|) = \max_{|\phi\rangle_{B|AC}} |\langle\phi|\psi\rangle|^2 = \max_{|\phi\rangle_{B|AC}}\left|\sum_i \lambda_i\,\langle\phi|\psi_i, i\rangle\right|^2 \tag{2.46a}$$

$$= \left|\frac{1}{\sqrt{n}}\sum_{i,j}\lambda_i\,\langle\psi_j, j|\psi_i, i\rangle\right|^2 = \tfrac{1}{n}\sum_{i,j}\lambda_i\lambda_j, \tag{2.46b}$$

where we have used the Schmidt decomposition of $|\psi\rangle_{ABC}$ from Eq. (2.32) in the second equality, and the fact that the particular maximally entangled state maximising the expression is clearly $|\phi\rangle_{B|AC} = \sum_j \frac{1}{\sqrt{n}}\,|\psi_j\rangle_{AC}\,|j\rangle_B$ (n is the smaller of the dimensions of B and AC). Subtracting λ_0^2 in Eq. (2.45) rescales this entanglement fidelity so that it is zero when the state is separable in the $B|AC$ partition (when $\lambda_0 = 1$, $\lambda_{i>0} = 0$).

Therefore, we again have a bottleneck inequality: the entanglement of the particle that mediates the interactions (B) limits the rate at which entanglement can be generated between the others (A and C). As in the three-qubit case, the bound on the rate implicitly goes to zero when systems A and C become maximally entangled, since they can not then be entangled with B and the rescaled entanglement fidelity on the right hand side of Eq. (2.45) goes to zero.

However, unlike the bottleneck inequality of Eq. (2.21) derived in Section 2.1, which assumed qubit subsystems, the inequality of Eq. (2.45) for the entanglement fidelity is valid for tripartite systems of arbitrary dimension. Therefore, it can be applied to more than just tripartite chains. In a large

system of many interacting particles in which the particles do not all interact directly with each other (i.e. the interaction network is not fully connected), it is always possible to group the particles into three subsets, A, B and C, such that no particle in A interacts directly with a particle in C (see Fig. 2.2). The bottleneck inequality then applies to the three compound systems formed from these groups of particles.

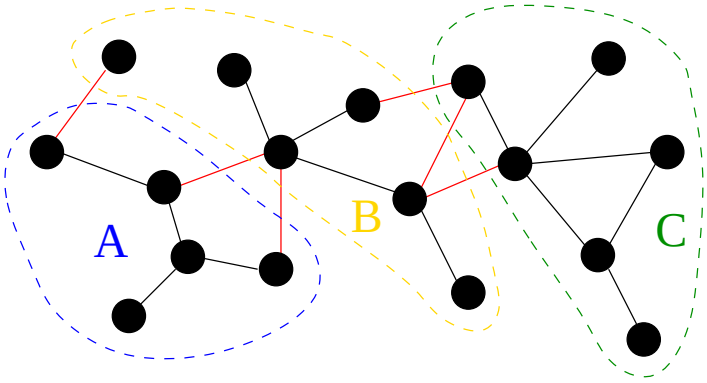

Figure 2.2: In any network of interacting particles that is not fully connected, the particles can be grouped into three sets, A, B and C, such that particles in A and C do not interact directly, but only indirectly via particles in B. The general bottleneck inequality of Eq. (2.45) then applies to these sets, with the Hamiltonian given by interactions that cross the boundaries of the sets. The entanglement of set B with the rest acts as a bottleneck to entanglement generation between sets A and C. For example, for the sets A, B and C shown in the figure, the Hamiltonian is given by the interactions shown in red.

Of course, considering an entire set of particles as a single system loses any information about the structure of the interactions within that set. In the next chapter, we will derive a set of entanglement *rate equations* for large sets of interacting particles that take the complete structure of the interaction network into account.

Chapter 3

Entanglement Rate Equations

3.1 Motivation from Chemistry

In the previous chapter, we showed that a particle mediating indirect interactions between two others acts as a bottleneck for entanglement generation between them, and remarked that the general result of Section 2.2 can be applied to *any* system of interacting particles by dividing those particles into three sets. This bottleneck result captures one aspect of the entanglement dynamics of an interacting system. However, it neglects all details of the structure of the interactions within a set of particles. Additionally, the result only applies to systems whose overall state is pure, as demonstrated by the results of Chapter 1.

Consider a complex chemical reaction, whose reaction mechanism consists of many intermediate steps. The reaction mechanism describes the steps by which reactants are transformed, via successive intermediate compounds, into the final products. The rate at which a compound is produced depends on the amounts of its immediate precursors that are present, as described by the rate equation for that step. The complete reaction is described by a complete set of such *rate equations*, that is a set of coupled differential equations, one for each step in the reaction mechanism. Significantly, if one step in the reaction is slow, it limits the rate of later steps, as the concentrations of some of the necessary compounds will be low. For example, writing the reaction shown in Fig. 3.1 as

$$A + B \rightleftharpoons C + D \rightleftharpoons E + B, \tag{3.1}$$

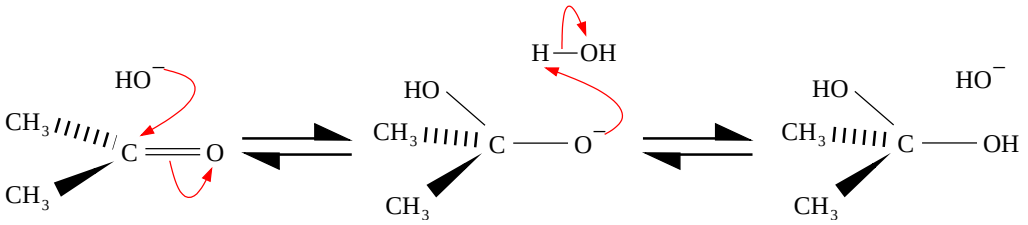

Figure 3.1: The reaction mechanism for a chemical reaction (in this case the base-catalyzed hydration of a ketone) describes the steps by which reactants are transformed into the final products. Each step in the reaction mechanism gives rise to a rate equation: a differential equation relating the change in the concentrations of the step's products to the concentrations of its reactants. The entire reaction is therefore described by a set of coupled differential equations.

it is described by the set of rate equations

$$\frac{\mathrm{d}[A]}{\mathrm{d}t} = -k_1[A][B] + k_2[C][D], \tag{3.2a}$$

$$\frac{\mathrm{d}[B]}{\mathrm{d}t} = -k_1[A][B] + k_2[C][D] + k_3[C][D] - k_4[E][B], \tag{3.2b}$$

$$\frac{\mathrm{d}[C]}{\mathrm{d}t} = k_1[A][B] - k_2[C][D] - k_3[C][D] + k_4[E][B], \tag{3.2c}$$

$$\frac{\mathrm{d}[D]}{\mathrm{d}t} = k_1[A][B] - k_2[C][D] - k_3[C][D] + k_4[E][B], \tag{3.2d}$$

$$\frac{\mathrm{d}[E]}{\mathrm{d}t} = k_3[C][D] - k_4[E][B], \tag{3.2e}$$

where the k's are constants and the square brackets denote concentrations. If one step in the reaction mechanism is particularly slow, it can act as a bottleneck for later steps, and this *rate determining step* will effectively determine the overall reaction rate.

We would like to derive an analogous description of entanglement flow. Of course, we have seen in Chapter 1 that there is a sense in which entanglement does not flow at all. Nevertheless, after introducing a generalization of the entanglement fidelity in Section 3.2.1, we will derive a set of *entanglement rate equations* in Section 3.2.2 that are somewhat analogous to the rate equations describing a chemical reaction, thereby demonstrating that the intuitive concept of entanglement *flow* does have a physical meaning, despite the results of Chapter 1, and indeed can be made quantitative.

The entanglement rate equations apply to arbitrary networks of interacting particles, including systems in mixed states, and take into account the full structure of the interaction network. We will see that, true to the chemical reaction analogy, if one step in the entanglement flow is slow, it limits the

rate of subsequent steps. In Section 3.3, we will demonstrate for some simple examples how the complete set of entanglement rate equations is derived. In Section 3.4, we prove that the curves saturating the inequalities in a set of rate equations constitute upper bounds on all possible evolutions of the entanglement fidelities. Finally, we consider long-range entanglement generation in Section 3.3, discussing two different entanglement distribution protocols in the context of the rate equations, before using the results of Section 3.4 to prove a universal lower-bound on the time required to entangle the end spins in a spin chain (or, more precisely, the scaling of that time with the length of the chain).

3.2 Derivation of the Entanglement Rate Equations

3.2.1 Generalising the Entanglement Fidelity

As in Chapter 2, we will need to measure the entanglement of arbitrary-dimensional mixed states, and once again we will use the entanglement fidelity for this (see Definition 2.3 and the discussion preceding it). A natural question is how entanglement flows when generating bipartite entanglement between two particular particles in the network. This, along with Definition 2.4 of the entanglement fidelity in terms of an optimization over local unitaries, motivates the following generalization of the entanglement fidelity to subsystems.

Definition 3.1 (Entanglement Fidelity (generalization))
Given a bipartite Hilbert space $\mathcal{H}_a \otimes \mathcal{H}_b$ of two particles a and b, and given a bipartite state ρ_{AB} of a (generally larger) bipartite Hilbert space $\mathcal{H}_A \otimes \mathcal{H}_B$, where $\mathcal{H}_a \subseteq \mathcal{H}_A$ and $\mathcal{H}_b \subseteq \mathcal{H}_B$, the entanglement fidelity of ρ_{AB} relative to subsystem ab is defined by

$$F(\rho_{AB}) = \max_{U_A, U_B} \langle\phi|_{ab} \operatorname{Tr}_{/ab} \left[U_A \otimes U_B \rho_{AB} U_A^\dagger \otimes U_B^\dagger \right] |\phi\rangle_{ab} \,, \tag{3.3}$$

where U_A and U_B are unitary operations on $\mathcal{H}_A$ and $\mathcal{H}_B$, $\operatorname{Tr}_{/ab}$ denotes the partial trace over everything other than *$\mathcal{H}_a \otimes \mathcal{H}_b$, and $|\phi\rangle_{ab}$ is any fixed maximally entangled state on $\mathcal{H}_a \otimes \mathcal{H}_b$.*

This generalized entanglement fidelity has the same operational interpretation as the usual entanglement fidelity: it is the maximum overlap with a maximally entangled state that is achievable using only local unitaries. If $\mathcal{H}_A = \mathcal{H}_a$ and $\mathcal{H}_B = \mathcal{H}_b$, it reduces to the standard definition of Section 2.2.1.

It also has the desirable property that tracing over a subsystem can not increase the entanglement fidelity. More precisely, if A and B are subsystems of A' and B' respectively, where A' and A contain a common subsystem a, and similarly B' and B contain a common subsystem b, the entanglement fidelity with respect to ab of the state of $A'B'$ is always larger than or equal to the entanglement fidelity with respect to ab of the state $\rho_{AB} = \mathrm{Tr}_{/AB}[\rho_{A'B'}]$ of AB. This is immediately clear from Definition 3.1, as the maximization in the AB case is over a subset of the unitaries maximized over in the $A'B'$ case.

3.2.2 Rate Equations for Entanglement Fidelities

We consider a system of particles S whose interactions are described by some interaction network, an example of which is shown in Fig. 3.2. We can write the system Hamiltonian as $H = \sum_{ij} H_{ij}$, where the sum is over particles i and j that are connected by an interaction. Given two particular particles a and b, we choose any two sets of particles, A and B, that contain a and b (see Fig. 3.2).

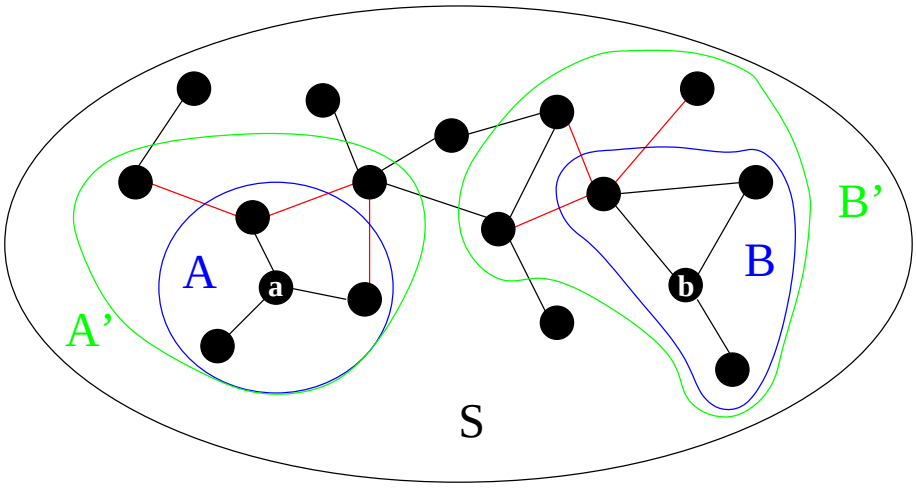

Figure 3.2: A network of interacting particles, showing the sets involved in the entanglement rate equations. Interactions crossing the boundaries of A and B are indicated in red.

The first part of the derivation of the entanglement rate equations is a generalization of the analysis in Section 2.2 leading to Eq. (2.31). Applying Uhlmann's theorem (Theorem 2.5) to the entanglement fidelity with respect

to the ab subsystem of the state ρ_{AB}, we have

$$F_{AB} = \max_{U_A,U_B} \left| \langle\phi|_{ab} \underbrace{\mathrm{Tr}_{/ab}[U_A \otimes U_B \; \rho_{AB}(t) \; U_A^\dagger \otimes U_B^\dagger]}_{\sigma_{ab}} |\phi\rangle_{ab} \right|^2 \tag{3.4a}$$

$$= \max_{\substack{U_A,U_B \\ |\chi\rangle}} \left| \langle\chi| \, U_A \otimes U_B \, |\psi\rangle \right|^2 \tag{3.4b}$$

$$= \langle\bar{\chi}| \, \bar{U}_A \otimes \bar{U}_A \, |\psi\rangle^2 . \tag{3.4c}$$

The state $U_A \otimes U_B \, |\psi\rangle$ is a purification of σ_{ab}. (We are retaining the unitaries, rather than incorporating them into one of the purifications, for later convenience. Strictly speaking, they should be extended to $\mathcal{H}_{AB} \otimes \mathcal{H}'$ (see Theorem 2.5) and written $U_A \otimes U_B \otimes \mathbb{1}_{\mathcal{H}'}$. In the interests of economy, we drop all $\mathbb{1}_{\mathcal{H}'}$'s from now on.) Using the freedom to fix one of the purifications in Uhlmann's theorem, we can choose $|\psi\rangle$ to be a purification of the overall state ρ_S of the entire system, which guarantees that $U_A \otimes U_B \, |\psi\rangle$ is a purification of σ_{ab}, as required.

The state $|\chi\rangle$ is a purification of the maximally entangled state $|\phi\rangle$, i.e. $|\chi\rangle = |\phi\rangle_{ab} \, |\vartheta\rangle$ is an extension of a maximally entangled state on $\mathcal{H}_a \otimes \mathcal{H}_b$ to the larger Hilbert space $\mathcal{H}_{AB} \otimes \mathcal{H}'$. The maximization over $|\chi\rangle$ then reduces to a maximization over $|\vartheta\rangle$. We have introduced the notation $|\bar{\chi}\rangle$, $\bar{U}_A$ and $\bar{U}_B$ to denote the particular state and unitaries achieving the maximum, and have dropped the absolute value as the global phase of $|\bar{\chi}\rangle$ can always be chosen to make the expression real and positive.

If the system evolves under the Hamiltonian $H = \sum_{ij} H_{ij}$ for an infinitesimal time δt, the state evolves to $\rho_{AB}(t + \delta t) = \mathrm{Tr}_{/AB}[e^{-iH\delta t} \rho_S(t) \, e^{iH\delta t}]$. δt is infinitesimal, so the exponential can be decomposed as a product $e^{-iH\delta t} = \prod_{ij} e^{-iH_{ij}\delta t} + O(\delta t^2)$. Since unitaries acting locally on A or B can not change the entanglement fidelity, only those interactions H_{ij} involving at least one particle outside A and B give a first-order contribution to the evolution of the entanglement fidelity. Therefore, we need only include that smaller set of interactions. Letting

$$\mathrm{d}U = \exp\left(-i\delta t \sum_{\substack{i \in S \\ j \in \bar{A} \cup B}} H_{ij} \right) \tag{3.5}$$

be the resulting (infinitesimal) unitary evolution operator, the entanglement fidelity after the evolution becomes

$$F_{AB}(t + \delta t) = \max_{\substack{V_A,V_B \\ |\zeta\rangle}} \left| \langle\zeta| \, (V_A \otimes V_B) \, \mathrm{d}U \, |\xi\rangle \right|^2 , \tag{3.6}$$

where we have used Uhlmann's theorem again. V_A and V_B are unitaries arising from Definition 3.1 of the entanglement fidelity, the state $|\zeta\rangle$ is again simply an extension of $|\phi\rangle$ to $\mathcal{H}_{AB} \otimes \mathcal{H}'$, and $(V_A \otimes V_B)\,\mathrm{d}U\,|\xi\rangle$ can be chosen to be any fixed purification of the two-qubit density operator

$$\tau_{ab} = \mathrm{Tr}_{/ab}\left[(V_A \otimes V_B)\,\mathrm{d}U \rho_S\,\mathrm{d}U^\dagger\,(V_A \otimes V_B)^\dagger\right]. \tag{3.7}$$

Again making use of this freedom, and recalling that we previously chose $|\psi\rangle$ to be a purification of ρ_S, we can choose $|\xi\rangle$ to be the same state: $|\xi\rangle = |\psi\rangle$, which automatically satisfies the condition that $(V_A \otimes V_B)\,\mathrm{d}U\,|\xi\rangle$ must be a purification of τ_{ab}.

The state $|\bar{\chi}\rangle$ and unitaries $\bar{U}_A$ and $\bar{U}_B$ were defined to be those maximizing Eq. (3.4b). Thus, by definition,

$$\langle\bar{\chi}|\,\bar{U}_A \otimes \bar{U}_B\,|\psi\rangle \geq |\langle\chi|\,U_A \otimes U_B\,|\psi\rangle| \tag{3.8}$$

for all $|\chi\rangle$, U_A and U_B. In particular, this is true for infinitesimal changes, e.g. $|\bar{\chi}\rangle + \delta t\,|\chi^\perp\rangle$ where $|\chi^\perp\rangle$ is orthogonal to $|\chi\rangle$. Thus $\langle\chi^\perp|\,\bar{U}_A \otimes \bar{U}_B\,|\psi\rangle \leq 0$. However, if this were strictly negative for some $|\chi^\perp\rangle$, then $-|\chi^\perp\rangle$ would make it positive. Therefore

$$\langle\chi^\perp|\,\bar{U}_A \otimes \bar{U}_B\,|\psi\rangle = 0. \tag{3.9}$$

Similarly, considering the unitaries $\bar{U}_A$ and $\bar{U}_B$, by definition

$$\left\langle\bar{\xi}\right|\bar{U}_A \otimes \bar{U}_B\,|\psi\rangle \geq \left\langle\bar{\xi}\right|U_A \otimes U_B\,|\psi\rangle \tag{3.10}$$

for any U_A and U_B, in particular for infinitesimal changes such as $U_A = \bar{U}_A \exp(-iH_A\mathrm{d}t) = \bar{U}_A - i\mathrm{d}t\bar{U}_A H_A$. Thus, by a very similar argument,

$$\langle\bar{\chi}|\,\bar{U}_A H_A \otimes \bar{U}_B\,|\psi\rangle = \langle\bar{\chi}|\,\bar{U}_A \otimes \bar{U}_B H_B\,|\psi\rangle = 0. \tag{3.11}$$

Equation (3.6) for the entanglement fidelity at time $t + \delta t$ must tend to Eq. (3.4b) (the corresponding expression for time t) as $\delta t \to 0$, so $|\zeta\rangle = |\bar{\chi}\rangle + \delta t\,|\chi^\perp\rangle$ and $V_{A/B} = \bar{U}_{A/B}(\mathbb{1} + i\delta t H_{A/B})$, where $H_{A/B}$ are Hermitian operators on A or B respectively. Using this, expanding $\mathrm{d}U = \mathbb{1} - i\delta t H + O(\delta t^2)$ (where H is the sum of interactions involving at least one particle in A or B), and making use of Eq. (3.9) and Eq. (3.11), we have

$$F_{AB}(t + \delta t) = \langle\bar{\chi}|\,\bar{U}_A \otimes \bar{U}_B(\mathbb{1} - i\delta t H)\,|\psi\rangle^2 + O\left(\delta t^2\right). \tag{3.12}$$

That is, the state $|\bar{\chi}\rangle$ and unitaries $\bar{U}_A$ and $\bar{U}_B$ maximizing Eq. (3.4b) also maximizes Eq. (3.6), to first order in δt.

From the definition of the infinitesimal evolution operator $\mathrm{d}U$ in Eq. (3.5), the Hamiltonian H in Eq. (3.12) currently includes all interactions involving at

least one particle in A or B. By expanding H as a sum over these two-particle interactions, we can again use Eq. (3.9) and Eq. (3.11) to show that only interactions that *cross the boundary* of A or B need to be included to give the entanglement fidelity to first order in δt:

$$F_{AB}(t+\delta t) = \langle\bar{\chi}|\, \bar{U}_A \otimes \bar{U}_B \Big(\mathbb{1} - i\delta t \sum_{i\notin A, j\in A} H_{ij} - i\delta t \sum_{i\notin B, j\in B} H_{ij} \Big) |\psi\rangle^2 + O\left(\delta t^2\right). \quad (3.13)$$

Now, as global phases were chosen to make it real and positive,

$$\langle\bar{\chi}|\, \bar{U}_A \otimes \bar{U}_B\, |\psi\rangle = \sqrt{F_{AB}(t)}. \quad (3.14)$$

Expanding the square in Eq. (3.13) and only retaining first-order terms in δt, we arrive at a first expression for the time-derivative of the entanglement fidelity:

$$\dot{F}_{AB} = \sqrt{F_{AB}} \cdot \frac{1}{i} \sum_{\substack{i\notin A\cup B \\ j\in A\cup B}} (\langle\varphi|\, H_{ij}\, |\psi\rangle - \langle\psi|\, H_{ij}\, |\varphi\rangle), \quad (3.15)$$

where $|\varphi\rangle = \bar{U}_A^\dagger \otimes \bar{U}_B^\dagger\, |\bar{\chi}\rangle$, and we have dropped the explicit time-dependence of $F_{AB}(t)$ for notational simplicity. (Cf. Eq. (2.31) of Section 2.2).

To proceed, we will need the following proposition by Fan-Hoffman, which we will use in the proof of the subsequent Lemma.

Proposition 3.2 (Fan-Hoffman)
For any matrix X, the ordered singular values $\sigma_i^\downarrow$ of X, where $\sigma_1^\downarrow \geq \sigma_2^\downarrow \geq \cdots \geq \sigma_n^\downarrow$, are individually greater than or equal to the ordered eigenvalues $r_i^\downarrow$ of $\operatorname{Re} X = (X + X^\dagger)/2$, that is $\sigma_i^\downarrow \geq r_i^\downarrow$ for all i.

Proof See Bhatia [1997]. □

Note that the eigenvalues of $\operatorname{Re} X$ can be negative, in which case the absolute values of the eigenvalues do *not* necessarily obey Proposition 3.2.

Lemma 3.3 (Matrix Trace Inequality for Real and Imaginary Parts)
For any matrix X,

$$(\operatorname{Tr}|X|)^2 - (\operatorname{Tr}[\operatorname{Re} X])^2 \geq \operatorname{Tr}[(\operatorname{Im} X)^2], \quad (3.16)$$

where

$$|X| = \sqrt{XX^\dagger}, \quad \operatorname{Re} X = \frac{X + X^\dagger}{2}, \quad \operatorname{Im} X = \frac{X - X^\dagger}{2i}. \quad (3.17)$$

Proof Assume initially that $\mathrm{Tr}[\mathrm{Re}\, X]$ is non-negative. Defining P (N) to be the set of indices of the positive semi-definite (negative) eigenvalues of $\mathrm{Re}\, X$,

$$\begin{aligned}
&\sum_{i\neq j}(\sigma_i\sigma_j - r_i r_j)\\
&= \sum_{i\neq j}\sigma_i\sigma_j - \sum_{\substack{i,j\in P\\ i\neq j}} r_i r_j + \sum_{\substack{i\in P\\ j\in N}} |r_i|\,|r_j| + \sum_{\substack{i\in N\\ j\in P}} |r_i|\,|r_j| - \sum_{\substack{i,j\in N\\ i\neq j}} |r_i|\,|r_j| \qquad (3.18a)\\
&= \sum_{\substack{i\,\mathrm{or}\,j\notin P\\ i\neq j}}\sigma_i\sigma_j + \sum_{\substack{i\in P\\ j\in N}} |r_i|\,|r_j| + \sum_{\substack{i,j\in P\\ i\neq j}}\left(\sigma_i^{\downarrow}\sigma_j^{\downarrow} - r_i^{\downarrow}r_j^{\downarrow}\right) + \sum_{i\in N}|r_i|\left(\sum_{j\in P}|r_j| - \sum_{\substack{j\in N\\ j\neq i}}|r_j|\right). \qquad (3.18b)
\end{aligned}$$

The first two terms are clearly positive, the third is positive by Proposition 3.2, and the last by the assumption that $\mathrm{Tr}[\mathrm{Re}\, X] \geq 0$. Thus

$$\sum_{i\neq j}(\sigma_i\sigma_j - r_i r_j) \geq 0. \qquad (3.19)$$

Now, making use of this,

$$\begin{aligned}
&\left(\mathrm{Tr}\,|X|\right)^2 - \left(\mathrm{Tr}[\mathrm{Re}\, X]\right)^2\\
&\quad= \left(\sum_i \sigma_i\right)^2 - \left(\sum_i r_i\right)^2 \qquad (3.20a)\\
&\quad= \sum_i \sigma_i^2 - \sum_i r_i^2 + 2\sum_{i\neq j}(\sigma_i\sigma_j - r_i r_j) \qquad (3.20b)\\
&\quad\geq \sum_i \sigma_i^2 - \sum_i r_i^2 \qquad (3.20c)\\
&\quad= \mathrm{Tr}\left[XX^{\dagger}\right] - \mathrm{Tr}[(\mathrm{Re}\, X)^2] = \mathrm{Tr}[(\mathrm{Im}\, X)^2] \qquad (3.20d)
\end{aligned}$$

where in the last line we have expanded $X = \mathrm{Re}\, X + i\,\mathrm{Im}\, X$, and used the fact that $\mathrm{Re}\, X$ and $\mathrm{Im}\, X$ are both Hermitian and that the trace of their commutator is 0.

For completeness, we can remove the assumption $\mathrm{Tr}[\mathrm{Re}\, X] \geq 0$ by noting that, if there existed a matrix X with $\mathrm{Tr}[\mathrm{Re}\, X] < 0$ such that the Lemma did not hold, then the matrix $-X$ would also violate the Lemma. But then $\mathrm{Tr}[\mathrm{Re}(-X)] \geq 0$, so the Lemma must hold for all matrices. □

The term H_{ij} in the Hamiltonian acts only on the particles i and j.

Therefore

$$\frac{1}{i}\big(\langle\varphi|\,H_{ij}\,|\psi\rangle - \langle\psi|\,H_{ij}\,|\varphi\rangle\big) \tag{3.21a}$$
$$= \mathrm{Tr}\Big[H_{ij}\cdot\frac{1}{i}\,\mathrm{Tr}_{\not ij}\big[|\psi\rangle\langle\varphi| - |\varphi\rangle\langle\psi|\big]\Big] \tag{3.21b}$$
$$= 2\,\mathrm{Tr}\left[H_{ij}\,\mathrm{Im}\,X_{ij}\right] \quad \text{where } X_{ij} = \mathrm{Tr}_{/ij}\,|\psi\rangle\langle\varphi| \tag{3.21c}$$
$$\leq 2\,|\mathrm{Tr}[H_{ij}\,\mathrm{Im}\,X_{ij}]| \tag{3.21d}$$
$$\leq 2\sqrt{\mathrm{Tr}[H_{ij}^2]}\sqrt{\mathrm{Tr}[(\mathrm{Im}\,X_{ij})^2]} \tag{3.21e}$$
$$\leq 2\,\|H_{ij}\|_{\mathrm{HS}}\sqrt{(\mathrm{Tr}\,|X_{ij}|)^2 - (\mathrm{Tr}[\mathrm{Re}\,X_{ij}])^2}, \tag{3.21f}$$

using the Cauchy-Scwartz inequality for the Hilbert-Schmidt inner-product and the fact that H_{ij} and $\mathrm{Im}\,X_{ij}$ are both Hermition in the penultimate line, and Lemma 3.3 in the last line. $\|\bullet\|_{\mathrm{HS}}$ denotes the Hilbert-Schmidt norm, and

$$X_{ij} = \mathrm{Tr}_{/ij}\,|\psi\rangle\langle\varphi|\,. \tag{3.22}$$

Finally, we need to relate the quantities under the square-root to entanglement fidelities. Firstly,

$$(\mathrm{Tr}[\mathrm{Re}\,X_{ij}])^2 = (\mathrm{Re}(\mathrm{Tr}\,|\psi\rangle\langle\varphi|))^2 = \langle\varphi|\psi\rangle^2 = F_{AB}(t), \tag{3.23}$$

where the last equality follows from Eq. (3.4) for F_{AB} because global phases were chosen to make $\langle\varphi|\psi\rangle$ real and positive (recall from Eq. (3.15) that $|\varphi\rangle = \bar{U}_A^\dagger \otimes \bar{U}_B^\dagger\,|\bar{\chi}\rangle$).

Secondly, from Eq. (3.15) for the time-derivative of $\dot{F}_{AB}$, we are only interested in the H_{ij} that act on one particle j within one of the sets A or B, and one particle i that is outside both. If j is in A, define the sets $A'_i = A \cup i$ and $B'_i = B$. If it is in B, define $A'_i = A$, $B'_i = B \cup i$. We apply Uhlmann's theorem again, this time to the entanglement fidelity $F_{A'_iB'_i}$ of $\rho_{A'_iB'_i}$ with respect to the same particles a and b (see Fig. 3.2), and again choose the same state $|\psi\rangle$ (a purification of the overall system state) for one

of the purifications. So long as A_i' and B_i' are disjoint, we have

$$F_{A_iB_i'} = \max_{\substack{V_{A_i'}, V_{B_i'} \\ |\zeta\rangle}} \left|\langle\zeta| V_{A_i'} \otimes V_{B_i'} |\psi\rangle\right|^2 \tag{3.24a}$$

$$\geq \max_{V_{A_i'}, V_{B_i'}} \left|\langle\varphi| V_{A_i'} \otimes V_{B_i'} |\psi\rangle\right|^2 \tag{3.24b}$$

$$\geq \max_{U_{ij}} \left|\langle\varphi| U_{ij} |\psi\rangle\right|^2 \tag{3.24c}$$

$$= \max_{U_{ij}} \left|\mathrm{Tr}[U_{ij} X_{ij}]\right|^2 \tag{3.24d}$$

$$= \mathrm{Tr}[|X_{ij}|]^2. \tag{3.24e}$$

An inequality appears each time we restrict the maximization. In the second line, we choose the particular state $|\varphi\rangle$ defined in Eq. (3.15) and drop the maximization over states. U_{ij} is a unitary acting on particles i and j, which must both be within either A_i' or B_i' if those sets are disjoint. The maximization over unitaries $V_{A_i'}$ and $V_{B_i'}$ acting on A_i' and B_i' is therefore restricted to a smaller set of unitaries U_{ij} in the third line. The final equality follows from the following simple Lemma:

Lemma 3.4 (Variational Characterization of the Absolute Value) *For any matrix X*

$$\max_U \left|\mathrm{Tr}[UX]\right| = \mathrm{Tr}\left|X\right|, \tag{3.25}$$

where $|X| = \sqrt{XX^\dagger}$.

Proof This is easily proven using the polar decomposition $X = |X|\, V$, where V is unitary:

$$|\mathrm{Tr}[XU]| = |\mathrm{Tr}[|X|VU]| = \left|\mathrm{Tr}\left[|X|^{1/2} |X|^{1/2} VU\right]\right| \tag{3.26a}$$

$$\leq \sqrt{\mathrm{Tr}\,|X|}\sqrt{\mathrm{Tr}\left[U^\dagger V^\dagger |X| VU\right]} = \mathrm{Tr}\,|X| \tag{3.26b}$$

using the Cauchy-Schwartz inequality for the Hilbert-Schmidt inner product, followed by the cyclic property of the trace. Equality is obtained by choosing $U = V^\dagger$. □

If A_i' and B_i' have a particle in common (and it must be particle i if they do), then the inequality of Eq. (3.24c) is not valid. We can instead bound

$$\mathrm{Tr}[|X_{ij}|]^2 = |\langle\varphi|\psi\rangle| \leq 1. \tag{3.27}$$

Thus, using Eq. (3.23) and Eq. (3.24e) (A_i' and B_i' disjoint) or Eq. (3.27) (A_i' and B_i' overlapping) in Eq. (3.21f), and substituting the result in

Eq. (3.15) for the time-derivative of $\dot{F}_{AB}$, we arrive at the entanglement rate equation:

$$\dot{F}_{AB}(t) \leq 2 \sum_{\substack{i \notin A, j \in A; \\ i \notin B, j \in B}} \|H_{ij}\|_{\mathrm{HS}} \sqrt{F_{AB}(t)} \sqrt{F_{A'_i B'_i}(t) - F_{AB}(t)}, \tag{3.28}$$

where $F_{A'_i B'_i}$ is defined to be 1 if sets A'_i and B'_i overlap.

3.3 Examples of Entanglement Rate Equations

Given two particles a and b in an interacting system, the entanglement rate equation of Eq. (3.28) can be used to build up a complete set of rate equations for the system, in the following way: starting from the equation for $\dot{F}_{ab}$, recursively write down the equations for the entanglement fidelities appearing on the right-hand side of Eq. (3.28). The equations will trace through the interaction network, so that the sets A'_i and B'_i appearing on the right-hand side grow as the iteration procedes. Once those sets overlap, we have $F_{A'_i B'_i} = 1$, and the corresponding term of the rate equation will not introduce any new entanglement fidelities, so the recursion eventually terminates.

Let us use this procedure to derive a set of rate equations for some simple examples. First, consider a chain of N interacting particles, as shown in Fig. 3.3, with Hamiltonian

$$H = \sum_{i=1}^{N-1} H_{i,i+1}. \tag{3.29}$$

Figure 3.3: A chain of interacting particles.

We label the particles from 1 to N, and label particles 1 and N at the ends of the chain by a and b (see Fig. 3.3). If A is the set of particles 1 to n and B is the set of particles $N-m+1$ to N, we will denote by $F_{n,m}$ the entanglement fidelity of the state ρ_{AB} with respect to subsystem ab (Definition 3.1). That is, $F_{n,m}$ is the entanglement fidelity between n adjacent particles at the left and m adjacent particles at the right of the chain.

From the rate equation (Eq. (3.28)) derived in Section 3.2.2, the complete set of rate equations for the entanglement fidelities relative to a and b includes one rate equation for each $F_{n,m}$:

$$\dot{F}_{n,m} = 2\,\|H_{n,n+1}\|_{\mathrm{HS}}\,\sqrt{F_{n,m}}\sqrt{F_{n+1,m}-F_{n,m}} + 2\,\|H_{N-m,N-m+1}\|_{\mathrm{HS}}\,\sqrt{F_{n,m}}\sqrt{F_{n,m}-F_{n,m+1}}, \quad (3.30)$$

where $F_{n,m} = 1$ for $n + m > N$. Graphically, we can represent this by the directed graph shown in Fig. 3.4, where the edges connect fidelities for which the derivative of the fidelity at the start of the edge depends explicitly on the fidelity at the end.

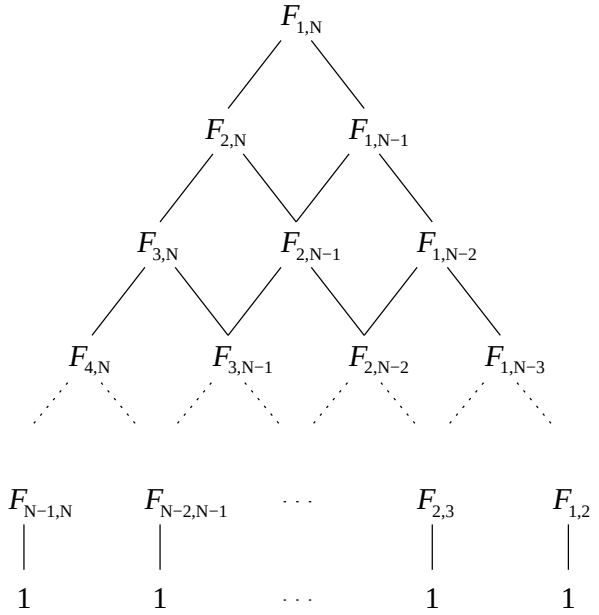

Figure 3.4: The directed graph representing the entanglement rate equations for a chain of N particles.

Now consider the interaction network shown in Fig. 3.5, with Hamiltonian

$$H = H_{a,c} + H_{a,d} + H_{c,b} + H_{b,d} \quad (3.31)$$

The complete set of rate equations for the entanglement fidelities relative to

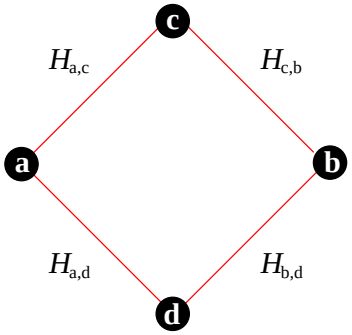

Figure 3.5: A simple, diamond interaction network, with two interaction paths between any two particles.

the particles labelled a and b is

$$\dot{F}_{a,b} = \sqrt{F_{a,b}}\Big(\|H_{ac}\|_{\mathrm{HS}} \sqrt{F_{(ac),b} - F_{a,b}} + \|H_{ad}\|_{\mathrm{HS}} \sqrt{F_{(ad),b} - F_{a,b}} + \|H_{bc}\|_{\mathrm{HS}} \sqrt{F_{a,(bc)} - F_{a,b}} + \|H_{bd}\|_{\mathrm{HS}} \sqrt{F_{a,(bd)} - F_{a,b}}\Big) \tag{3.32a}$$

$$\dot{F}_{(ac),b} = \sqrt{F_{(ac),b}}\Big(\|H_{cb}\|_{\mathrm{HS}} \sqrt{1 - F_{(ac),b}} + \|H_{ad}\|_{\mathrm{HS}} \sqrt{F_{(acd),b} - F_{a,b}}\Big) \tag{3.32b}$$

$$\vdots$$

$$\dot{F}_{(acd),b} = (\|H_{cb}\|_{\mathrm{HS}} + \|H_{db}\|_{\mathrm{HS}}) \sqrt{F_{(acd),b}} \sqrt{1 - F_{(ac),b}}\,, \tag{3.32c}$$

$$\vdots$$

which can be represented graphically by the directed graph shown in Fig. 3.6. Note how this differs from the graph for the chain (Fig. 3.4). For instance, the graph in Fig. 3.6 contains loops, reflecting the fact that the interaction network contains a loop (Fig. 3.5).

3.4 Bounds on Entanglement Flow

Many different, locally-inequivalent states (on the level of operations on the individual particles) can have the same entanglement fidelity, so it is not a priori clear that the curves that saturate the inequalities in the rate equations at all points in time correspond to the evolution that minimizes the time required to reach a given value of the entanglement fidelity. It could be that initially increasing the fidelity slower than allowed by the rate equations

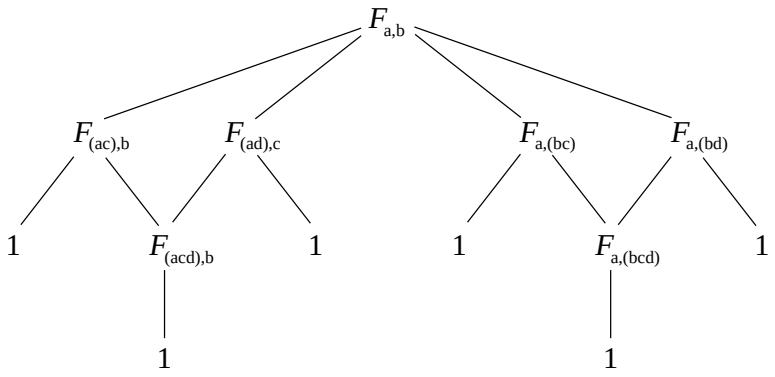

Figure 3.6: The directed graph representing the entanglement rate equations for the diamond network shown in Fig. 3.5.

would enable us to later increase the fidelity at a much faster rate, reducing the time required to reach the desired value of the fidelity.

In fact, this is not the case: the fastest possible evolution that satisfies the entanglement rate equations is that which saturates them at each point in time. Therefore, the rate equations can be used to bound the entanglement fidelities themselves, not only their derivatives: the solutions to the rate equations of Eq. (3.28) with the inequality replaced by an equality provide upper bounds on all possible evolutions of the entanglement fidelities. (Note that this evolution is not necessarily physically possible. The rate equations only provide upper bounds, not corresponding lower bounds. Indeed, in Section 3.5, we will give an example of a system in which the bounds produced by the rate equations are demonstrably not tight.)

Theorem 3.5 (Entanglement Rate Equation Bounds)
The curves $f_{AB}(t)$ that saturate the inequalities in a complete set of rate equations for a system are upper bounds on any evolution of the entanglement fidelities $F_{AB}(t)$ of the system.

Proof The proof is by induction on sets A and B. $f_{AB}(t)$ is the solution of the rate equation (Eq. (3.28)) with the inequality replaced by equality; it is the solution to the differential equation

$$\frac{\mathrm{d}f_{AB}}{\mathrm{d}t} = 2 \sum_{\substack{i \notin A, j \in A; \\ i \notin B, j \in B}} \|H_{ij}\|_{\mathrm{HS}} \sqrt{f_{AB}(t)} \sqrt{f_{A'_i B'_i}(t) - f_{AB}(t)}, \tag{3.33}$$

with $f_{A'_i B'_i} = 1$ if sets A'_i and B'_i overlap, and the same boundary conditions as for the F_{AB}'s. Assume that the $f_{A'_i B;i}(t)$'s *are* upper bounds on $F_{A'_i B;i}(t)$.

That is, for all times t, and for all particles i connected by an interaction to either the set A or the set B,

$$f_{A'_iB'_i}(t) \geq F_{A'_iB'_i}(t). \tag{3.34}$$

If $f_{AB}(t)$ is *not* an upper bound on $F_{AB}(t)$, then $F_{AB}(t)$ must cross $f_{AB}(t)$ at some point $t = t_0$, i.e.

$$F_{AB}(t_0) = f_{AB}(t_0) \tag{3.35a}$$

and

$$\left.\frac{\mathrm{d}^k F_{AB}}{\mathrm{d}t^k}\right|_{t_0} > \left.\frac{\mathrm{d}^k f_{AB}}{\mathrm{d}t^k}\right|_{t_0} \tag{3.35b}$$

for some k^{th}-order derivative of $F_{AB}(t)$ at t_0.

First, take the case in which Eq. (3.35b) involves the first-order derivative, $k = 1$. $F_{AB}(t)$ must still satisfy the inequality in the rate equation (Eq. (3.28)). Using that, along with Eq. (3.35a), Eq. (3.34) and Eq. (3.33), gives

$$\begin{aligned}
\dot{F}_{AB}(t_0) &\leq 2\sum_{\substack{i\notin A, j\in A;\\ i\notin B, j\in B}} \sqrt{F_{AB}(t_0)}\sqrt{F_{A'_iB'_i}(t_0) - F_{AB}(t_0)} && \text{(3.36a)}\\
&= 2\sum_{\substack{i\notin A, j\in A;\\ i\notin B, j\in B}} \sqrt{f_{AB}(t_0)}\sqrt{F_{A'_iB'_i}(t_0) - f_{AB}(t_0)} && \text{(3.36b)}\\
&\leq 2\sqrt{f_{AB}(t_0)}\sqrt{f_{A'_iB'_i}(t_0) - f_{AB}(t_0)} && \text{(3.36c)}\\
&= \dot{f}_{AB}(t_0), && \text{(3.36d)}
\end{aligned}$$

which contradicts Eq. (3.35b) for $k = 1$. Therefore, assuming the $f_{A'_iB'_i}(t)$ bound the $F_{A'_iB'_i}(t)$ from above, $F_{AB}(t)$ can not cross $f_{AB}(t)$.

Now take the case in which Eq. (3.35b) relating the derivatives of f_{AB} and F_{AB} involves a higher-order derivative, $k > 1$. There then exists some $\delta > 0$ for which, for any $0 < \epsilon < \delta$, the first-order derivative obeys

$$\dot{F}_{AB}(t_0 + \epsilon) > \dot{f}_{AB}(t_0 + \epsilon) + O(\epsilon^{k-1}). \tag{3.37}$$

As all derivatives of $F_{AB}(t)$ and $f_{AB}(t)$ up to oder k are equal at t_0, we also have

$$F_{AB}(t_0 + \epsilon) = f_{AB}(t_0 + \epsilon) + O(\epsilon^k). \tag{3.38}$$

But, from Eq. (3.28), and using Eq. (3.38), Eq. (3.34) and Eq. (3.33), we have

$$\dot{F}_{AB}(t_0+\epsilon) \leq 2 \sum_{\substack{i\notin A, j\in A;\\ i\notin B, j\in B}} \sqrt{F_{AB}(t_0+\epsilon)}\sqrt{F_{A'_iB'_i}(t_0+\epsilon) - F_{AB}(t_0+\epsilon)} \tag{3.39a}$$

$$= 2 \sum_{\substack{i\notin A, j\in A;\\ i\notin B, j\in B}} \sqrt{f_{AB}(t_0+\epsilon)}\sqrt{F_{A'_iB'_i}(t_0+\epsilon) - F_{AB}(t_0+\epsilon)} + O(\epsilon^k) \tag{3.39b}$$

$$\leq 2 \sum_{\substack{i\notin A, j\in A;\\ i\notin B, j\in B}} \sqrt{f_{AB}(t_0+\epsilon)}\sqrt{f_{A'_iB'_i}(t_0+\epsilon) - f_{AB}(t_0+\epsilon)} + O(\epsilon^k) \tag{3.39c}$$

$$= \dot{f}_{AB}(t_0) + O(\epsilon^k). \tag{3.39d}$$

As this upper bound on $\dot{F}_{AB}(t_0+\epsilon)$ is second order in ϵ, whereas the lower bound of Eq. (3.37) is first order in ϵ, we can always choose ϵ small enough to produce a contradiction between the two. Therefore, given our initial assumption that the $f_{A'_iB'_i}(t)$ bound the $F_{A'_iB'_i}(t)$ from above, we again have that $F_{AB}(t)$ can not cross $f_{AB}(t)$. Therefore, if all the $f_{A'_iB'_i(t)}$ are upper bounds on the $F_{A'_iB'_i(t)}$, then $f_{AB}(t)$ is an upper bound on $f_{AB}(t)$.

The recursive procedure for building up a complete set of rate equations (described at the start of this section) terminates when it reaches sets A'_i and B'_i that overlap, in which case $F_{A'_iB'_i} = 1$ by definition. The initial step in the induction is trivially provided by the corresponding definition $f_{A'_iB'_i} = 1$, since clearly $f_{A'_iB'_i} \geq F_{A'_iB'_i}$ in this case, fulfilling the inductive assumption of Eq. (3.34) and completing the proof. □

3.5 Long-Distance Entanglement Generation

How quickly can entanglement be created over large distances? The question is both theoretically interesting and experimentally important. Many quantum information processing tasks require entanglement, and the faster this can be produced, the less the system will suffer from decoherence. Quantum computing algorithms often generate large amounts of entanglement during their execution, so determining how fast entanglement can be generated can also provide bounds on algorithm complexity [Ambainis, 2002].

Imagine for example a chain of interacting spins, initially in a completely separable state, where the goal is to create a maximally entangled state between the end spins. How does the time required to create the maximally entangled state scale with the length of the chain? Whereas the corresponding question when sending information along the chain can immediately be

answered by appealing to causality (the scaling must be linear), creating entanglement arbitrarily fast does not violate any causality principle. The entanglement can not be used to transfer information without sending additional classical information, and this classical communication guarantees that any causality constraints are satisfied.

In Section 3.5.1, we review some existing entanglement generation schemes in the context of the entanglement rate equations derived in Section 3.2.2. Then, in Section 3.5.2, we use the rate equations to derive a universal limit on how fast entanglement can be generated, or more precisely, how the time required to entangle two particles scales with the size of the system.

3.5.1 Entanglement Generation Schemes

How fast entanglement can be generated depends, of course, on how we are able to manipulate the system, for example the spins in a spin chain. We can identify three basic scenarios, differing in the amount of local control that is allowed:

LOCC Local operations on individual spins (including measurements), and classical communication from one spin to another are allowed "for free"; they do not count towards the evolution time.

Fast local unitaries Local unitary operations on individual spins are allowed "for free", but measurements and classical communication are not.

No local control No local operations are allowed; the only control over the system is the ability to switch on the interactions initially, and switch them off again at some later time.

The LOCC scenario assumes we neglect the time required for classical communication of the measurement outcomes between different spins. This can be justified on theoretical grounds, since classical communication can not create entanglement, and it makes sense to consider the interactions as the resource. In many physical implementations, it is also reasonable on pragmatic grounds: classical communication is usually much easier to implement than quantum operations.

It turns out that measurement is a very powerful resource. In the LOCC scenario, the end qubits in a qubit chain (a chain of spin-$\frac{1}{2}$ particles) can be maximally entangled in a time *independent* of the length of the chain. Though not discussed in the context of entanglement generation, Briegel and

Raussendorf [2001] showed that a cluster state can be created in a chain in constant time, and local measurements on a cluster state allow a Bell-state to be projected out on any desired pair of qubits, including the end pair [Verstraete et al., 2004].

However, if the interactions are really the *only* non-local resource, then any classical communication must also be implemented via the interactions, and local measurements are of no benefit. The LOCC scenario then becomes equivalent to the fast local unitary scenario.

If we can apply local unitary operations on any spin in the chain, then we can efficiently simulate evolution under any Hamiltonian [Nielsen et al., 2002, Wocjan et al., 2002]. Again, it is reasonable to discount local resources, which in this scenario means neglecting the time required to carry out the local unitaries (hence the "fast local unitary" approximation). And again, this can also be justified on physical grounds, since local unitaries are typically much faster than interactions. In Section 3.5.1, we will use the entanglement rate equations to derive a lower bound on the time required to entangle the end spins of a chain in this scenario.

Khaneja and Glaser [2002] have developed an interesting protocol for state transfer along qubit chains, in the context of NMR spectroscopy, which can easily be transformed into an entanglement generation scheme that satisfies the requirements of the fast local unitary scenario. First the middle qubits are entangled, then the state of each middle qubit is encoded into a three-qubit state. The encoded states are transferred along the chain towards the ends, where they are decoded again. The protocol requires local unitaries to be applied at discrete points in time during the evolution.

The evolution of the entanglement fidelities is shown in Fig. 3.7, clearly showing the initial encoding phase followed by the encoded states being moved step-by-step along the chain. It achieves a surprising three-fold speed-up over the trivial swapping protocol for entanglement generation in a chain with the same interactions (entangle the middle qubits; move to the ends by swapping), though the scaling of the time with the length of the chain is still linear, as in the trivial protocol.

Finally, we may have no local control over the spins, only retaining the ability to switch on and off all the interactions in the chain. Christandl et al. [2004] developed a state-transfer protocol for qubit chains in this scenario, and Yung et al. [2003] have given a simple extension to entanglement generation. The only local control required is fixing the coupling strengths between different qubits, which must be inhomogeneous. Figure 3.8 shows the entanglement dynamics for the odd chain-length protocol of Yung et al. [2003].

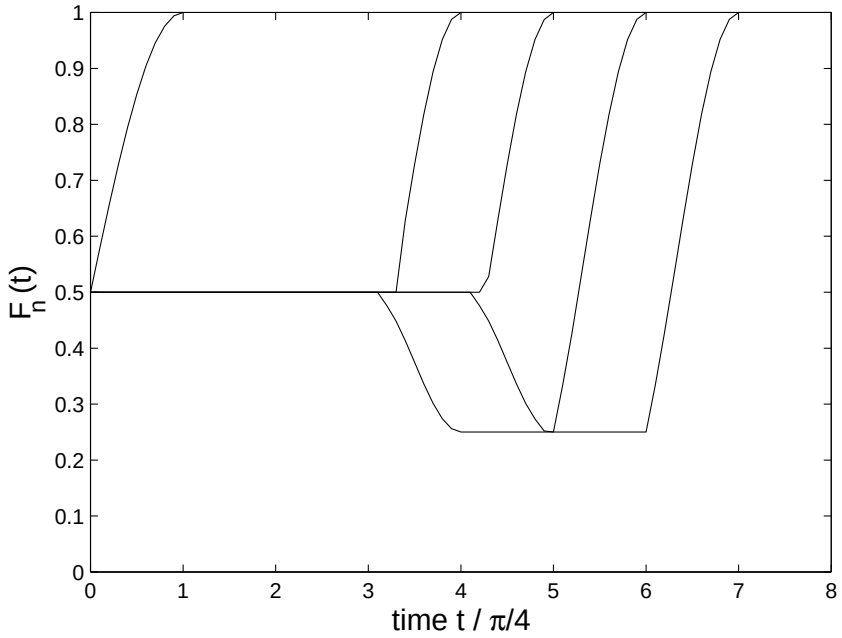

Figure 3.7: Entanglement dynamics in the entanglement generation protocol based on the state transfer scheme of Khaneja and Glaser [2002], for a chain of ten qubits. Successive curves show the evolution of entanglement fidelities F_5 through F_1, numbered as in Section 3.3.

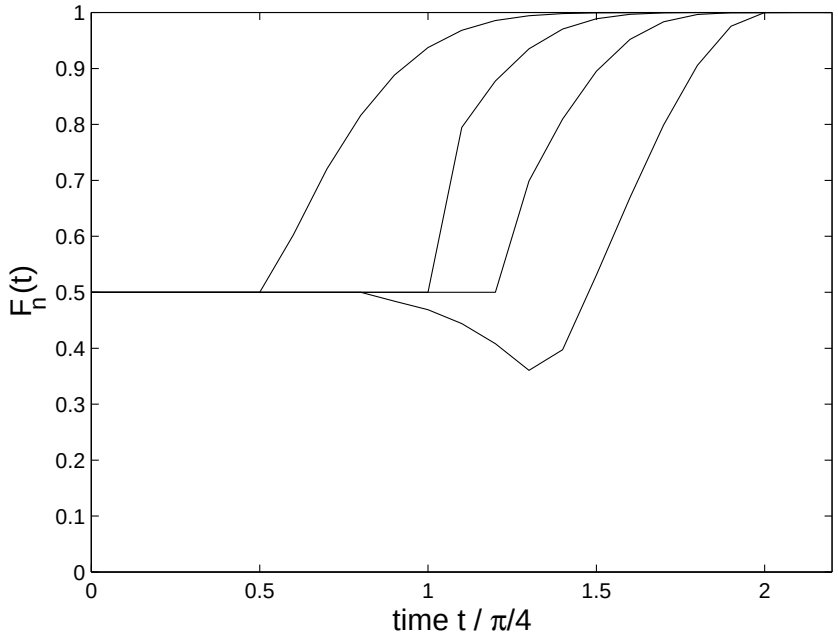

Figure 3.8: Entanglement dynamics in the entanglement generation scheme for odd chain lengths from Yung et al. [2003], here for nine qubits. Successive curves show the evolution of entanglement fidelities F_4 through F_1, numbered as in Section 3.3. (Note that the times can not be compared with those in Fig. 3.7, since interaction strengths in Christandl et al. [2004] and Yung et al. [2003] are not normalized.)

Comparing with Fig. 3.7, it is clear that the entanglement dynamics are quite different. The scheme does not involve moving a state step-by-step along the chain. Rather, the entangled state spreads out over the entire chain, then refocusses at the ends. If the strongest coupling strength is normalized to some fixed value, then the time to entangle the end spins using this scheme again scales linearly with the length of the chain.*

Osborne and Linden have also developed a protocol for state transfer in qubit chains, which could be adapted to entanglement generation, involving limited local control over a vanishingly small (in the limit of large chain lengths) number of qubits at each end of the chain [Osborne and Linden, 2003].

3.5.2 Bounds on Entanglement Generation

In this section, we use the entanglement rate equations to prove a lower bound on the time required to entangle the ends of a spin chain, in the fast local unitary scenario described in the previous section. The bound is universal: it is valid for all interaction Hamiltonians, and any sequence of local unitary operations. It can also easily be applied to general interaction networks, though the bound may be rather weak in that case.

There is a trivial linear *upper* bound on the scaling with the length of the chain of the time needed to entangle the ends, achieved for example by simply entangling the first two spins, then moving half of the entangled pair along the chain using swap operations. On the other hand, the entanglement rate equations upper-bound the evolution of the entanglement fidelities, and the entanglement fidelity of a state is equal to 1 if and only if that state is maximally entangled. So solving the rate equations for a spin chain would provide a *lower* bound on the time required to create a maximally entangled state.

Unfortunately, the set of non-linear coupled differential equations arising from the rate equations (Eq. (3.33)) has no known analytic solution. Solving numerically would provide numerical lower bounds, but we are interested in the asymptotic scaling behaviour, which requires an analytic solution. Instead, we will derive analytic upper bounds on the solutions, which we use to give an upper bound on the asymptotic scaling.

First, number the spins in the chain from 1 to N, and label the spins 1 and N at the ends of the chain by a and b (see Fig. 3.3). Let A and B be sets

*Christandl et al. [2004] state that the time required to entangle the end qubits is independent of the length of the chain. For this to be the case, the interaction strengths must be increased for longer chains. If the interaction strengths are normalized, the scaling is linear.

of n neighbouring spins at the two ends of the chain, i.e. A contains spins 1 to n and B contains spins $N-n+1$ to N. Let A' and B' be similar sets of $n+1$ neighbouring spins. The rate equation for the entanglement fidelity with respect to ab of the AB system is

$$\dot{F}_{AB} \leq 2\,\|H_{n,n+1}\|\,\sqrt{F_{AB}}\sqrt{F_{A'_{n+1}B}-F_{AB}} + 2\,\|H_{N-n,N-n+1}\|\,\sqrt{F_{AB}}\sqrt{F_{AB'_{N-n}}-F_{AB}}, \quad (3.40)$$

where the set A'_i, defined in Section 3.2.2, contains spin i in addition to all the spins in A, similarly for set B'_i, and we have dropped the explicit time-dependence from the notation for the entanglement fidelities.

A'_{n+1} and B'_{N-n} are subsystems of A' and B', respectively, as are A and B. From the discussion in Section 3.2.1, the entanglement fidelities $F_{A'_{n+1}B}$ and $F_{AB_{N-n}}$ with respect to the common subsystem ab are always less than or equal to $F_{A'B'}$:

$$F_{A'_{n+1}B}(t) \leq F_{A'B'}(t), \quad (3.41a)$$

$$F_{AB'_{N-n}}(t) \leq F_{A'B'}(t) \quad (3.41b)$$

at any time t during the evolution. If we assume that the strengths of all the interactions in the spin chain are equal to the same value $\|H_{ij}\|$, then from Eq. (3.40) we have

$$\dot{F}_{AB}(t) \leq 4\,\|H_{ij}\|\,\sqrt{F_{AB}(t)}\sqrt{F_{A'B'}(t)-F_{AB}(t)}. \quad (3.42)$$

We now relabel the entanglement fidelity for system AB where A and B contain n spins by F_n, so that there are $\lfloor N/2 \rfloor$ fidelities in total, F_1 to $F_{\lfloor N/2 \rfloor}$ ($\lfloor \bullet \rfloor$ denotes rounding down to the nearest integer). F_1 is then the entanglement fidelity of the two spins a and b, $F_{\lfloor N/2 \rfloor}$ is the entanglement fidelity of the entire chain split into two down the middle (with one particle in the middle left over if N is odd), and we define $F_{\lfloor N/2 \rfloor+1} = 1$. From Eq. (3.42), a complete set of rate equations is then given by the coupled set of differential equations

$$[equations]\dot{F}_n(t) \leq 4\,\|H_{ij}\|\,\sqrt{F_n(t)}\sqrt{F_{n+1}(t)-F_n(t)}, \quad n = 1 \ldots \lfloor N/2 \rfloor. \quad (3.43)$$

If the chain starts in a completely separable state $|\psi_1\rangle\,|\psi_2\rangle \cdots |\psi_N\rangle$, then all entanglement fidelities are initially equal to 1/2, since the entanglement fidelity of a separable pure state is 1/2, giving us the boundary conditions

$$F_n(0) = \frac{1}{2}, \quad n = 1 \ldots \lfloor N/2 \rfloor. \quad (3.44)$$

Let $f_n(t)$ be the solutions of the set of differential equations given by replacing the inequality in Eq. (3.43) with an equality:

$$\dot{f}_n(t) = 4 \|H_{ij}\| \sqrt{f_n(t)} \sqrt{f_{n+1}(t) - f_n(t)}, \quad n = 1 \ldots \lfloor N/2 \rfloor, \tag{3.45}$$

with $f_{\lfloor N/2 \rfloor + 1}(t) = 1$ and the same boundary conditions $f_n(0) = 1/2$ as for the entanglement fidelities themselves. Applying Theorem 3.5 of Section 3.4, the $f_n(t)$ are bounds on the entanglement fidelities $F_n(t)$:

$$F_n(t) \leq f_n(t), \text{ for all times } t. \tag{3.46}$$

Now, since the entanglement fidelity, and hence $f_n(t)$, is bounded by 1, from Eq. (3.45) we have

$$\dot{f}_n(t) \leq 4 \|H_{ij}\| \sqrt{f_{n+1}(t) - f_n(t)}. \tag{3.47}$$

We can use the argument of Theorem 3.5 to prove that, if $u_n(t)$ is an upper bound on $f_n(t)$, then the solution $u_{n-1}(t)$ to the differential equation

$$\dot{u}_{n-1}(t) = 4 \|H_{ij}\| \sqrt{u_n(t) - u_{n-1}(t)} \tag{3.48}$$

(where we have dropped the initial square-root) is an upper bound on $f_{n-1}(t)$, so long as the boundary conditions on u_{n-1} satisfy

$$u_{n-1}(0) \geq f_{n-1}(0). \tag{3.49}$$

Now assume there is a $u_n(t)$ of the form

$$u_n(t) = \frac{4 \|H_{ij}\|^2 t^2}{a_n} + \frac{1+\epsilon}{2} \tag{3.50}$$

that is an upper bound on $f_n(t)$ for some positive constants a_n and ϵ. The differential equation Eq. (3.48) for $u_{n-1}(t)$ then has a solution of the same form as $u_n(t)$ (as can be seen by direct substitution) with the same ϵ, and a coefficient a_{n-1} given by the recursion relation

$$a_{n-1} = \frac{a_n}{2} + \frac{a_n}{2} \sqrt{1 + \frac{4}{a_n}}. \tag{3.51}$$

Since $u_{n-1}(0) = (1+\epsilon)/2$, which is greater than the boundary condition $f_{n-1}(0) = F_{n-1}(0) = 1/2$, it satisfies the requirement in Eq. (3.49). Therefore $u_{n-1}(t)$ is an upper bound on $f_{n-1}(t)$ by the argument above.

All that remains is the initial step in the induction: that there is indeed a bound $u_{\lfloor N/2 \rfloor}(t)$ on $f_{\lfloor N/2 \rfloor}(t)$ with the form assumed in Eq. (3.50), for some

constants $a_{\lfloor N/2 \rfloor}$ and ϵ. (Recall that $F_{\lfloor N/2 \rfloor}$ is the entanglement fidelity of the entire spin chain split into two halves.) Fortunately, the differential equation for $f_{\lfloor N/2 \rfloor}$ from Eq. (3.45) (recall that $f_{\lfloor N/2 \rfloor + 1} = 1$ by definition),

$$\dot{f}_{\lfloor N/2 \rfloor}(t) = 4 \left\| H_{ij} \right\| \sqrt{f_{\lfloor N/2 \rfloor}(t)} \sqrt{1 - f_{\lfloor N/2 \rfloor}(t)}, \tag{3.52}$$

can be solved analytically [Kamke, 1961]. One set of solutions has the form

$$f_{\lfloor N/2 \rfloor}(t) = \sin^2(2 \left\| H_{ij} \right\| t + \phi) \tag{3.53}$$

with ϕ an arbitrary constant. There is also a trivial solution: $f_{\lfloor N/2 \rfloor}(t) = 1$. Since the chain starts in a completely separable pure state, the initial condition is $f_{\lfloor N/2 \rfloor}(0) = 1/2$, and the solution we require is

$$f_{\lfloor N/2 \rfloor}(t) = \begin{cases} \sin^2(2 \left\| H_{ij} \right\| t + \pi/4) & 2 \left\| H_{ij} \right\| t \leq \pi/4 \\ 1 & 2 \left\| H_{ij} \right\| t > \pi/4. \end{cases} \tag{3.54}$$

The two parts of the solution reflect the fact that, once the entanglement fidelity has reached its maximum value of 1, there is nothing to be gained by further evolution, and the interactions affecting $F_{\lfloor N/2 \rfloor}$ (namely the interaction connecting the two halves of the chain) should be switched off.

Knowing an explicit solution for $f_{\lfloor N/2 \rfloor}(t)$, it is easy to find a bound $u_{\lfloor N/2 \rfloor}(t)$ with the appropriate form. To make the algebra simpler, we can upper-bound $f_{\lfloor N/2 \rfloor}(t)$ by $2 \left\| H_{ij} \right\| t + 1/2$. Thus a $u_{\lfloor N/2 \rfloor}(t)$ with the form given in Eq. (3.50) that satisfies $u_{\lfloor N/2 \rfloor}(t) \geq 2 \left\| H_{ij} \right\| t + 1/2$ will suffice as the initial step in the induction. Solving this inequality, namely

$$\frac{4 \left\| H_{ij} \right\|^2 t^2}{a_{\lfloor N/2 \rfloor}} + \frac{1 + \epsilon}{2} \geq 2 \left\| H_{ij} \right\| t + \frac{1}{2}, \tag{3.55}$$

leads to the relation $a_{\lfloor N/2 \rfloor} \leq 2\epsilon$. Any positive $a_{\lfloor N/2 \rfloor}$ and ϵ satisfying this will give an appropriate $u_{\lfloor N/2 \rfloor}(t) \geq f_{\lfloor N/2 \rfloor}(t) \geq F_{\lfloor N/2 \rfloor}(t)$, and will guarantee that $u_{\lfloor N/2 \rfloor}(0) = (1+\epsilon)/2 \geq f_{\lfloor N/2 \rfloor}(0) = F_{\lfloor N/2 \rfloor}(0) = 1/2$. Thus we have shown that an upper bound on $F_{\lfloor N/2 \rfloor}(t)$ with the appropriate form exists, which completes the proof. For neatness, we can let $\epsilon \to 0$, so that $u_n(0) \to F_n(0) = 1/2$ and $a_{\lfloor N/2 \rfloor} \to 0$ (as used to give the curve $u_1(t)$ shown in Fig. 3.9).

Solving $u_1(t) = 1$ will give a *lower* bound on the time required for $f_1(t)$ to reach 1, which is itself a lower bound on the time T_{ent} required for the entanglement fidelity of the end two qubits F_1 to reach 1, or equivalently, for the end qubits to become maximally entangled. We are interested in the scaling of T_{ent} for large chain lengths, when the coefficient a_1 becomes large. Rather than solving $u_1(t) = 1$ explicitly to obtain the bound, it is simpler

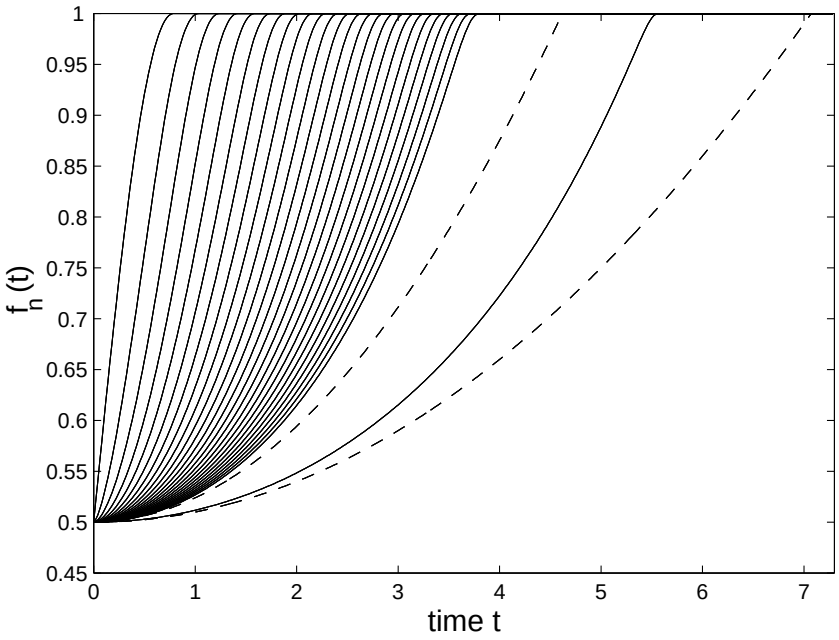

Figure 3.9: Numerically calculated entanglement fidelity curves $f_n(t)$ saturating the inequalities in the entanglement rate equations for a chain (Eq. (3.43)). The final solid curve is for the entanglement fidelity $f_1(t)$ of the end qubits in a chain of lengh $N = 100$, and the dashed curves show the corresponding upper and lower bounds $u_1(t)$ and $l_1(t)$.

to Taylor expand the square-root in the recursion relation of Eq. (3.51) in order to show that the relation asymptomatically approaches $a_{n-1} = a_n + 1$, or equivalently $a_1 = a_{\lfloor N/2 \rfloor} + \lfloor N/2 \rfloor$, as $N \to \infty$. Thus, for large N, choosing $\epsilon \to 0$ so that $a_{\lfloor N/2 \rfloor} \to 0$ as before, the equation we must solve tends to

$$u_1(t) = \frac{4 \|H_{ij}\|^2 t^2}{\lfloor N/2 \rfloor} + \frac{1+\epsilon}{2} = 1. \tag{3.56}$$

Therefore, the bound on T_{ent} tends to

$$T_{\text{ent}} \geq \frac{1}{2 \|H_{ij}\|} \sqrt{\frac{\lfloor N/2 \rfloor}{2}}, \tag{3.57}$$

a square-root scaling with chain length (see Fig. 3.10).

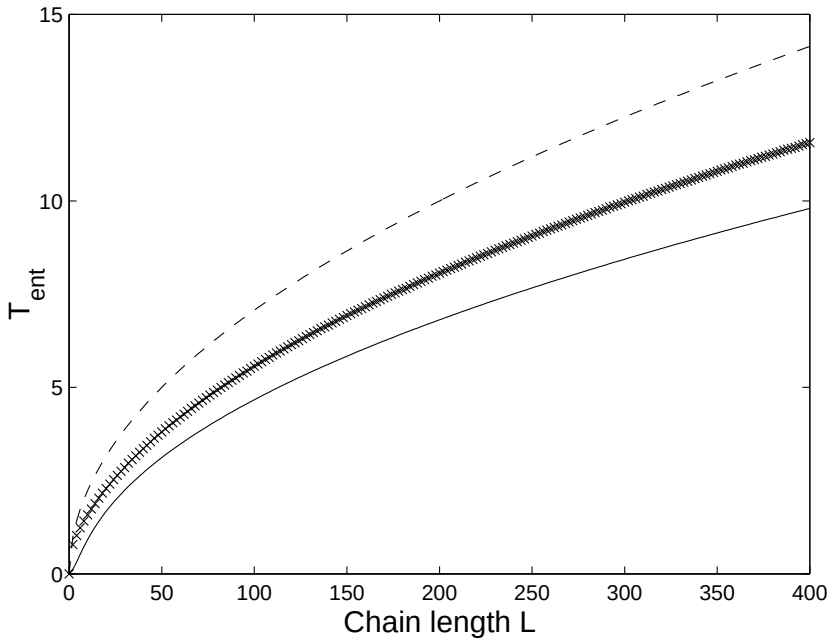

Figure 3.10: Scaling with chain-length N of the time T_{ent} required to create a maximally entangled state between the ends. The points show numerical results obtained by saturating the rate equations (Eq. (3.43)) with all interaction strengths $\|H_{ij}\| = /2$. The solid and dashed curves show the analytic lower and upper bounds, $T_{\text{ent}} \geq \sqrt{\lfloor N/2 \rfloor /2}$ and $T_{\text{ent}} \leq \sqrt{\lfloor N/2 \rfloor}$ respectively, from Eq. (3.57) and Eq. (3.59).

We have loosened many an inequality during the proof of this square-root bound. Could the rate equations of Eq. (3.43) give a tighter bound? We can use essentially the same proof with the inequalities reversed to prove that a square-root bound is the best that can be obtained.

Instead of using $F_n(t) \leq 1$ to substitute for the first square-root in the differential equation for $f_n(t)$ (Eq. (3.45)), we use $F_n(t) \geq 1/2$, which is valid when the $F_n(t)$ saturate the inequalities in the rate equations, that is when $F_n(t) = f_n(t)$. Then, solutions of

$$\dot{l}_{n-1}(t) = 2\sqrt{2}\,\|H_{ij}\|\,\sqrt{l_n(t) - l_{n-1}(t)} \tag{3.58}$$

are *lower* bounds on $f_{n-1}(t)$. We can rescale the time $\tau = t/\sqrt{2}$ so that the differential equation for $l_n(\tau)$ has the same form as that for $u_n(t)$ in the previous proof (Eq. (3.48)). Assuming solutions of the form $l_n(\tau) = \tau^2/a_n$, solving the resulting recursion relation, and proving there is a *lower* bound on $f_n(t)$ of the appropriate form, leads to an *upper* bound on the scaling, for any evolution saturating the rate equations. For large chain lengths, the bound tends to

$$T_{\text{ent}} \leq \frac{\sqrt{\lfloor N/2 \rfloor}}{2\,\|H_{ij}\|}, \tag{3.59}$$

which is also a square-root scaling. Therefore, the square-root bound we have derived is, up to a $\sqrt{2}$ numerical factor, the best that can be obtained from Eq. (3.43) (see Fig. 3.10). However, as Eq. (3.43) are weaker than the original rate equations (Eq. (3.40)), it is possible that they could provide tighter bounds by another method.

In fact, subsequent to the publication of these results in Cubitt et al. [2005], Bravyi et al. [2006] proved a *linear* lower-bound on the scaling of T_{ent} with chain length. Whilst their approach did not involve deriving anything analogous to the entanglement rate equations, it does suggest that the inequality in the rate equation (Eq. (3.28)) is not tight. Whether or not it is possible to derive a tighter version of the entanglement rate equations (from which the linear scaling follows) remains an interesting open problem.

Chapter 4

Engineering Correlation and Entanglement Dynamics

How does entanglement, or more generally, how do correlations evolve in large, interacting systems? How easily can the evolution be controlled and manipulated? The entanglement rate equations derived in Chapter 3 provide limits that are necessarily obeyed by any physically realizable evolution. In Section 3.5 of the same chapter, we considered various scenarios that differed in the amount of control there was over the system. In this chapter, we will investigate what evolution actually occurs in an explicit example of an interacting, many-body system, and demonstrate that we can achieve a surprising amount of control over the evolution of correlations, even when control over the system is severely limited.

In Section 4.1, we discuss the practical motivation behind the study of correlation and entanglement dynamics in large systems, in particular the connection with quantum repeaters. The exact solution for the time-evolution of the simple model system introduced in Section 4.2 leads, in Section 4.3, to analytical results for the evolution of string (Section 4.3.1) and two-point (Section 4.3.2) correlations. These two-point connected correlation functions provide a lower bound on the localizable entanglement, the relevant figure of merit for quantum repeaters. We also derive an alternative, though somewhat less interesting, analytic bound using Grassmann calculus in Section 4.4. Finally, in Section 4.5, we use these analytic results to investigate how the correlation and entanglement dynamics can be controlled and engineered, even when control over the system is limited to changing its two global, external parameters.

4.1 Physical and Practical Motivation

Practical applications of entanglement often involve distributing different parts of an entangled state to different parties, each of whom may be spatially separated from the others. This raises the problem of how to reliably distribute entangled states across large distances. If the quantum channel connecting two parties (a fibre optic cable, for example, or even free space) is noisy, the entanglement will typically be destroyed before it arrives. The standard solution is to use a *quantum repeater* [Briegel et al., 1998, Dür et al., 1999a].

A quantum repeater consists of a chain of nodes that exchange entanglement with their neighbours, then use entanglement swapping [Żukowski et al., 1993] to successively transfer the entanglement to the ends of the chain. Because the entangled pairs are sent over shorter distances, and entanglement purification can be used after each step to recover a smaller number of higher-quality entangled pairs from a larger number of imperfect pairs before proceeding to the next step, this allows entanglement to be distributed over arbitrarily large distances (at the expense of introducing more intermediate nodes).

Physically, the nodes of the quantum repeater will be quantum systems (e.g. atoms in optical cavities) that interact with their neighbours according to some Hamiltonian (e.g. by exchanging photons) [Cirac et al., 1997, Kraus and Cirac, 2004, van Enk et al., 1998]. Therefore, instead of actively exchanging and swapping entanglement, it has been suggested that correlations inherent in the ground state of this interacting system could be used to distribute long-range entanglement ([Verstraete et al., 2004]). However, cooling to the ground state is an unrealistic prospect in many systems. Why not instead use the non-equilibrium *dynamics* to distribute the entanglement?

We can model the physical quantum repeater by a spin chain. Clearly, if we want to distribute entanglement using the system's dynamics, we must initially prepare it in a non-equilibrium state, that is, one that is not an eigenstate of the Hamiltonian. The state must therefore contain excitations. It is instructive to consider what happens if a single, localized, low-energy excitation is created on top of an equilibrium state, for example by flipping a single spin. Since the low-energy excitations take the form of spin waves, this excitation will create a spin wave that propagates along the chain. Thus the correlation and entanglement dynamics produced by a single, localized excitation can be understood as nothing other than propagation of spin waves.

The spin-wave propagation is completely determined by the dispersion relation given by the system's spectrum. The form of the dispersion relation will typically depend on external, physical parameters of the system (e.g. the strength of an external magnetic field). Thus already in this setup, we

can manipulate the external parameters to control the dispersion relation, and hence control the propagation of correlations. For example, changing the gradient of the dispersion relation will change the speed at which the correlations propagate.

However, as just described, this idea of using the dispersion relation as a means to control correlation dynamics is unrealistic. The ground state will typically be highly correlated and difficult to prepare, and it requires precise local control over the system in order to create the local excitation and break the translational symmetry. With that level of local control, more sophisticated quantum-repeater setups are already possible. Moreover, it is not clear that the correlations will remain localised; they are likely to disperse rapidly as they propagate.

Therefore, we will extend the idea to systems prepared in translationally invariant, uncorrelated initial states, which can be created more easily. For example, the fully polarised state with all spins down, $|\downarrow\rangle|\downarrow\rangle\ldots|\downarrow\rangle$, can be prepared by applying a large, external magnetic field in the $-z$ direction. As the initial state will be far from the ground state, it will contain many excitations. The correlation dynamics is then the result of the propagation and interference of a large number of spin waves at many different frequencies. Nonetheless, we will show analytically for a simple, example spin model that the system can be engineered so that correlations propagate in well-defined, localised wave packets, with little dispersion. The external parameters can then be used to control the propagation of these correlation packets.

4.1.1 Localizable Entanglement

The natural figure of merit for a quantum repeater (and other systems which distribute entanglement via a sequence of nodes) is the *localizable entanglement* L of its state before measurements are performed [Popp et al., 2005]. It is defined as the maximum entanglement that can be created between two nodes (the end two, for example) using any local operations, including measurements, on the other nodes, averaged over all possible measurement outcomes. This definition is not complete until we specify which entanglement measure is used to quantify the entanglement of the pure states produced by the measurements, and which local measurement operations (von Neumann measurements, positive operator value measures, etc.) are allowed.

As we will deal with systems of qubits, we will use the localizable concurrence, restricted to von Neumann measurements. (The concurrence of a two-qubit state was defined in Section 2.1.2 of Chapter 2.) Note that any other definition of the localizable concurrence (e.g. allowing more general POVM measurements) will be bounded from below by this, as von Neumann

measurements are a special case of any more general set of measurements.

Definition 4.1 (Localizable Concurrence)
The localizable concurrence between qubits 1 *and* N *of an* N*-qubit state* $|\psi\rangle$ *is defined by*

$$L(\psi) = \max_{\{\mathcal{P}_i^{s_i}\}} \left(\sum_{\substack{s_2 \ldots s_{N-1} \\ =0,1}} C\Big(\bigotimes_{i=2}^{N-1} \mathcal{P}_i^{s_i} \, |\psi\rangle \Big) \right), \tag{4.1}$$

where $\{\mathcal{P}_i^{s_i}\}_{s_i=0,1}$ *is a complete set of projectors on qubit* i *defining a von Neumann measurement on that qubit, the maximization is over all possible sets of projectors (all von Neumann measurements), and the concurrence* C *is given by Definition 2.1.*

4.2 Time-Evolution in the XY–Model

As a simple, yet reasonably realistic example of an interacting, many-body system, we will consider the XY–model for a chain of N spin-$\frac{1}{2}$ particles, which we label sequentially:

$$H_{XY} = -\frac{1}{2} \sum_{l=1}^{N} \left(\frac{1+\gamma}{2} \sigma_l^x \sigma_{l+1}^x + \frac{1-\gamma}{2} \sigma_l^y \sigma_{l+1}^y + \lambda \sigma_l^z \right), \tag{4.2}$$

where the σ's are the usual Pauli operators acting on the sites denoted by their subscripts. The parameter λ can be interpreted as the strength of an external magnetic field aligned along the z–direction, whilst γ controls the anisotropy of the interactions. (Note that, following the standard spin-model convention, throughout this chapter tensor products will be denoted by juxtaposing operators, without any explicit tensor product symbol "$\otimes$" in between, and operators act as the identity on systems that are not indicated explicitly by subscripts. Thus $\sigma_l^x \sigma_{l+1}^x$ denotes $\mathbb{1}_{i<l} \otimes \sigma_l^x \otimes \sigma_{l+1}^x \otimes \mathbb{1}_{j>l+1}$.)

As discussed in Section 4.1, we start the system in a translationally invariant, completely separable state. For example, the state with all spins down: $|\psi\rangle = |\downarrow\rangle|\downarrow\rangle \cdots |\downarrow\rangle$, is easily prepared by initially applying a large, external magnetic field, corresponding to setting the parameter λ to a large, negative value. Suddenly reducing or switching off the magnetic field will then start the system evolving.

The XY–Hamiltonian can be diagonalized by a well-known sequence of Jordan-Wigner, Fourier, and Bogoliubov transformations [Lieb et al., 1961]. First, the spin-operators are transformed into fermionic operators using the Jordan-Wigner transformation. We will use a non-standard (though

equivalent) formulation involving Majorana rather than fermionic operators. For each site l in the spin-chain, define the Majorana operators

$$x_l = \prod_{m<l} \sigma_m^z \sigma_l^x, \quad p_l = \prod_{m<l} \sigma_m^z \sigma_l^y \tag{4.3}$$

which are Hermitian and obey the canonical anti-commutation relations

$$\{x_m, x_n\} = \{p_n, p_m\} = 2\delta_{m,n}, \tag{4.4a}$$
$$\{x_m, p_n\} = 0. \tag{4.4b}$$

Rewriting the Hamiltonian in terms of these operators corresponds to applying the Jordan-Wigner transformation, but in terms of Majorana fermions rather than the usual Jordan-Wigner fermions c_l, $c_l^\dagger$ (the two are related by $x_l = c_l^\dagger + c_l$ and $p_l = (c_l^\dagger - c_l)/i$):

$$H_{XY} = \frac{i}{2} \sum_l \left(\frac{1+\gamma}{2} p_l x_{l+1} - \frac{1-\gamma}{2} x_l p_{l+1} + \lambda x_l p_l \right) \tag{4.5}$$

As the system is translationally invariant, the next step is to apply a Fourier transform:

$$x_k = \sum_l \left(\cos\left(\frac{2\pi kl}{N}\right) x_l - \sin\left(\frac{2\pi kl}{N}\right) p_l \right) \tag{4.6a}$$
$$p_k = \sum_l \left(\sin\left(\frac{2\pi kl}{N}\right) x_l + \cos\left(\frac{2\pi kl}{N}\right) p_l \right), \tag{4.6b}$$

which leads to

$$H_{XY} = -\frac{i}{2} \sum_k \varepsilon_k \big(\cos\theta_k (x_k p_k - p_k x_k) - \sin\theta_k (x_k x_{-k} - p_k p_{-k}) \big) \tag{4.7}$$

where

$$\cos\theta_k = \frac{\cos(\frac{2\pi k}{N}) - \lambda}{\varepsilon_k} \tag{4.8}$$

and ε_k is defined below. (We have dropped some terms that are inversely proportional to N, since we will later take the limit $N \to \infty$.)

The Fourier transform is a canonical transformation (it preserves the anti-commutation relations), so it can also be written as an orthogonal transformation of the Majorana operators (see Section A.5.5 in Appendix A):

$$\boldsymbol{\kappa} = O_{\mathrm{FT}}\, \boldsymbol{r}, \tag{4.9}$$

where $\boldsymbol{r}$ is a vector of the original Majorana operators ($r_{2l-1} = x_l$ and $r_{2l} = p_l$), $\boldsymbol{\kappa}$ its Fourier transform (ordered so that $\kappa_{4k-3} = x_k$, $\kappa_{4k-2} = p_k$, $\kappa_{4k-2} = x_{-k}$ and $\kappa_{4k} = p_{-k}$), and

$$O_{\mathrm{FT}} = \begin{pmatrix} F_{0,0} & \cdots & F_{0,N} \\ \vdots & \ddots & \vdots \\ F_{N,0} & \cdots & F_{N,N} \end{pmatrix}, \quad F_{k,l} = \begin{pmatrix} c_{k,l} & -s_{k,l} \\ s_{k,l} & c_{k,l} \end{pmatrix}. \tag{4.10}$$

Finally, using the Bogoliubov transformation,

$$\gamma_k^x = u_k x_k - v_k p_{-k}, \tag{4.11a}$$
$$\gamma_k^p = u_k p_k - v_k x_{-k}, \tag{4.11b}$$

where

$$u_k = \cos\left(\frac{\theta_k}{2}\right), \tag{4.12a}$$
$$v_k = \sin\left(\frac{\theta k}{2}\right), \tag{4.12b}$$

the Hamiltonian takes the diagonal form:

$$H_{XY} = -\frac{i}{4}\sum_k \varepsilon_k \left(\gamma_k^x \gamma_k^p - \gamma_k^p \gamma_k^x\right), \tag{4.13}$$

with spectrum

$$\varepsilon_k = \sqrt{\left(\cos\left(\frac{2\pi k}{N}\right) - \lambda\right)^2 + \gamma^2 \sin^2\left(\frac{2\pi k}{N}\right)}. \tag{4.14}$$

Like the Fourier transform, the Bogoliubov transformation preserves the canonical anti-commutation relations, and can also be written as an orthogonal transformation among Majorana operators (see Section A.5.5 in Appendix A):

$$\boldsymbol{\gamma} = O_{\mathrm{Bog}}\, \boldsymbol{\kappa} \tag{4.15}$$

where $\boldsymbol{\kappa}$ was defined in Eq. (4.9), $\boldsymbol{\gamma}$ is the corresponding vector of Bogoliubov-transformed operators ($\gamma_{4k-3} = \gamma_k^x$, $\gamma_{4k-2} = \gamma_k^p$, $\gamma_{4k-2} = \gamma_{-k}^x$ and $\gamma_{4k} = \gamma_{-k}^p$), and

$$O_{\mathrm{Bog}} = \bigoplus_{k=1}^{N/2} \begin{pmatrix} u_k & 0 & 0 & -v_k \\ 0 & u_k & -v_k & 0 \\ 0 & v_k & u_k & 0 \\ v_k & 0 & 0 & u_k \end{pmatrix} \tag{4.16a}$$

with

$$c_{k,l} = \cos\left(\frac{2\pi kl}{N}\right), \tag{4.16b}$$

$$s_{k,l} = \sin\left(\frac{2\pi kl}{N}\right). \tag{4.16c}$$

The Hamiltonian in Eq. (4.13) is quadratic in terms of the Majorana operators, and is reminiscent of a squeezing operation in quantum optics, except that the mode operators γ_k^x and γ_k^p are fermionic rather than bosonic:

$$H = \sum_{m,n} h_{m,n}\gamma_m\gamma_n, \tag{4.17a}$$

$$h = -\frac{i}{4}\bigoplus_{k=1}^{N}\begin{pmatrix} 0 & \varepsilon_k \\ -\varepsilon_k & 0 \end{pmatrix}. \tag{4.17b}$$

Therefore, the time-evolution of the Majorana operators in the Heisenberg picture also corresponds to a canonical transformation, represented by the orthogonal transformation

$$\boldsymbol{\gamma}(t) = O(t)\boldsymbol{\gamma}, \tag{4.18a}$$

where

$$O(t) = e^{At}, \quad A_{m,n} = -i(h_{m,n} - h_{n,m}) \tag{4.18b}$$

$$= \bigoplus_{k=1}^{N}\begin{pmatrix} \cos(\varepsilon_k t) & -\sin(\varepsilon_k t) \\ \sin(\varepsilon_k t) & \cos(\varepsilon_k t) \end{pmatrix}. \tag{4.18c}$$

The initial state, with all spins down, is the vacuum state of the x_l, p_l Majorana operators. Since the vacuum of any set of fermionic operators is a fermionic Gaussian state, it can be represented by its *covariance matrix*, or matrix of two-point correlations (see Section A.5 in Appendix A):

$$\Gamma_{m,n} = \langle [r_m, r_n] \rangle, \tag{4.19a}$$

$$\Gamma_{\text{vac}} = \bigoplus^{N}\begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}. \tag{4.19b}$$

An equivalent way of stating this is to appeal to Wick's theorem (Theorem A.28 of Appendix A), which states that all higher-order correlation functions for the vacuum state of any set of fermionic operators can be re-expressed in terms of two-point correlations.

Thus the initial state is Gaussian in terms of the x_l, p_l operators, the Hamiltonian is quadratic in terms of the γ_k^x, γ_k^p operators, and these sets of operators are connected by canonical transformations (namely the Fourier and Bogoliubov transformations). Transformations that preserve the anti-commutation relations leave Gaussian states Gaussian. Therefore, the state of the system remains Gaussian at all times, and the time-evolution can be analysed using the fermionic Gaussian state formalism, described in detail in Appendix A.

Putting all this together, the state at time t, as represented by its covariance matrix $\Gamma(t)_{m,n} = \langle [r_n(t), r_m(t)] \rangle$, is given by

$$\Gamma(t) = O_{\mathrm{FT}}^T\, O_{\mathrm{Bog}}^T\, O(t)\, O_{\mathrm{Bog}}\, \Gamma_{\mathrm{vac}}\, O_{\mathrm{Bog}}^T\, O(t)^T\, O_{\mathrm{Bog}}\, O_{\mathrm{FT}} \tag{4.20a}$$

$$= \begin{pmatrix} G_0 & G_1 & \cdots & G_N \\ G_{-1} & G_0 & & \vdots \\ \vdots & & \ddots & \vdots \\ G_{-N} & \cdots\cdots & & G_0 \end{pmatrix} \tag{4.20b}$$

where, in the thermodynamic limit $N \to \infty$ with $\frac{2\pi k}{N} \to \phi$, we have

$$G_x = \int_{-\pi}^{\pi} \mathrm{d}\phi \begin{pmatrix} g_0 & g_1 \\ g_{-1} & g_0 \end{pmatrix} \tag{4.20c}$$

$$g_0 = iS \sin(\phi x) \sin\big(2t\varepsilon(\phi)\big) \tag{4.20d}$$

$$g_{\pm 1} = 2CS \sin(\phi x) \sin^2\big(t\varepsilon(\phi)\big) \pm \cos(\phi x)\Big(C^2 + S^2 \cos\big(2t\varepsilon(\phi)\big)\Big) \tag{4.20e}$$

$$C = \frac{\cos(\phi) - \lambda}{\varepsilon(\phi)}, \quad S = \frac{\gamma \sin(\phi)}{\varepsilon(\phi)}, \quad x = m - n. \tag{4.20f}$$

Note that the covariance matrix in Eq. (4.20b) is of Toeplitz form, as required by translational invariance.

4.3 Correlation and Entanglement Dynamics

Eq. (4.20) gives an analytic expression for the system state at any time t in terms of the Majorana operators x_l, p_l. However, we are interested in calculating correlation functions and entanglement for the original spin system, and not all of these can be calculated analytically from the covariance matrix.

We first calculate the simpler, albeit less well-motivated, *string* correlation functions, and show that their evolution can be described by wave-packet propagation. We then consider the more interesting spin-spin (or two-point)

correlations, and show that their evolution can be understood as multiple wave packets propagating and interfering simultaneously. Finally, we derive two bounds on the localizable entanglement between two spins, one based on correlation functions, the other on a Grassmann integral over fermionic phase-space.

4.3.1 String Correlation Functions

String correlations are important in certain spin models for revealing hidden order, not detected by two-point correlation functions (for example in the AKLT model [Affleck et al., 1988]). They take the form of strings of spin operators acting on adjacent spins, for instance

$$S_{ab}(x) = \left\langle \sigma_i^a \sigma_{i+1}^z \cdots \sigma_{i+n-1}^z \sigma_{i+x}^b \right\rangle, \tag{4.21}$$

where σ^a is one of the Pauli operators, as is σ^b. Due to translational invariance, the right hand side is independent of i.

For $a, b = x$ or y, the string correlations are given directly by elements of the covariance matrix. For example,

$$\begin{aligned} S_{xx}(x) &= S_{xx}(m-n) && \text{(4.22a)} \\ &= \left\langle \sigma_m^x \Big(\prod_{m<i<n} \sigma_i^z \Big) \sigma_n^x \right\rangle && \text{(4.22b)} \\ &= \frac{1}{i} \left\langle \prod_{i<m} \sigma_i^z \sigma_m^y \prod_{j<n} \sigma_j^z \sigma_n^x \right\rangle && \text{(4.22c)} \\ &= \frac{1}{i} \left\langle x_m x_n \right\rangle && \text{(4.22d)} \\ &= \frac{1}{i} \Gamma_{2m-1,2n-1}, && \text{(4.22e)} \end{aligned}$$

using the definition of the Majorana operators from Eq. (4.3) in the third equality, and the definition of the covariance matrix (Eq. (4.19a)) in the final one.

From Eq. (4.20) for the covariance matrix at time t, the evolution of this particular string correlation function is given by

$$S_{xx}(x,t) = \int_{-\pi}^{\pi} \mathrm{d}k \frac{S}{2} \cos\big(kx - 2t\epsilon_k\big) - \int_{-\pi}^{\pi} \mathrm{d}k \frac{S}{2} \cos\big(kx + 2t\epsilon_k\big), \tag{4.23}$$

where the system spectrum ε_k was defined in Eq. (4.14), S in Eq. (4.20f).

This has the form of an equation describing two wave packets with envelope $S/2$ propagating in opposite directions along the chain, according to a dispersion relation given by the spectrum ε_k. Thus the evolution of string correlations, although produced by the collective dynamics of a large number of excitations, can be explained by very simple physics.

4.3.2 ZZ–Correlations and Localizable Entanglement

We have seen that the evolution of string correlations is described by propagating wave packets. Does the same hold true for other correlation functions? In particular, does it hold for the more interesting two-point (connected) correlation functions?

Two-point connected correlation functions are given by the expectation values of products of two spin operators, with the classical part of the correlation subtracted. For instance

$$C_{zz}(x) = C_{zz}(m-n) = \langle \sigma_n^z \sigma_m^z \rangle - \langle \sigma_n^z \rangle \langle \sigma_m^z \rangle , \tag{4.24}$$

which, again, only depends on $x = m - n$ due to translational invariance.

From the definition of the Majorana operators (Eq. (4.3)), we have $\sigma_n^z = x_n p_n / i$, so that

$$\begin{aligned} C_{zz}(x) &= -\langle x_n p_n x_m p_m \rangle + \langle x_n p_n \rangle \langle x_m p_m \rangle && \text{(4.25a)} \\ &= \langle x_n x_m \rangle \langle p_n p_m \rangle - \langle x_n p_m \rangle \langle p_n x_m \rangle , && \text{(4.25b)} \end{aligned}$$

using Wick's theorem (Theorem A.28 of Appendix A) to expand the expectation value of the product of four Majorana operators as a sum of expectation values of pairs. These correspond to entries in the covariance matrix of Eq. (4.20) which leads, after some algebra, to the following expression for the time-evolution of this connected correlation function:

$$\begin{aligned} C_{zz}(x,t)^2 = &\left(\int_{-\pi}^{\pi} \mathrm{d}k \, \frac{S}{2} \sum_{s=0}^{1} \big((-1)^s \cos(kx + (-1)^{1-s} \, 2\varepsilon_k t) \big) \right)^2 \\ &+ \left(\int_{-\pi}^{\pi} \mathrm{d}k \, CS \big(\sin(kx) - \frac{1}{2} \sum_{s=0}^{1} \sin(kx + (-1)^s \, 2\varepsilon_k t) \big) \right)^2 \\ &- \left(\int_{-\pi}^{\pi} \mathrm{d}k \, \big(C^2 \cos(kx) + \frac{S^2}{2} \sum_{s=0}^{1} \cos(kx + (-1)^s \, 2\varepsilon_k t) \big) \right)^2 . \end{aligned} \tag{4.26}$$

The spectrum ε_k was defined in Eq. (4.14), S and C in Eq. (4.20f).

Although more complicated than the string correlations, this also takes the form of wave packets evolving according to the same dispersion relation, albeit multiple wave packets with different envelopes (three in each direction) propagating and interfering simultaneously. (The time-independent terms conspire to ensure that the correlations are zero initially, as expected.)

What about entanglement? The following theorem establishes a connection between two-point correlation functions and the localizable concurrence of Definition 4.1.

Theorem 4.2 (Localizable Concurrence Correlation Bound) *For an N-qubit pure state $|\psi\rangle$, the maximum two-point connected correlation function between qubits n and m is a lower bound on the localizable concurrence of $L_{n,m}(\psi)$ of those qubits:*

$$L_{n,m}(\psi) \geq \max_{\boldsymbol{a},\boldsymbol{b}} \langle\psi|\, \boldsymbol{a}\cdot\boldsymbol{\sigma}\, \boldsymbol{b}\cdot\boldsymbol{\sigma}\, |\psi\rangle \tag{4.27}$$

where $\boldsymbol{\sigma} = (\sigma^x, \sigma^y, \sigma^z)$ is the vector of Pauli matrices, $\boldsymbol{a}$ and $\boldsymbol{b}$ are real vectors with unit length, and the dot indicates the usual scalar product.

Proof See Popp et al. [2005]. □

It follows trivially from Theorem 4.2 that any two-point connected correlation function provides a lower bound on the localizable concurrence. In particular, the localizable concurrence $L(x,t)$ between two spins a distance x apart is bounded from below by the $C_{zz}(x,t)$ correlation function of Eq. (4.26), for all times t. Therefore, when distributing entanglement using a spin chain described by the XY–model, the wave-packet description of Eq. (4.26) guarantees that at least that amount of entanglement is present.

4.4 Grassmann Integral Bound

The definition of localizable concurrence (Definition 4.1) involves a maximization over local operations. Clearly, if we restrict these local operations to projective measurements in the σ^z basis, we will obtain a lower bound:

$$L_{n,m} \geq \sum_{\substack{s_1\ldots s_N \\ =0,1}} \left| \langle\psi| \Big(\prod_{1\leq i<n} \mathcal{P}_i^{s_i}\Big) \sigma_n^y \Big(\prod_{n<j<m} \mathcal{P}_j^{s_j}\Big) \sigma_m^y \Big(\prod_{m<k\leq N} \mathcal{P}_k^{s_k}\Big) |\psi^*\rangle \right| \tag{4.28}$$

where $\mathcal{P}_i^0 = |\uparrow\rangle\langle\uparrow|$ and $\mathcal{P}_i^1 = |\downarrow\rangle\langle\downarrow|$ are projectors on the spin-up/down state of the i^{th} spin.

Inserting factors of -1 in front of appropriate terms in the sum, then taking the modulus outside the sum, we obtain

$$L_{n,m} \geq \sum_{\substack{s_1 \ldots s_N \\ =0,1}} \left| \langle\psi| \Big(\prod_{1 \leq i < n} \mathcal{P}_i^{s_i} \Big) \sigma_n^y \Big((-1)^{\sum_j s_j} \prod_{n<j<m} \mathcal{P}_j^{s_j} \Big) \sigma_m^y \Big(\prod_{m<k \leq N} \mathcal{P}_k^{s_k} \Big) |\psi^*\rangle \right| \tag{4.29a}$$

$$\geq \left| \langle\psi| \, \sigma_n^y \Big(\prod_{n<i<j} \sigma_i^z \Big) \sigma_m^y \, |\psi^*\rangle \right| \tag{4.29b}$$

$$= |\langle\psi| \, x_n p_m \, |\psi^*\rangle| \, . \tag{4.29c}$$

This gives an alternative bound, also expressed in terms of Majorana operators, to the one derived in Section 4.3.2 from correlation functions. However, the expression in Eq. (4.29c) is *not* a covariance matrix element because of the complex conjugation in the ket.

In order to derive an explicit analytic expression for this bound, it is convenient to use a phase-space representation of the fermionic Gaussian states $|\psi\rangle$ and $|\psi^*\rangle$. This is described in detail in Appendix A, though for convenience we also restate the results as we need them in this section.

Using the fermionic P–representation (Definition A.21), we can write any density operator ρ for the state of N fermions in the over-complete basis formed by the fermionic coherent states:

$$\rho = \int \mathrm{d}\boldsymbol{\xi} P(\boldsymbol{\xi}) \, |\boldsymbol{\xi}\rangle\langle\boldsymbol{\xi}| \, , \tag{4.30}$$

where $|\boldsymbol{\xi}\rangle$ denotes the fermionic coherent state (Definition A.12) represented by the N–dimensional vector $\boldsymbol{\xi}$ of Grassmann numbers (see Theorem A.22). Equation (4.30) can in fact be taken as an implicit definition of the P–representation for fermionic states.

Rewriting Eq. (4.29c) inside a trace, and making use of the P–representation, we have

$$L_{n,m}^2 \geq |\langle\psi| \, x_n p_m \, |\psi^*\rangle| = \mathrm{Tr}\Big[|\psi\rangle\langle\psi| \, x_n p_m \, |\psi^*\rangle\langle\psi^*| \, x_n p_m \Big] \tag{4.31a}$$

$$= \mathrm{Tr}\left[\left(\int \mathrm{d}\boldsymbol{\xi} P_\psi(\boldsymbol{\xi}) \, |\boldsymbol{\xi}\rangle\langle\boldsymbol{\xi}| \right) x_n p_m \left(\int \mathrm{d}\boldsymbol{\eta} P_{\psi^*}(\boldsymbol{\eta}) \, |\boldsymbol{\eta}\rangle\langle\boldsymbol{\eta}| \right) x_n p_m \right] \tag{4.31b}$$

$$= \iint \mathrm{d}\boldsymbol{\xi} \mathrm{d}\boldsymbol{\eta} \Big(-P_\psi(\boldsymbol{\xi}) P_{\psi^*}(\boldsymbol{\eta}) \, \langle\boldsymbol{\xi}| (c_n^\dagger + c_n)(c_m^\dagger - c_m) \, |\boldsymbol{\eta}\rangle \, \langle -\boldsymbol{\eta}| (c_n^\dagger + c_n)(c_m^\dagger - c_m) \, |\boldsymbol{\xi}\rangle \Big) \tag{4.31c}$$

$$= \iint \mathrm{d}\boldsymbol{\xi}\mathrm{d}\boldsymbol{\eta}\Big(P_{\psi}(\boldsymbol{\xi}) P_{\psi^*}(\boldsymbol{\eta}) \langle \boldsymbol{\xi}|\boldsymbol{\eta}\rangle \langle -\boldsymbol{\eta}|\boldsymbol{\xi}\rangle \times (\bar{\xi}_n\bar{\xi}_m - \eta_n\eta_m - \bar{\xi}_n\eta_m - \bar{\xi}_m\eta_n + \delta_{n,m}) \times (\xi_n\xi_m - \bar{\eta}_n\bar{\eta}_m + \xi_n\bar{\eta}_m + \xi_m\bar{\eta}_n - \delta_{n,m}) \Big), \tag{4.31d}$$

where $P_{\psi}(\xi)$ and $P_{\psi^*}(\eta)$ are the P–representations of states $|\psi\rangle$ and $|\psi^*\rangle$. The $\langle -\boldsymbol{\eta}|$ in Eq. (4.31c) comes from anti-commuting one of the fermionic coherent states through the rest of the expression to remove the trace (see Eq. (A.74) in Appendix A). Equation (4.31d) is obtained by normal-ordering the creation and annihilation operators c_i and $c_i^\dagger$, and using the fact that coherent states are right eigenstates of all annihilation operators and left eigenstates of all creation operators.

Define $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$ to be the "real" representations of the "complex" Grassmann vectors $\boldsymbol{\xi}$ and $\boldsymbol{\eta}$ respectively, that is

$$\boldsymbol{\alpha} = \Omega \begin{pmatrix} \xi_1 \\ \bar{\xi}_1 \\ \vdots \\ \xi_N \\ \bar{\xi}_N \end{pmatrix}, \quad \boldsymbol{\beta} = \Omega \begin{pmatrix} \eta_1 \\ \bar{\eta}_1 \\ \vdots \\ \eta_N \\ \bar{\eta}_N \end{pmatrix}, \tag{4.32}$$

where the matrix

$$\Omega = \bigoplus^{N} \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ i & -i \end{pmatrix} \tag{4.33}$$

converts from the "complex" to the "real" Grassmann representation. From Result A.13 of Appendix A, which gives an expression for the overlap of two fermionic coherent states, we have

$$\langle \boldsymbol{\xi}|\boldsymbol{\eta}\rangle \langle -\boldsymbol{\eta}|\boldsymbol{\xi}\rangle = \langle \boldsymbol{\alpha}|\boldsymbol{\beta}\rangle \langle -\boldsymbol{\beta}|\boldsymbol{\alpha}\rangle = \exp\left(\boldsymbol{\alpha}^T\boldsymbol{\beta} - \boldsymbol{\alpha}^T\Gamma_{\mathrm{vac}}\boldsymbol{\alpha} - \boldsymbol{\beta}^T\Gamma_{\mathrm{vac}}\boldsymbol{\beta}\right). \tag{4.34}$$

Also, from Theorem A.27 of Appendix A, the "complex" P–representation of a fermionic Gaussian state with covariance matrix Γ (where Γ is still expressed in the "real" representation of Eq. (4.19a) that we have been using throughout) is given by the Gaussian form

$$P(\boldsymbol{\xi}) = \mathrm{Pf}(\Gamma - \Gamma_{\mathrm{vac}}) \exp\left(-\frac{1}{2}\boldsymbol{\xi}^\dagger\, \Omega^{-1}(\Gamma - \Gamma_{\mathrm{vac}})(\Omega^{-1})^\dagger \boldsymbol{\xi}\right). \tag{4.35}$$

Let Γ_{ψ} denote the (real representation of the) covariance matrix corresponding to the system state $|\psi\rangle$, which at time t is given by Eq. (4.20). Let Γ_{ψ^*} denote the (real representation of the) covariance matrix corresponding

to $|\psi^*\rangle$. Now, our initial state (the state with all spins down) obviously has only real coefficients in the spin basis, and the original spin Hamiltonian of Eq. (4.2) is also invariant under complex conjugation in that basis. The only effect of the complex conjugation is to reverse the direction of time in the evolution operator $\exp(-iHt)$. Thus Γ_{ψ^*} is given by Eq. (4.20) with t replaced by $-t$.

Using Eq. (4.34) and Eq. (4.35) in Eq. (4.31d), the bound on the localizable entanglement becomes

$$L_{n,m}^2 \geq \iint \mathrm{d}\boldsymbol{\xi}\mathrm{d}\boldsymbol{\eta}\, p_{n,m}(\boldsymbol{\xi},\boldsymbol{\eta})\times \exp\left(\left(\boldsymbol{\xi}^\dagger\ ,\ \boldsymbol{\eta}^\dagger\right)(\Omega^{-1}\oplus\Omega^{-1})M(\Omega^{-1}\oplus\Omega^{-1})^\dagger\begin{pmatrix}\boldsymbol{\xi}\\ \boldsymbol{\eta}\end{pmatrix}\right), \quad (4.36a)$$

with

$$M = \begin{pmatrix}(\Gamma_\psi - \Gamma_{\mathrm{vac}})^{-1} + 2\Gamma_{\mathrm{vac}} & -\mathbb{1} \\ \mathbb{1} & (\Gamma_{\psi^*} - \Gamma_{\mathrm{vac}})^{-1} + 2\Gamma_{\mathrm{vac}}\end{pmatrix} \quad (4.36b)$$

and

$$p_{n,m}(\boldsymbol{\xi},\boldsymbol{\eta}) = (\bar{\xi}_n\bar{\xi}_m - \eta_n\eta_m - \bar{\xi}_n\eta_m - \bar{\xi}_m\eta_n + \delta_{n,m})\times (\xi_n\xi_m - \bar{\eta}_n\bar{\eta}_m + \xi_n\bar{\eta}_m + \xi_m\bar{\eta}_n - \delta_{n,m}) \quad (4.36c)$$

The integrand in Eq. (4.36a) is the product of a multi-dimensional Gaussian with a fourth-order polynomial $p_{n,m}$ in the Grassmann variables. As such, it can be carried out using a general result for Grassmann integrals involving the product of an even monomial with a Gaussian.

Result 4.3 (Even Grassmann Monomial times Gaussian Integral) *The Grassmann integral of the product of an even order monomial with a Gaussian is given by*

$$\int \mathrm{d}\boldsymbol{\alpha}\exp\left(-\frac{1}{2}\boldsymbol{\alpha}^T A\,\boldsymbol{\alpha}\right)\alpha_{i_1}\alpha_{i_2}\ldots\alpha_{i_{2n-1}}\alpha_{i_{2n}} = \mathrm{Pf}(A)\,\frac{1}{4}\sum_{\mathcal{P}}\mathrm{sgn}(\pi)A^{-1}_{\pi_1,\pi_2}\ldots A^{-1}_{\pi_{2n-1},\pi_{2n}}. \quad (4.37)$$

The sum is over all permutations π of the indices $i_1,\ldots,i_{2n}$, where $\mathrm{sgn}(\pi)$ *denotes the signum of the permutation.* $\mathrm{Pf}(A)$ *denotes the Pfaffian of the anti-symmetric matrix A, equal to the square-root of its determinant.*

Proof See Result A.6 in Section A.2.3. □

As fermionic covariance matrices are always anti-symmetric, the matrix M defined in Eq. (4.36b) is also anti-symmetric. With the elements of M given by Eq. (4.36b) and those of the covariance matrices by Eq. (4.20) from Section 4.2, applying Result 4.3 to the integral in Eq. (4.36a) for each term in the polynomial $p_{n,m}(\boldsymbol{\xi}, \boldsymbol{\eta})$, gives (after no inconsiderable algebra) a final expression for the bound on the localizable entanglement:

$$\begin{aligned} LE(x,t) \geq & \\ \prod_k & \left(\cos(2\varepsilon_k t) - iC\sin(2\varepsilon_k t)\right) \times \\ & \frac{1}{N}\sum_k \Bigg(\frac{2\big(C + \exp(2i\varepsilon_k t)(1-C)\big)\big(C\cos(kx) - S\sin(kx)\big) - 2\cos(kx)}{\cos(2\varepsilon_k t) - iC\sin(2\varepsilon_k t)} \\ & \qquad + 2\left(S\sin(kx) - C\cos(kx)\right)\Bigg). \end{aligned} \tag{4.38}$$

This bound depends only on the separation $x = m - n$ of the two spins, as expected from the translational invariance of the system. (The system spectrum ε_k was defined in Eq. (4.14), S and C in Eq. (4.20f), both in Section 4.2.)

Unlike the correlation-based bound of Section 4.3.2, this has no simple physical interpretation. What is worse, it tends to zero in the thermodynamic limit $N \to \infty$ for $t \neq 0$, since the product becomes an infinite product of terms with magnitude less than one. However, for short times and finite chains, it is typically tighter than the correlation bound given by Eq. (4.26) (see Fig. 4.1).

It should be pointed out though that, as we neglected terms of order $1/N$ during the diagonalization of the XY–Hamiltonian (Eq. (4.7)), the expression of Eq. (4.20) for the covariance matrix, and hence also the bound we have just derived, is only strictly valid in the thermodynamic limit. However, the numerical results shown in Fig. 4.2 indicate that Eq. (4.38) is a good approximation to the exact bound given by Eq. (4.29c), even for relatively small N.

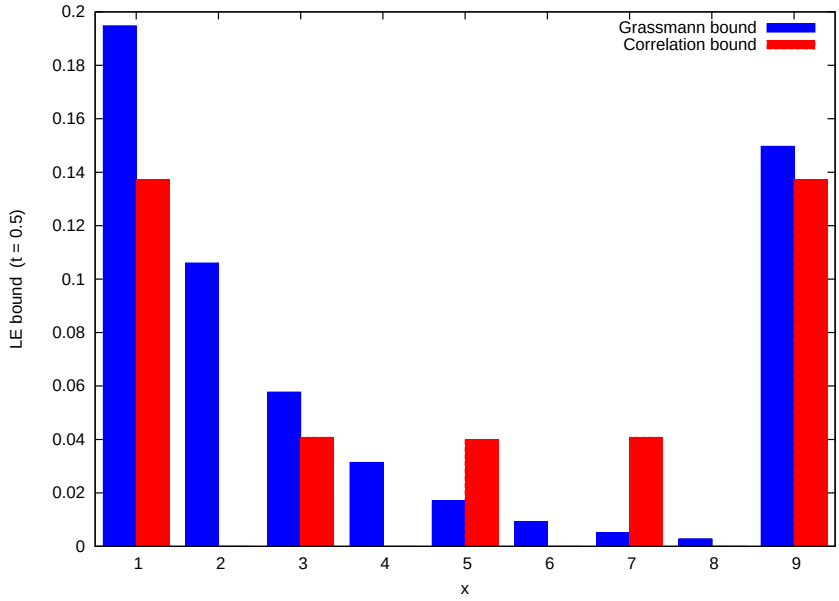

Figure 4.1: Bound on the localizable entanglement in the $\gamma = 1$, $\lambda = 0.2$ XY–model at $t = 0.5$ (blue), from Eq. (4.38). The corresponding correlation-function bound from Eq. (4.26) (red) is weaker for certain values of x, especially low values and at small times t (as shown here).

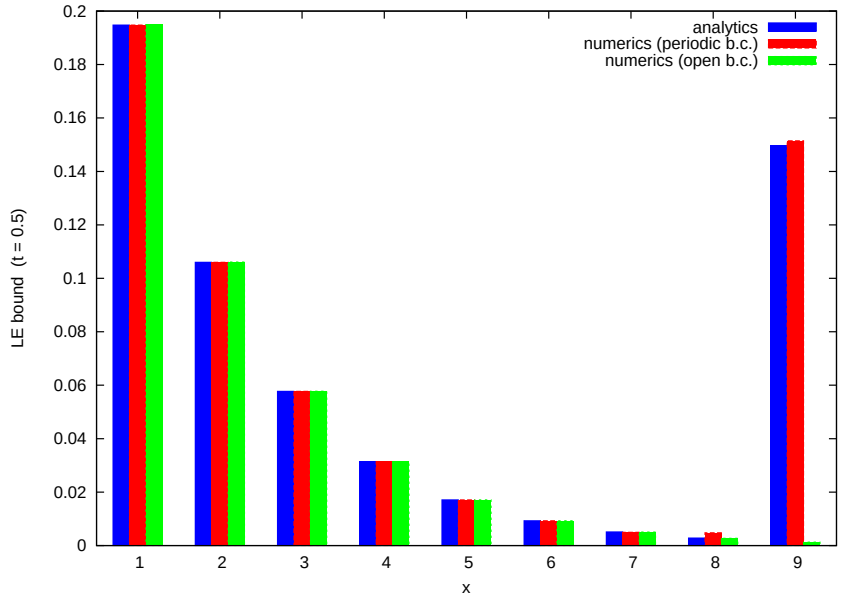

Figure 4.2: Bound on the localizable entanglement in the $\gamma = 1$, $\lambda = 0.2$ XY–model at $t = 0.5$ (blue), from Eq. (4.38). Numerical calculations of $\langle\psi|\,\sigma_1^y \prod_{1<j<n} \sigma_j^z \sigma_{1+x}^y\,|\psi^*\rangle$ (on which the analytic bound is based, and which is a lower bound on the localizable entanglement even for finite chains) are shown in red (periodic boundary conditions) and green (open boundary conditions), for a finite chain of length 10. Despite the fact that the analytic bound is only strictly valid in the thermodynamic limit, it remains a very good approximation up to times for which the boundary conditions become significant (as is already happening at $x = 9$ in the open boundary condition case).

4.5 Engineering the Correlation Dynamics

The aim of this chapter is not only to demonstrate that the dynamics of correlations and entanglement in a many-body system can show interesting behaviour, but moreover to demonstrate that the system can be deliberately *engineered* in order to produce useful behaviour. To this end, we will use the explicit analytic expressions for the correlation functions derived in Sections 4.3.1 and 4.3.2, and the corresponding correlation-based bound on the localizable entanglement of Section 4.3.2, to study to what extent the correlation dynamics can be controlled by changing the external parameters of our example system.

4.5.1 Correlation Wave Packets

Although the behaviour of the correlations described by Eq. (4.23) and Eq. (4.26) can be quite complex, the fact that both equations can be interpreted in terms of propagating wave packets allows us to make quantitative predictions as to the behaviour of the correlations.

String Correlations

Taking first the simpler, though less well-motivated, case of string correlations, Eq. (4.23) for the XX string correlations as a function of block size x and time t, which we reproduce here for convenience:

$$S_{xx}(x,t) = \int_{-\pi}^{\pi} \mathrm{d}k \frac{S}{2} \cos\big(kx - 2\epsilon_k t\big) - \int_{-\pi}^{\pi} \mathrm{d}k \frac{S}{2} \cos\big(kx + 2\epsilon_k t\big), \qquad (4.39)$$

has precisely the form of a solution to a one-dimensional wave equation, with dispersion relation

$$\omega(k) = 2\varepsilon_k = 2\sqrt{(\cos k - \lambda)^2 + \gamma^2 \sin^2 k}. \qquad (4.40)$$

(The spectrum ε_k of the XY-model was given in Eq. (4.14). The wave-number k runs from $-\pi$ to π, as usual.) The envelope of the wave packets is given by

$$f(k) = \frac{S}{2} = \frac{\gamma \sin(k)}{2\sqrt{(\cos k - \lambda)^2 + \gamma^2 \sin^2 k}}, \qquad (4.41)$$

plots of which are shown in Fig. 4.3. Note that both the dispersion relation and the wave-packet envelope depend on the system parameters γ and λ. For

given γ and λ, the peaks of the envelope are at

$$k = \begin{cases} \arccos(\lambda) & \lambda \leq 1 \\ \arccos\left(\frac{1}{\lambda}\right) & \lambda > 1, \end{cases} \tag{4.42}$$

At that value of k, the first and second derivatives of the dispersion relation are

$$v(k) = \frac{\mathrm{d}\omega(k)}{\mathrm{d}k} = \begin{cases} 2\gamma\lambda & \lambda \leq 1 \\ \frac{2\sqrt{\gamma^2+\lambda^2-1}}{\lambda} & \lambda > 1, \end{cases} \tag{4.43a}$$

$$d(k) = \frac{\mathrm{d}^2\omega(k)}{\mathrm{d}k^2} = \begin{cases} \frac{2(1-\gamma^2)\sqrt{1-\lambda^2}}{\gamma} & \lambda \leq 1 \\ \frac{2(1-\gamma^2)}{\lambda}\sqrt{\frac{\lambda^2-1}{\gamma^2+\lambda^2-1}} & \lambda > 1. \end{cases} \tag{4.43b}$$

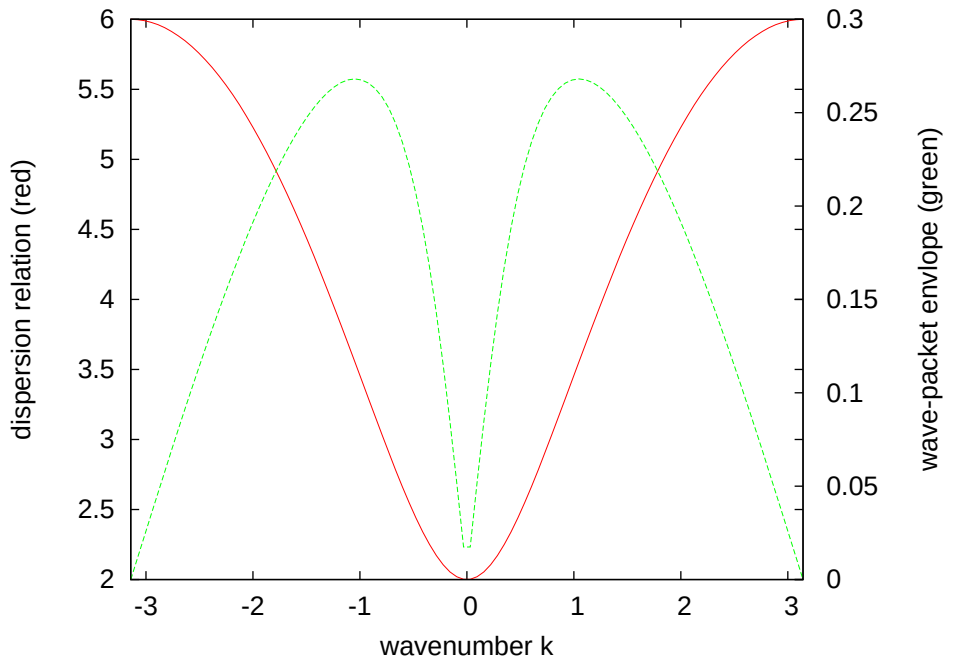

Figure 4.3: String correlation wave-packet envelope (green) from Eq. (4.41), for $\gamma = 1.1$, $\lambda = 2.0$. These correlation packets propagate according to the dispersion relation (red), given by Eq. (4.40). (Note that the green curve shows the *absolute value* of the envelope function; its sign is unimportant for our purposes.)

Therefore, we would expect the string correlation dynamics to be well approximated by a wave packet with envelope $f(k)$ from Eq. (4.41) propagating at a group velocity given by Eq. (4.43a) and dispersing at a rate given by

Eq. (4.43b). (There is also always a mirror-image wave packet propagating in the opposite direction.) For example, for $\gamma = 1.1$ and $\lambda = 2.0$, the wave-packet envelope is shown in Fig. 4.3, and Eq. (4.43a) and Eq. (4.43b) predict a group velocity $v \approx 2$ and dispersion rate $d \approx -0.18$. This closely matches the actual behaviour, shown in Fig. 4.4.

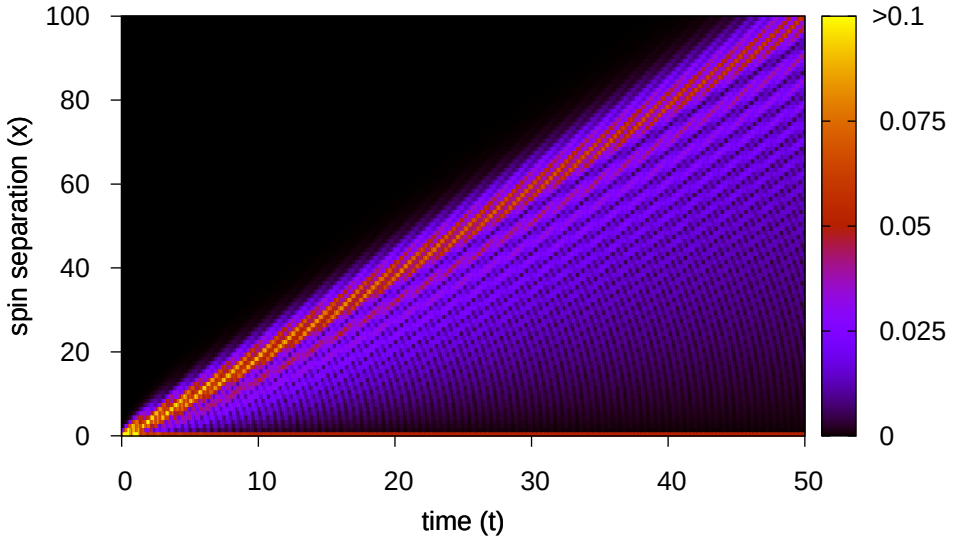

Figure 4.4: The wave packets defined by the envelope shown in Fig. 4.3 propagate according to the dispersion relation of Eq. (4.40). For those values of the system parameters ($\gamma = 1.1$, $\lambda = 2.0$), Eq. (4.43a) and Eq. (4.43b) predict a group velocity $v \approx 2$ and dispersion rate $d \approx -0.18$. This is in good agreement with the actual behaviour, which indeed shows little dispersion and a well-defined correlation velocity, with the correlations propagating a distance $x = 100$ in time $t = 50$.

Note that, when we talk about wave packets "propagating", we really mean that the string correlations in a block of spins of size x spread out over time according to the wave-packet description given above. Moreover, as the system is translationally invariant, this dynamics is repeated for every translation of that block along the chain: the string correlations of spins n to m are identical to those of spins $n+l$ to $m+l$ at all times, for any translation l.

Two-Point Connected Correlations

The study of two-point (or spin-spin) correlations is better motivated than that of string correlations (see Section 4.1), but it is also more complicated.

Equation (4.26) for the ZZ-correlation function,

$$C_{zz}(x,t)^2 = \left(\int_{-\pi}^{\pi} \mathrm{d}k\, \frac{S}{2} \sum_{s=0}^{1} \left((-1)^s \cos(kx + (-1)^{1-s}\, 2\varepsilon_k t)\right)\right)^2 \tag{4.44}$$
$$+ \left(\int_{-\pi}^{\pi} \mathrm{d}k\, CS\left(\sin(kx) - \frac{1}{2}\sum_{s=0}^{1} \sin(kx + (-1)^s\, 2\varepsilon_k t)\right)\right)^2$$
$$- \left(\int_{-\pi}^{\pi} \mathrm{d}k\, \left(C^2 \cos(kx) + \frac{S^2}{2}\sum_{s=0}^{1} \cos(kx + (-1)^s\, 2\varepsilon_k t)\right)\right)^2,$$

also resembles the solution to a one-dimensional wave-equation, but this time there are three wave packets (plus their mirror-images) propagating simultaneously. The dispersion relation is the same as for the string correlations (Eq. (4.40)), so again depends on the system parameters γ and λ, as do all three wave-packet envelopes. This shows that the complicated dynamics produced by the propagation and interference of a large number of fundamental excitations can lead to correlation dynamics that can be described by remarkably simple physics, at least in the example system we have analysed.

Unlike the string-correlation case, by "propagation" of two-point correlations wave packets, we mean something closer to propagation along the chain. If we take one spin n to be fixed, then correlations between that spin and a spin $n + x$ do indeed propagate along the physical spin-chain. However, like the string correlations, translational invariance dictates that the same dynamics is simultaneously occurring all along the chain, repeated for all translations $n + l$ and $n + x + l$ of the two spins.

In subsequent sections, we will investigate examples of the two-point correlation dynamics in more detail. As any two-point connected correlation function gives a lower bound on the localizable entanglement, the discussion also applies to entanglement, with the caveat that it only guarantees the behaviour of some minimum amount of entanglement. In the following, we will use the term "correlations" to refer to both correlation functions and (localizable) entanglement.

4.5.2 Engineering the Correlation Velocity

What form does the two-point correlation dynamics take for different values of the system parameters? For parameter regimes in which the dispersion relation is highly non-linear, and the wave-packet envelopes are very broad (in frequency-space), the correlations will disperse rapidly, and we would not expect to see any well-defined propagation of packets of correlation.

However, if we can find parameter regimes in which the dispersion relation is reasonably linear over a large frequency range, and all three wave-packet envelopes are concentrated around this linear region, then we would expect to see propagation of correlation packets along the chain at some well-defined group velocity, with the packets dispersing slowly as they propagate. In fact, such regimes do exist, as shown in Fig. 4.5, and our predictions are borne out by the actual correlation dynamics. This provides an explicit example of the behaviour suggested by the group-velocity bound on correlation propagation of Bravyi et al. [2006] (see Section 3.5.2).

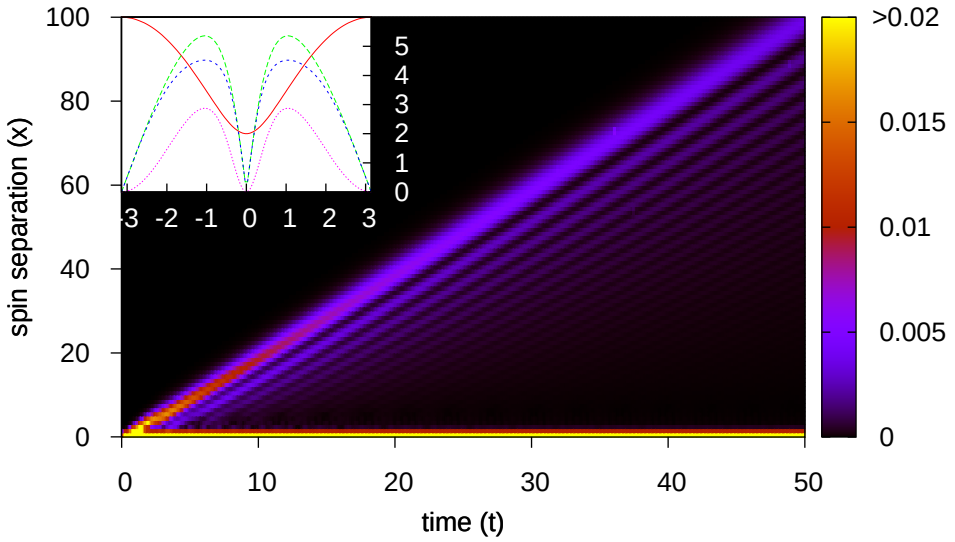

Figure 4.5: For $\gamma = 1.1$, $\lambda = 2.0$, all the wave-packet envelopes from Eq. (4.44) (dashed curves, inset) are similar in form, centred around a nearly linear region of the dispersion relation with gradient $v \approx 2$ (red curve, inset) and curvature $d \approx -0.1$. Thus the correlations propagate in well-defined packets at a speed given by the gradient, with little dispersion. (Note that the wave-packet envelopes have been scaled up by a factor of 10.)

Changing the system parameters within this regime changes the gradient of the dispersion relation. As long as the the wave-packet envelopes remain peaked in the fairly linear region of the dispersion relation, so that the correlation wave packets remain well-defined as they propagate, this allows us to engineer the correlation velocity of the system. Examples are shown in Figs. 4.6 and 4.7.

This demonstrates that it is possible to engineer useful correlation dynamics simply by controlling the *global* properties of the interactions, without requiring any detailed local control, either over the interactions or during the evolution.

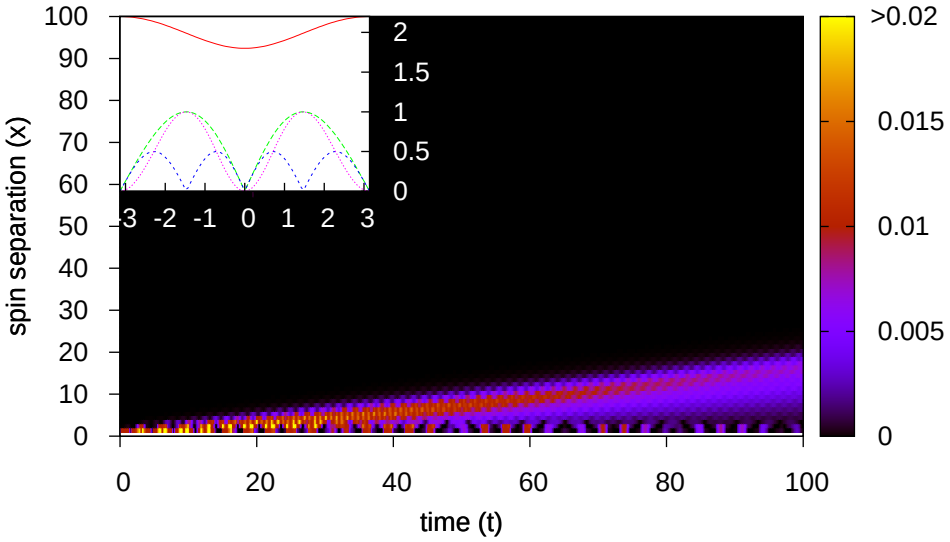

Figure 4.6: For $\gamma = 1.0$, $\lambda = 0.1$, the most prominent peaks in the wave-packet envelope (green and magenta curves, inset) are centred around a very linear region of the dispersion relation (red curve), with group velocity $v \approx 0.2$ and dispersion rate $d \approx 0$. Correlations therefore propagate slower than in Fig. 4.5 (note the horizontal scale). The less prominent peaks in the wave-packet envelope (blue curves, inset) propagate at a slightly slower speed $v \approx 1.5$, and cause the broadening of the correlation packets seen in the figure.

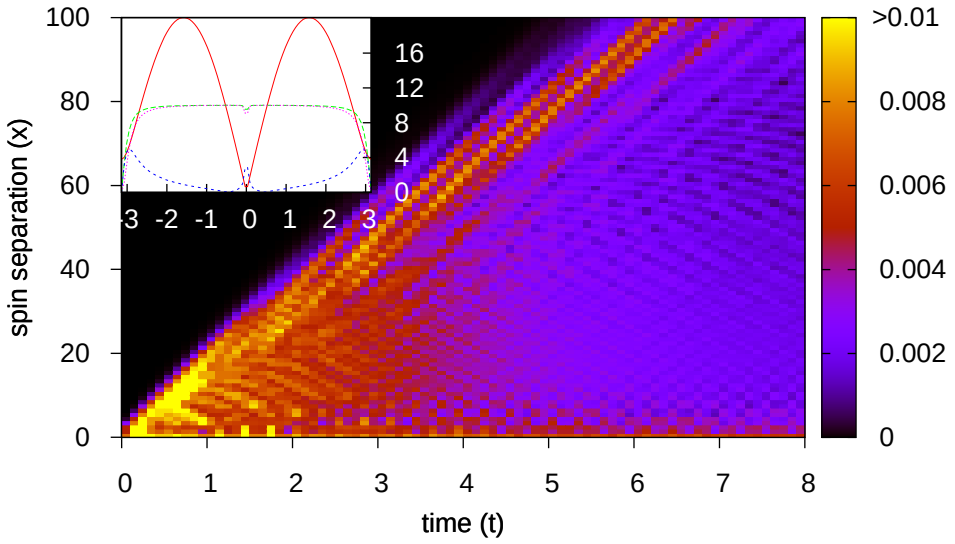

Figure 4.7: For $\gamma = 10.0$, $\lambda = 0.9$, the wave-packet envelopes are spread over the entire frequency range (dashed curves, inset). However, the dispersion relation (red curve, inset) is reasonably linear both for wave-numbers above and below $\pi/2$, with gradients $v \approx \pm 18$. As the envelopes are symmetric around $\pi/2$, the correlations still propagate at a well-defined speed $v \approx 18$, faster than in Fig. 4.6 (note the horizontal scale), though as expected from the very broad envelopes and larger dispersion rate $d \approx -9$ around the envelope peaks, they also show significantly more dispersion. (Note that the wave-packet envelopes have been scaled up by a factor of 10.)

4.5.3 Controlling the Correlation Packets

We have shown that the correlation velocity can be engineered by changing the system parameters, but an even more interesting possibility arises if we are able to change the system parameters *during* the evolution. The extent of the control over the system parameters will, of course, depend on the physical setup corresponding to the spin model, but it is not too unrealistic a requirement; the parameter λ in the XY–model, for example, can be interpreted as the strength of a uniform, external magnetic field, which can easily be controlled.

If the system parameters in our example system, the XY–model, are changed with time, the covariance matrix is no longer given by Eq. (4.20). Instead, it is given by a similar expression,

$$\Gamma(t) = O_{\mathrm{FT}}^{T}\, O(t)\, \Gamma_{\mathrm{vac}}\, O(t)^{T}\, O_{\mathrm{FT}}, \tag{4.45}$$

but with the orthogonal evolution operator $O(t)$ given by the time-ordered exponential

$$O(t) = T\left[\exp\left(\int_0^t \mathrm{d}t' A(t')\right)\right] = \lim_{h\to 0} \prod_{n=1}^{\lfloor t/h \rfloor} \exp\big(A(nh)\big). \tag{4.46}$$

(We have incorporated the Bogoliubov transformation into $O(t)$, as it depends on the system parameters which now depend on time.) The anti-symmetric matrix A corresponding to the Hamiltonian is now time-dependent, and is given by

$$A_{m,n} = -i(h_{m,n}(t) - h_{n,m}(t)), \quad H = \sum_{m,n} h_{m,n}(t) r_k r_{k'}, \tag{4.47}$$

where the decomposition of the Hamiltonian in terms of the Majorana operators r_k defined in Section 4.2 now changes with time.

In general, the time-ordering is essential, and we can not obtain an explicit analytic expression for the covariance matrix. But, if the system parameters, and hence $A(t)$, change slowly in time, dropping the time-ordering will give a good approximation to the evolution operator. In that case, from Eq. (4.46), $O(t)$ becomes the exponential of the time-averaged value of $A(t)$. The evolution is therefore determined simply by the running time-average of the Hamiltonian; that is, the state at time t is the same as that due to evolution under an XY–Hamiltonian with parameters given by the time-average:

$$\lambda = \frac{1}{t}\int_0^t \mathrm{d}t' \lambda(t'), \quad \gamma = \frac{1}{t}\int_0^t \mathrm{d}t' \lambda(t'). \tag{4.48}$$

If we remain in the parameter regime displaying well-defined wave-packet propagation (see Section 4.5.2), for which the relevant part of the dispersion relation is approximately linear, adjusting the parameters changes the gradient of the dispersion relation without significantly affecting its curvature. Thus, to a good approximation, we can slowly adjust the parameters to control the velocity of the wave packets as they propagate, allowing us to speed up and slow down the correlations. Figure 4.8, which shows results calculated by numerically approximating the time-ordered exponential of Eq. (4.46), confirms our analysis.

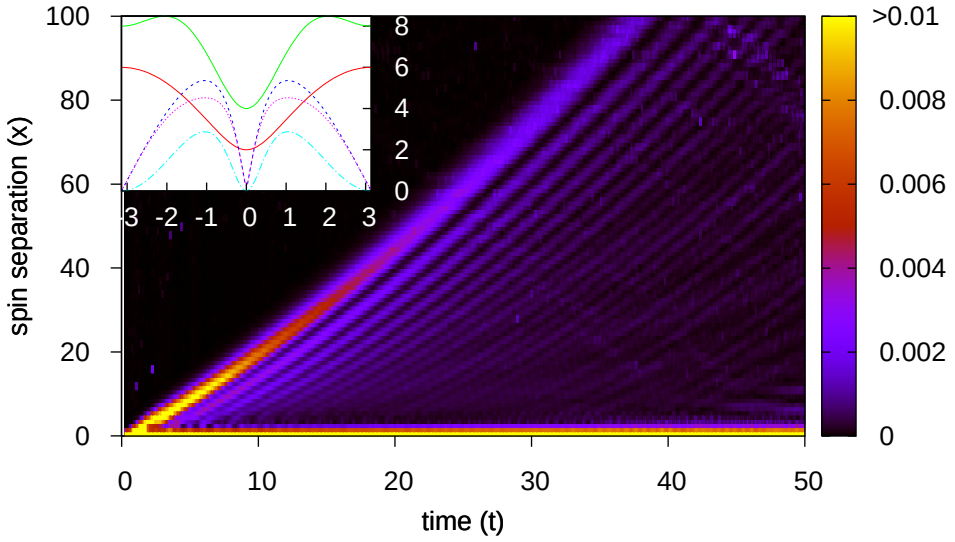

Figure 4.8: Starting from the parameters $\gamma = 1.1$, $\lambda = 2.0$ of Fig. 4.5, the parameters are smoothly changed to move from the solid to the dashed dispersion relation (inset), increasing the correlation speed. (Note that th wave-packet envelopes (non-dashed curves, inset) have been scaled up by a factor of 10.)

This implies we can achieve significant control over the propagation of correlations in a spin chain even during its evolution. And all this whilst only requiring external control over simple, global properties of the physical system.

4.5.4 Freezing Correlations

We can engineer a system so that correlations propagate through it in well-defined packets, and we can control the velocity of these packets as they are propagating. Clearly if we are interested in using the system to distribute entanglement, it would be useful to be able to freeze the correlations at any

desired location. One method would be to simply switch off the interactions. But, depending on the system, that may be difficult (if the spin model were realized in a solid-state system, for example, switching off the interactions would likely involve fabricating an entirely new system).

If we assume, as in the previous section, that the only control we have is over a few global parameters of the system (γ and λ in our XY–model example), can we still freeze correlations? Instead of changing the parameters continuously, we now consider changing them abruptly. The time-evolved covariance matrix in this scenario can be calculated analytically by the same methods used in the fixed-parameter scenario of Section 4.2.

Suppose the initial system parameters γ_0 and λ_0 are suddenly changed to γ_1 and λ_1 at time t_1. Until time t_1, the two-point ZZ correlation function will, of course, evolve according to Eq. (4.26) from Section 4.3.2, as before. After time t_1, the evolution becomes more complicated. The covariance matrix is given by

$$\Gamma(t) = O_{\mathrm{FT}}^T \, \mathcal{O} \Gamma_{\mathrm{vac}} \mathcal{O}^T \, O_{\mathrm{FT}} \tag{4.49}$$

$$\mathcal{O} = O_{\mathrm{Bog}_1}^T \, O(t - t_1) \, O_{\mathrm{Bog}_1} \, O_{\mathrm{Bog}_0}^T \, O(t_1) \, O_{\mathrm{Bog}_1}. \tag{4.50}$$

The Fourier transform and time-evolution operators O_{FT} and $O(t)$ are given by Eq. (4.10) and Eq. (4.18b), as before. The Bogoliubov transformation operators O_{Bog_0} and O_{Bog_1} are given by Eq. (4.16a) with parameters γ_0, λ_0 and γ_1, λ_1 respectively.

The ZZ correlation function is still given by Eq. (4.25b), and the analogue of Eq. (4.26) (the wave-packet description of the ZZ–correlations) separates into a sum of wave packets evolving in four different ways: those that initially evolve according to ϵ_0 and subsequently (after t_1) evolve according to ϵ_1, those that subsequently evolve according to $-\epsilon_1$, those that only start evolving at t_1, and those that that undergo no further evolution after t_1 (all of which can be seen in Fig. 4.9).

We are primarily interested in the terms whose evolution is "frozen" at time t_1, which for $t > t_1$ are given by

$$\begin{aligned} C_{zz}^{t_1}(x,t)^2 = &\left(\frac{1}{2} \int_{-\pi}^{\pi} \mathrm{d}k S_1 S_2 (C_1 S_2 - C_2 S_1) \sum_{\pm} \sin(kx \pm 2 t_1 \varepsilon_1) \right)^2 \\ &- \left(\frac{1}{2} \int_{-\pi}^{\pi} \mathrm{d}k S_1 C_2 (C_1 S_2 - C_2 S_1) \sum_{\pm} \cos(kx \pm 2 t_1 \varepsilon_1) \right)^2 . \end{aligned} \tag{4.51}$$

Since t does not appear on the right hand side, this expression clearly describes wave packets that propagate until time t_1 and then stop. Using these, we can

move correlations to the desired location, and then "quench" the system by abruptly changing the parameters, freezing the correlations at that location. An example of this is shown in Fig. 4.9.

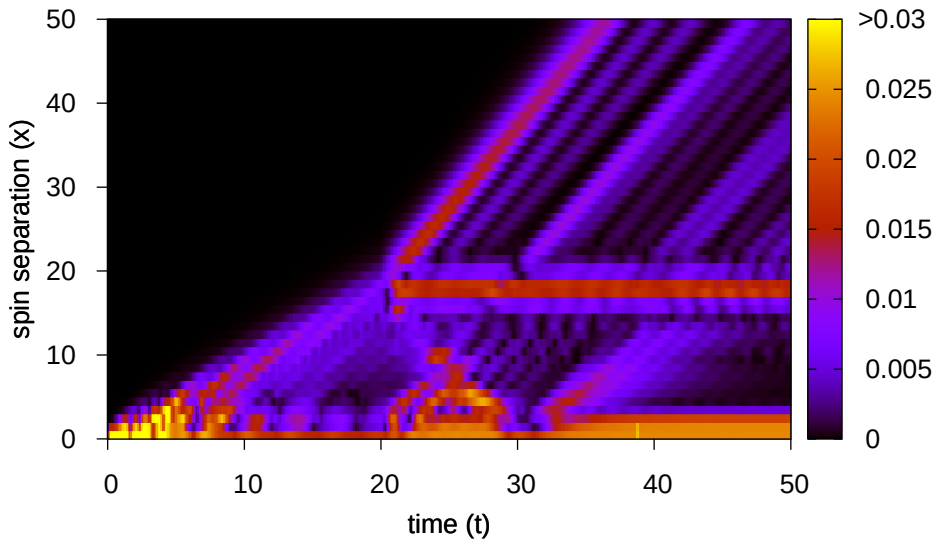

Figure 4.9: The system is initially allowed to evolve with parameters $\gamma_0 = 0.9$ $\lambda_0 = 0.5$, then "quenched" at time $t_1 = 20$ to $\gamma_1 = 0.1$, $\lambda_1 = 10$. Some of the correlations are frozen at the location they reached at t_1 (the bright line at $x \approx 18$ after $t = 20$). Others propagate according to the new dispersion relation, or are "reflected".

In conclusion, we have shown that by changing simple, global parameters, a many-body system can be engineered so that the correlation dynamics produced when the system is started in an easily prepared non-equilibrium state consists of correlation packets propagating along the system at a well-defined group-velocity. Not only can the correlation velocity be engineered in advance, but by changing the same global parameters during the evolution, the propagation speed can be controlled, and some of the correlation packets can even be frozen once they have reached a desired location. This level of control clearly has applications to entanglement distribution in the quantum-repeater-like setups discussed in Section 4.1.

Appendix A

Fermionic Gaussian States

In this appendix, we develop a theory of fermionic Gaussian states which parallels that used frequently in quantum optics for bosons, and give proofs for the results used in Chapter 4. There are numerous, equivalent ways to define fermionic Gaussian states. We will follow the approach used in the bosonic case, by first defining fermionic coherent states and displacement operators, and using these to construct a characteristic function and other phase-space representations for fermionic states. We can then define a Gaussian state to be one whose phase-space representations are Gaussian.

As will be discussed in Section A.1, this development will require the use of anti-commuting, or *Grassmann*, numbers. We give an overview of the algebra and calculus of Grassmann numbers in Section A.2, and derive a number of results that will be necessary later. In Section A.3, we show how Grassmann numbers can be used to define fermionic displacement operators, which generate coherent states when applied to the vacuum. These will allow us to construct phase-space representations of fermionic states in Section A.4. Finally, in Section A.5, we define fermionic Gaussian states and derive some of their properties.

A number of the results relating to Grassmann numbers, fermionic coherent states and phase-space representations can be found in Simons [2001, Chapters 3 and 4] and, especially, Cahill and Glauber [1999]. They are reproduced here in order to make this thesis reasonably self-contained, and in order to present alternative, simpler proofs for some of the results. These results then serve as the foundation on which the fermionic Gaussian state formalism is built.

A.1 Motivation

A coherent state of a collection of bosonic modes is defined to be any simultaneous eigenstate of all annihilation operators, and the set of all coherent states forms an over-complete basis for states of those modes. Bosonic coherent states are not only useful as a mathematical tool, but are also important physically: ignoring certain subtleties, they are the states produced by lasers.

Starting from coherent states, and the displacement operators that create them from the vacuum, it is possible to build various phase-space representations and define bosonic Gaussian states. Again, Gaussian states are relatively simple mathematically, but also physically relevant: they are precisely the states produced by linear optical elements (lasers, phase plates, interferometers, optical squeezers, etc.).

It is not unreasonable to suppose that a similar formalism might be useful in the case of fermions. Many calculations involving collections of fermionic modes have a very similar mathematical structure to the equivalent calculations for bosons. Also, experimentalists are now able to create atom lasers, atom interferometers, etc. If the atoms have half-integer nuclear spin, they will behave as fermions rather than bosons, so fermionic Gaussian states are increasingly relevant to today's experiments.

We can try to define coherent states in a similar way for a collection of fermionic modes: as simultaneous eigenstates of all fermionic annihilation operators. Then

$$c_i \left|\boldsymbol{\eta}\right\rangle = \eta_i \left|\boldsymbol{\eta}\right\rangle, \tag{A.1}$$

However, we run into trouble due to the anti-commutation relations

$$\{c_i, c_j\} = \delta_{ij} \tag{A.2}$$

obeyed by the annihilation operators. We get a different sign depending on which order we apply two annihilation operators c_i and $c_{j\neq i}$:

$$\begin{array}{ccccc} c_i c_j \left|\boldsymbol{\eta}\right\rangle & = & c_i \eta_j \left|\boldsymbol{\eta}\right\rangle & = & \eta_j \eta_i \left|\boldsymbol{\eta}\right\rangle \\ \| & & & \nparallel & \\ -c_j c_i \left|\boldsymbol{\eta}\right\rangle & = & -c_j \eta_i \left|\boldsymbol{\eta}\right\rangle & = & -\eta_i \eta_j \left|\boldsymbol{\eta}\right\rangle \end{array} \tag{A.3}$$

(the final equality in each row follows from linearity of the annihilation operators). This shows that fermionic coherent states do not exist in the fermionic Fock space constructed from c_i and $c_i^\dagger$.

There is a standard solution to this inconvenience. Notice that Eq. (A.3) would be consistent if the eigenvalues η_i and η_j were *anti-commuting* numbers,

as we would then have

$$\eta_j \eta_i \left|\boldsymbol{\eta}\right\rangle = -\eta_i \eta_j \left|\boldsymbol{\eta}\right\rangle. \tag{A.4}$$

Such anti-commuting numbers can be defined, and are called Grassmann numbers (see Section A.2). By formally expanding the fermionic Fock space algebra to include these Grassmann numbers, we can avoid the problems of Eq. (A.3), and Eq. (A.1) becomes consistent.

The expanded Fock space includes the original fermionic Fock space, so it includes all physical fermionic states. However, it also includes states that are not physical. For example, anticipating the more rigorous treatment of Grassmann algebra in Section A.2, the fact that the coherent states are unphysical is easily seen by calculating the expectation value of $c_i + c_i^\dagger$. Taking the Hermitian conjugate of Eq. (A.4), the dual $\left\langle\boldsymbol{\eta}\right|$ of a coherent state $\left|\boldsymbol{\eta}\right\rangle$ is seen to be a left eigenstate of all the *creation* operators:

$$\left(c_i \left|\boldsymbol{\eta}\right\rangle\right)^\dagger = \left\langle\boldsymbol{\eta}\right| c_i^\dagger = \bar{\eta}\, c_i^\dagger, \tag{A.5}$$

where complex conjugation of Grassmann numbers is formally defined to be a mapping between pairs of Grassmanns: $\eta^* = \bar{\eta}$. The expectation value is then given by

$$\left\langle\boldsymbol{\eta}\right| c_i + c_i^\dagger \left|\boldsymbol{\eta}\right\rangle = \eta_i + \bar{\eta}_i. \tag{A.6}$$

The operator $c_i + c_i^\dagger$ is Hermitian, so it corresponds to a physical observable. As such, its expectation value should be a real number for any state of the system. However, $\eta_i + \bar{\eta}_i$ is a Grassmann number (it anti-commutes with any other Grassmann η_j). Therefore, the coherent states can not be physical states of the fermionic modes. Nonetheless, fermionic coherent states prove to be a very useful mathematical tool.

A.2 Grassmann Algebra and Calculus

A.2.1 Grassmann Numbers and their Algebra

Grassmann numbers are entities that anti-commute with other Grassmann numbers, but commute with ordinary numbers. We will denote Grassmann numbers by lower-case Greek letters (e.g. η, ξ), and ordinary numbers by lower-case Roman letters (e.g. a, b). Their (anti-)commutation relations are then given by:

$$\{\eta, \xi\} = \eta\xi + \xi\eta = 0, \tag{A.7a}$$

$$[a, \eta] = a\eta - \eta a. \tag{A.7b}$$

In particular, this implies that, for any Grassmann number η,

$$\eta^2 = 0. \tag{A.8}$$

Grassmann multiplication is defined to be distributive over addition:

$$x(y+z) = xy + xz, \tag{A.9a}$$
$$(x+y)z = xz + yz, \tag{A.9b}$$

when x, y and z are any combination of Grassmann and ordinary numbers (and, later, operators). Note that there is no ordering of Grassmann numbers: it is meaningless to ask whether η is larger or smaller than η.

A set of n mutually anti-commuting Grassmann numbers $\{\eta_i\}$ therefore generates a 2^n–dimensional *Grassmann algebra* over a field of ordinary numbers (which we will always take to be the field of complex numbers). The fact that the algebra has dimension 2^n can easily be seen from the fact that any non-zero product of Grassmann numbers can include each η_i at most once.

As we will often need to deal with products of Grassmanns and fermionic operators, we *define* them to anti-commute too:

$$\{\eta_i, c_j\} = \eta_i c_j + c_j \eta_i = 0, \tag{A.10a}$$
$$\{\eta_i, c_j^\dagger\} = \eta_i c_j^\dagger + c_j \eta_i^\dagger = 0. \tag{A.10b}$$

This, along with ordinary multiplication of operators and complex numbers, allows us to extend the Grassmann algebra to include fermionic operators as well.* Note that a Grassmann number "cancels out" the anti-commutativity of a fermionic operator, so that the product $\eta_i c_j$ commutes with everything in the algebra.

We also need to define complex conjugation of Grassmann numbers, so that we can take complex conjugates of products of Grassmanns with complex numbers, or Hermitian conjugates of their products with operators. To do this, we arbitrarily associate pairs of Grassmann numbers from the set $\{\eta_i\}$ and define them to be complex conjugates of each other (which means our set must contain an even number of Grassmanns). To reflect this pairing in our notation, we denote the members of a pair by η_i and $\bar{\eta}_i$, so that

$$\eta_i^* = \bar{\eta}_i, \quad \bar{\eta_i}^* = \eta_i. \tag{A.11}$$

The sum of a Grassmann number and its conjugate, $\eta + \bar{\eta}$, is clearly invariant under complex conjugation, and we will call these "real" Grassmann numbers.

*Were we to include bosonic operators too, defining them to commute with Grassmanns, we would end up with a superalgebra, as used in supersymmetric theories of fundamental particle physics.

We will generally reserve the letters α and β for "real" Grassmann numbers, whereas all other lower-case Greek letters will, unless otherwise indicated, be used for the usual Grassmann numbers.

Finally, for convenience, we arbitrarily define both complex and Hermitian conjugation to reverse the order of products containing Grassmanns:

$$(\eta_i \eta_j)^* = \bar{\eta}_j \bar{\eta}_i = -\bar{\eta}_i \bar{\eta}_j, \tag{A.12a}$$

$$(\eta_i c_j)^\dagger = c_j^\dagger \bar{\eta}_i = -\bar{\eta}_i c_j^\dagger. \tag{A.12b}$$

A.2.2 Grassmann Calculus

The calculus of Grassmann variables is remarkably simple, once one has become accustomed to the anti-commutation, due to the fact that the square of any Grassmann number vanishes. The most general function of a single Grassmann variable must have the form

$$f(\eta) = a + b\eta, \quad a, b \in \mathbb{C} \tag{A.13}$$

and, more generally, any function of n Grassmann variables can contain at most 2^n terms. This implies that the series expansion of any Grassmann function terminates after a finite number of terms and is exact. Thus a brute-force way to deal with any Grassmann function is to calculate its Taylor expansion in full, reducing it to a multi-linear form which is easy to deal with.

Differentiation with respect to a Grassmann variable is defined in analogy with the complex case.

Definition A.1 (Grassmann Differentiation)
Differentiation with respect to a Grassmann variable η_j is defined by

$$\frac{\partial \eta_i}{\partial \eta_j} = \partial_{\eta_i} \eta_j = \delta_{ij}, \tag{A.14}$$

where η_i is also a Grassmann variable.

Note however that, in order to be consistent with the anti-commutation relations, the differential operator ∂_{η_i} must itself anti-commute:

$$\partial_{\eta_i} \eta_j \eta_i = -\partial_{\eta_i} \eta_i \eta_j = -\eta_j = -\eta_j \partial_{\eta_i} \eta_i. \tag{A.15}$$

Definition A.2 (Grassmann Integration)
Integration over a Grassmann variable η is defined by

$$\int \mathrm{d}\eta = 0, \quad \int \mathrm{d}\eta_i \eta_j = \delta_{ij}, \tag{A.16}$$

called the Berezin *integral [Berezin, 1966].*

This definition of integration is *exactly equivalent* to Definition A.1 of Grassmann differentiation:

$$\int \mathrm{d}\eta f(\eta) = \int \mathrm{d}\eta(a + b\eta) = b = \frac{\partial}{\partial\eta}(a + b\eta) = \partial_\eta f(\eta), \tag{A.17}$$

so Grassmann differentiation and integration are the *same* operation.

We will denote multi-dimensional integrals over all variables contained in a vector of Grassmann variables $\boldsymbol{\eta}$ by

$$\int \mathrm{d}\boldsymbol{\eta} \equiv \int \prod_i \mathrm{d}\eta_i. \tag{A.18}$$

We will often take integrals over conjugate pairs of Grassmann variables, and for convenience we denote this

$$\int \mathrm{d}^2\eta \equiv \iint \mathrm{d}\bar{\eta}\mathrm{d}\eta. \tag{A.19}$$

Note that the conjugated variable comes first. Similarly, we will often encounter vectors $\boldsymbol{\eta}$ containing a set of Grassmann variables η_i and their conjugates $\bar{\eta}_i$. The integral over these will be denoted by

$$\int \mathrm{d}^2\boldsymbol{\eta} \equiv \int \prod_i \mathrm{d}^2\eta_i = \int \prod_i \mathrm{d}\bar{\eta}_i\mathrm{d}\eta_i. \tag{A.20}$$

Many results from the calculus of complex variables carry over to the Grassmann calculus, but one must be careful about the ordering of the variables. For example, in differentiating the product of two Grassmann functions $f(\boldsymbol{\eta})g(\boldsymbol{\eta})$, any part of $f(\boldsymbol{\eta})$ that anti-commutes with Grassmann numbers picks up an extra sign from commuting the differential operator past it.

Result A.3 (Grassmann Differentiation Product Rule)
The Grassmann derivative of the product of two functions $f(\boldsymbol{\eta})$ and $g(\boldsymbol{\eta})$, is given by

$$\partial_{\eta_i}\big(f(\boldsymbol{\eta})g(\boldsymbol{\eta})\big) = \big(\partial_{\eta_i} f(\boldsymbol{\eta})\big)g(\boldsymbol{\eta}) + \big(f_+(\boldsymbol{\eta}) - f_-(\boldsymbol{\eta})\big)\big(\partial_{\eta_i} g(\boldsymbol{\eta})\big), \tag{A.21}$$

where $f(\boldsymbol{\eta}) = f_+(\boldsymbol{\eta}) + f_-(\boldsymbol{\eta})$ is the decomposition of the function f into a part $f_+(\boldsymbol{\eta})$ that commutes with all Grassmann numbers, and a part $f_-(\boldsymbol{\eta})$ that anti-commutes.

From this, we can derive the Grassmann version of integration by parts.

Result A.4 (Grassmann Integration by Parts)

$$\int \mathrm{d}\boldsymbol{\eta}\big(\partial_{\eta_i} f(\boldsymbol{\eta})\big) g(\boldsymbol{\eta}) = -\int \mathrm{d}\boldsymbol{\eta}\big(f_+(\boldsymbol{\eta}) - f_-(\boldsymbol{\eta})\big)\partial_{\eta_i} g(\boldsymbol{\eta}). \tag{A.22}$$

Proof The integral of a derivative always vanishes for Grassmann variables, as the derivative will be missing that variable. Therefore, we have

$$\int \mathrm{d}\boldsymbol{\eta}\frac{\partial}{\partial\eta_i}\big(f(\boldsymbol{\eta})g(\boldsymbol{\eta})\big) = \int \mathrm{d}\boldsymbol{\eta}\left(\frac{\partial f(\boldsymbol{\eta})}{\partial\eta_i} g(\boldsymbol{\eta}) + f(-\boldsymbol{\eta})\frac{\partial g(\boldsymbol{\eta})}{\partial\eta_i}\right) = 0, \tag{A.23}$$

from which the result follows immediately. □

A.2.3 Gaussian Grassmann Integrals

The prototypical Gaussian Grassmann integral can be calculated by expanding the exponential as a (finite) series, then applying the rules of Grassmann integration described in the previous section. For $\eta_i \neq \eta_j$ and complex a, we have

$$\begin{aligned}
\iint \mathrm{d}\eta_i \mathrm{d}\eta_j\, e^{-a\eta_i\eta_j} &= \iint \mathrm{d}\eta_i \mathrm{d}\eta_j (1 - a\eta_i\eta_j) && \text{(A.24a)}\\
&= -a \iint \mathrm{d}\eta_i \mathrm{d}\eta_j \eta_i\eta_j && \text{(A.24b)}\\
&= a \iint \mathrm{d}\eta_i \eta_i \mathrm{d}\eta_j \eta_j && \text{(A.24c)}\\
&= a. && \text{(A.24d)}
\end{aligned}$$

The ordering of the variables is important, since a similar manipulation leads to

$$\iint \mathrm{d}\eta_j \mathrm{d}\eta_i\, e^{-a\eta_i\eta_j} = -a. \tag{A.25}$$

This can be extended to a general formula for multi-dimensional Gaussian Grassmann integrals.

Result A.5 (Gaussian Grassmann Integral (1))
Given a vector of Grassmann variables $\boldsymbol{\eta}$, a fixed Grassmann vector $\boldsymbol{\nu}$, and a fixed, complex, invertible, real anti-symmetric matrix A,

$$\int \mathrm{d}\boldsymbol{\eta} \exp\left(-\frac{1}{2}\boldsymbol{\eta}^T A\boldsymbol{\eta} + \boldsymbol{\nu}^T\boldsymbol{\eta}\right) = \mathrm{Pf}\, A \exp\left(-\frac{1}{2}\boldsymbol{\nu}^T A^{-1}\boldsymbol{\nu}\right), \tag{A.26}$$

where $\mathrm{Pf}\, A$ *denotes the Pfaffian of A,*

$$\mathrm{Pf}\, A = \sqrt{\det A}. \tag{A.27}$$

Proof We start by "shifting" the integration variables:

$$\begin{aligned}
&\int \mathrm{d}\boldsymbol{\eta} \exp\left(-\frac{1}{2}\,\boldsymbol{\eta}^T A \boldsymbol{\eta} + \boldsymbol{v}^T \boldsymbol{\eta}\right) \\
&\quad = \int \mathrm{d}\boldsymbol{\eta} \exp\left(-\frac{1}{2}(\boldsymbol{\eta}^T + \boldsymbol{v}^T A^{-1}) A (\boldsymbol{\eta} + A^{-1}\boldsymbol{v}) + \frac{1}{2}\boldsymbol{v}^T A^{-1}\boldsymbol{v}\right) \qquad \text{(A.28a)} \\
&\quad = \int \mathrm{d}\boldsymbol{\eta'} \exp\left(-\frac{1}{2}\,\boldsymbol{\eta'}^T A \boldsymbol{\eta'}\right) \exp\left(\frac{1}{2}\boldsymbol{v}^T A^{-1}\boldsymbol{v}\right), \qquad \text{(A.28b)}
\end{aligned}$$

where $\boldsymbol{\eta'} = \boldsymbol{\eta} + A^{-1}\boldsymbol{\nu}$ and the Jacobian matrix for this transformation is given by

$$J_{ij} = \frac{\partial \eta_i'}{\partial \eta_j} = \delta_{ij}. \qquad \text{(A.29)}$$

The exponentials in Eq. (A.28b) can be separated because both exponents contain only terms with an even number of Grassmann variables, so they commute.

Now any $2N \times 2N$ real anti-symmetric matrix A can be brought into a block-diagonal form by an orthogonal transformation,

$$OAO^T = \Sigma = \begin{pmatrix} 0 & \lambda_1 & & & & \\ -\lambda_1 & 0 & & & & \\ & & 0 & \lambda_2 & & \\ & & -\lambda_2 & 0 & \ddots & \\ & & & \ddots & \ddots & \lambda_N \\ & & & & -\lambda_N & 0 \end{pmatrix}, \qquad \text{(A.30)}$$

where the λ's are real (see Gantmacher [2000, Section XI.4]). For a $(2N+1) \times (2N+1)$ matrix, there will be an extra row and column of zeros. The Pfaffian of A is defined as

$$\mathrm{Pf}\, A = \prod_i \lambda_i. \qquad \text{(A.31)}$$

From Eq. (A.30), the eigenvalues of A occur in pairs $\pm i\lambda_i$, so

$$\det A = \prod_i \lambda_i^2 = (\mathrm{Pf}\, A)^2. \qquad \text{(A.32)}$$

Using Eq. (A.30),

$$\int \mathrm{d}\boldsymbol{\eta'} \exp\left(-\frac{1}{2}\boldsymbol{\eta}^T O^T \Sigma O \boldsymbol{\eta}\right) = \int \mathrm{d}\boldsymbol{\eta''} \exp\left(-\frac{1}{2}\boldsymbol{\eta''}^T \Sigma \boldsymbol{\eta''}\right), \qquad \text{(A.33)}$$

where we have "rotated" to a new set of Grassmann variables $\boldsymbol{\eta''} = O\boldsymbol{\eta}$, and the Jacobian is given by

$$J'_{ij} = \frac{\partial \eta''_i}{\partial \eta'_j} = O_{ij}, \quad \det J' = \det O = 1. \tag{A.34}$$

Finally, substituting Eq. (A.33) in Eq. (A.28b),

$$\begin{aligned}
&\int \mathrm{d}\boldsymbol{\eta} \exp\left(-\frac{1}{2}\,\boldsymbol{\eta}^T A \boldsymbol{\eta} + \boldsymbol{v}^T \boldsymbol{\eta}\right) \\
&\quad = \exp\left(\frac{1}{2}\boldsymbol{v}^T A^{-1} \boldsymbol{v}\right) \int \mathrm{d}\boldsymbol{\eta} \exp\left(-\frac{1}{2}\boldsymbol{\eta''}^T \Sigma \boldsymbol{\eta''}\right) && \text{(A.35a)} \\
&\quad = \exp\left(\frac{1}{2}\boldsymbol{v}^T A^{-1} \boldsymbol{v}\right) \int \mathrm{d}\boldsymbol{\eta} \exp\left(-\frac{1}{2}\sum_i (\lambda_i \eta_i \eta_{i+1} - \lambda_i \eta_{i+1} \eta_i)\right) && \text{(A.35b)} \\
&\quad = \exp\left(\frac{1}{2}\boldsymbol{v}^T A^{-1} \boldsymbol{v}\right) \prod_i \int \mathrm{d}\eta_i \mathrm{d}\eta_{i+1} e^{-\lambda_i \eta_i \eta_{i+1}} && \text{(A.35c)} \\
&\quad = \mathrm{Pf}\, A \, \exp\left(\frac{1}{2}\boldsymbol{v}^T A^{-1} \boldsymbol{v}\right). && \text{(A.35d)}
\end{aligned}$$

The final line follows from Eq. (A.31) for the Pfaffian of A and the Gaussian integral formula of Eq. (A.25). □

By taking successive derivatives of this result with respect to the components of $\boldsymbol{v}$ at $\boldsymbol{v} = 0$, we obtain expressions for the product of a Grassmann monomial with a Gaussian.

Result A.6 (Grassmann Monomial times Gaussian Integral)
Given an n^{th} order monomial of Grassmann variables, $\prod_{i=1}^n \eta_i$, and a fixed, complex, invertible, real anti-symmetric matrix A,

$$\begin{aligned}
&\int \mathrm{d}\boldsymbol{\eta} \exp\left(-\frac{1}{2}\boldsymbol{\eta}^T A\, \boldsymbol{\eta}\right) \prod_{i=1}^n \eta_i \\
&\qquad = \begin{cases} 0 & n \text{ odd} \\ \left(\frac{1}{2}\right)^{n/2} \mathrm{Pf}\, A \sum_\pi \left(\mathrm{sgn}(\pi) \prod_{i=1}^{n/2} A^{-1}_{\pi(2i-1),\pi(2i)}\right) & n \text{ even.} \end{cases}
\end{aligned} \tag{A.36}$$

The sum is over all permutations π of the indices $i = 1, \ldots, n$, where $\mathrm{sgn}(\pi)$ denotes the signum of the permutation. $\mathrm{Pf}(A)$ denotes the Pfaffian of the anti-symmetric matrix A, equal to the square root of its determinant.

Proof Taking derivatives of the left hand side of the Gaussian integral formula of Eq. (A.26) with respect to components $\nu_1, \ldots, \nu_n$ at $\boldsymbol{\nu} = 0$,

$$\partial_{\nu_1} \ldots \partial_{\nu_n} \Big|_{\boldsymbol{v}=0} \int \mathrm{d}\boldsymbol{\eta} \exp\left(-\tfrac{1}{2}\,\boldsymbol{\eta}^T A \boldsymbol{\eta} + \boldsymbol{\nu}^T \boldsymbol{\eta}\right) = \int \mathrm{d}\boldsymbol{\eta} \exp\left(-\tfrac{1}{2}\boldsymbol{\eta}^T A \boldsymbol{\eta}\right) \prod_{i=1}^{n} \eta_i. \tag{A.37}$$

Assume first that n is even. Taking the same derivative of the right hand side of Eq. (A.26), and denoting permutations of the labels $1 \ldots n$ by π,

$$\begin{aligned}
&\partial_{\nu_1} \ldots \partial_{\nu_n} \operatorname{Pf} A \, \exp\left(-\frac{1}{2}\,\boldsymbol{\nu}^T A^{-1} \boldsymbol{\nu}\right)\Big|_{\boldsymbol{\nu}=0} \\
&\quad = \operatorname{Pf} A \, \partial_{\nu_1} \ldots \partial_{\nu_n} \prod_{p,q} \left(1 - \frac{1}{2}\nu_p A^{-1}_{pq} \nu_q\right)\Big|_{\boldsymbol{\nu}=0} \qquad \text{(A.38a)} \\
&\quad = \operatorname{Pf} A \, \left(\frac{1}{2}\right)^{n/2} \partial_{\nu_1} \ldots \partial_{\nu_n} \sum_{\pi} \left(\prod_{i=1}^{n/2} \left(-\nu_{\pi(2i-1)} A^{-1}_{\pi(2i-1),\pi(2i)} \nu_{\pi(2i)}\right)\right)\Big|_{\boldsymbol{\nu}=0} \qquad \text{(A.38b)}
\end{aligned}$$

dropping terms whose derivatives vanish,

$$\begin{aligned}
&= \operatorname{Pf} A \, \left(\frac{1}{2}\right)^{n/2} \times \\
&\qquad \sum_{\pi} \left(\operatorname{sgn}(\pi) \partial_{\nu_{\pi(1)}} \ldots \partial_{\nu_{\pi(n)}} \prod_{i=1}^{n/2} \left(-\nu_{\pi(2i-1)} A^{-1}_{\pi(2i-1),\pi(2i)} \nu_{\pi(2i)}\right)\Big|_{\boldsymbol{\nu}=0}\right) \qquad \text{(A.38c)}
\end{aligned}$$

anti-commuting the differential operators ∂_{ν_i} to order them according to the permutation π,

$$\begin{aligned}
&= \operatorname{Pf} A \, \left(\frac{1}{2}\right)^{n/2} \times \\
&\qquad \sum_{\pi} \left(\operatorname{sgn}(\pi) \prod_{i=1}^{n/2} \partial_{\nu_{\pi(2i-1)}} \partial_{\nu_{\pi(2i)}} \left(-\nu_{\pi(2i-1)} A^{-1}_{\pi(2i-1),\pi(2i)} \nu_{\pi(2i)}\right)\Big|_{\boldsymbol{\nu}=0}\right) \qquad \text{(A.38d)} \\
&= \operatorname{Pf} A \, \left(\frac{1}{2}\right)^{n/2} \sum_{\pi} \operatorname{sgn}(\pi) \prod_{i=1}^{n/2} A^{-1}_{\pi(2i-1),\pi(2i)}. \qquad \text{(A.38e)}
\end{aligned}$$

Terms that do not involve all $\nu_1, \ldots, \nu_n$ will vanish in Eq. (A.38b) due to the rules of Grassmann differentiation (Definition A.1). Terms involving one or more extra $\nu_{i>n}$ will vanish when evaluating the term at $\boldsymbol{\nu} = 0$.

If on the other hand n is odd, it becomes impossible to pair up the differential operators ∂_{ν_i}, and the terms inside the product in Eq. (A.38b) will always either include an extra $\nu_{i>n}$ and vanish when evaluated at $\boldsymbol{\nu} = 0$, or will be missing one of the $\nu_{1\leq i\leq n}$ and vanish due to the rules of Grassmann differentiation (Definition A.1). □

A final variant of Gaussian Grassmann integration leads to an integral formula that will be useful later. (Conjugates of Grassmann variables will be used for later notational convenience; since conjugation is simply an arbitrary mapping between pairs of Grassmann numbers, and the complex conjugation operation is never used here, the result would be equally valid if all the variables were written without the conjugates.)

Result A.7 (Gaussian Grassmann Integral (2))
Given a vector $\boldsymbol{\eta}$ containing Grassmann variables η_i and their conjugates $\bar{\eta}_i$, and fixed Grassmann vectors $\boldsymbol{\mu}$ and $\boldsymbol{\nu}$,

$$\int \mathrm{d}^2\boldsymbol{\eta} \exp\left(-\sum_i \bar{\eta}_i\eta_i + \bar{\nu}_i\eta_i + \bar{\eta}_i\mu_i\right) = e^{\sum_i \bar{\nu}_i\mu_i}. \tag{A.39}$$

Proof "Shifting" the integration variables to

$$\eta_i' = \eta_i + \mu_i, \quad \bar{\eta}_i = \bar{\eta}_i + \bar{\nu}_i \tag{A.40}$$

with Jacobian matrix

$$J_{ij} = \frac{\partial\eta_i'}{\partial\eta_j} = \delta_{ij}, \tag{A.41}$$

we have

$$\int \mathrm{d}^2\boldsymbol{\eta} \exp\left(-\sum_i \bar{\eta}_i\eta_i + \bar{\nu}_i\eta_i + \bar{\eta}_i\mu_i\right) \tag{A.42a}$$

$$= \int \mathrm{d}^2\boldsymbol{\eta} \exp\left(-\sum_i (\bar{\eta}_i + \bar{\nu}_i)(\eta_i + \mu_i) - \bar{\nu}_i\mu_i\right) \tag{A.42b}$$

$$= e^{\sum_i \bar{\nu}_i\mu_i} \int \mathrm{d}^2\boldsymbol{\eta}' e^{-\sum_i \bar{\eta}_i'\eta_i'} \tag{A.42c}$$

$$= e^{\sum_i \bar{\nu}_i\mu_i}, \tag{A.42d}$$

where the final equality follows from the Gaussian integral formula given in Eq. (A.24a). □

A.2.4 Grassmann Fourier Transforms

A Grassmann integral transformation that is analogous to the usual Fourier transform of complex variables can be defined, and will be important in deriving the most useful of the fermionic phase-space representations in Section A.4. We give here the results that will be required. For a more complete treatment of Grassmann Fourier transforms, see Cahill and Glauber [1999].

Definition A.8 (Grassmann Fourier Transform)
The Grassmann Fourier transform $\tilde{f}(\boldsymbol{\xi})$ of a function $f(\boldsymbol{\eta})$ of Grassmann variables η_i and their conjugates $\bar{\eta}_i$ is defined by

$$\tilde{f}(\boldsymbol{\xi}) = \int \mathrm{d}^2\boldsymbol{\eta} \exp\left(-\sum_i \bar{\eta}_i \xi_i + \eta_i \bar{\xi}_i\right) f(\boldsymbol{\eta}). \tag{A.43}$$

A Grassmann analogue of the Dirac δ–function can also be defined.

Lemma A.9 (Grassmann $\boldsymbol{\delta}$–Function)
The function of Grassmann variables η_i given by

$$\delta(\boldsymbol{\eta} - \boldsymbol{\xi}) = \prod_i (\eta_i - \xi_i) = \int \mathrm{d}^2\boldsymbol{\mu} \exp\left(-\sum_i (\eta_i - \xi_i)\mu_i\right) \tag{A.44}$$

has the property that, for any function $f(\boldsymbol{\eta})$,

$$\int \mathrm{d}\boldsymbol{\eta}\, \delta(\boldsymbol{\eta} - \boldsymbol{\xi}) f(\boldsymbol{\eta}) = \int \mathrm{d}\boldsymbol{\eta}\, \delta(\boldsymbol{\xi} - \boldsymbol{\eta}) f(\boldsymbol{\eta}) = f(\boldsymbol{\xi}). \tag{A.45}$$

Proof For any monomial $(\prod_{j<i} \eta_i)\eta_i(\prod_{k>i} \eta_k)$ containing the variable η_i, where the product over j contains n variables,

$$\int \mathrm{d}\eta_i\, (\eta_i - \xi_i) \prod_{j,k\neq i} \eta_j \eta_i \eta_k = \int \mathrm{d}\eta_i\, (\eta_i - \xi_i)(-1)^n \eta_i \prod_{j,k\neq i} \eta_j \eta_k \tag{A.46a}$$

$$= \int \mathrm{d}\eta_i\, \eta_i \xi_i (-1)^n \prod_{j,k\neq i} \eta_j \eta_k \tag{A.46b}$$

$$= \xi_i (-1)^n \prod_{j,k\neq i} \eta_j \eta_k \tag{A.46c}$$

$$= \prod_{j,k\neq i} \eta_j \xi_i \eta_k. \tag{A.46d}$$

The same final result Eq. (A.45) is obtained if $(\eta_i - \xi_i)$ is changed to $(\xi_i - \eta_i)$, as two factors of -1 cancel. Since any Grassmann function can be expanded

as a finite sum of such monomials, we can carry out the integrals over η_i in Eq. (A.45) one by one, using Eq. (A.46) for each monomial in the expansion. This will successively replace each η_i by ξ_i, resulting in Eq. (A.45).

To show that the two expressions for the δ–function in Eq. (A.44) are equivalent, all that is necessary is to expand the exponential and do the integration using the Gaussian integral formula of Eq. (A.24a):

$$\int \mathrm{d}^2\boldsymbol{\mu} \exp\Big(-\sum_i (\eta_i - \xi_i)\mu_i\Big) = \prod_i \int \mathrm{d}\mu_i \exp(1 - (\eta_i - \xi_i)\mu_i) = \prod_i (\eta_i - \xi_i). \tag{A.47}$$

□

The Grassmann δ–function can be used to prove the following relation that we will need later to derive one of the fermionic phase-space representations. It is analogous to a version of Parseval's theorem for normal Fourier transforms.

Lemma A.10 (Grassmann Parseval Theorem)
For functions $f(\boldsymbol{\eta})$ and $g(\boldsymbol{\eta})$ of Grassmann variables η_i and $\bar{\eta}_i$, and their Grassmann Fourier transforms $\tilde{f}(\boldsymbol{\xi})$ and $\tilde{g}(\boldsymbol{\xi})$,

$$\int \mathrm{d}\boldsymbol{\xi} \tilde{f}(\boldsymbol{\xi}) \tilde{g}(-\boldsymbol{\xi}) = \int \mathrm{d}\boldsymbol{\eta} f(\boldsymbol{\eta}) g(\boldsymbol{\eta}). \tag{A.48}$$

Proof This is easily proven by inserting the integral form of the Grassmann δ–function from Lemma A.9:

$$\begin{aligned}
&\int \mathrm{d}^2\boldsymbol{\xi}\, \tilde{f}(\boldsymbol{\xi})\, \tilde{g}(-\boldsymbol{\xi}) \\
&\quad = \int \mathrm{d}^2\boldsymbol{\xi} \int \mathrm{d}^2\boldsymbol{\eta}\, e^{-\sum_i \bar{\eta}_i\xi_i + \eta_i\bar{\xi}_i} f(\boldsymbol{\eta}) \int \mathrm{d}^2\boldsymbol{\mu}\, e^{\sum_i \bar{\mu}_i\xi_i + \mu_i\bar{\xi}_i} g(\boldsymbol{\mu}) && \text{(A.49a)} \\
&\quad = \int \mathrm{d}^2\boldsymbol{\xi}\, \mathrm{d}^2\boldsymbol{\eta}\, \mathrm{d}^2\boldsymbol{\mu}\, e^{-\sum_i (\bar{\eta}_i - \bar{\mu}_i)\xi_i + (\eta_i - \mu_i)\bar{\xi}_i} f(\boldsymbol{\eta}) g(\boldsymbol{\mu}) && \text{(A.49b)} \\
&\quad = \int \mathrm{d}^2\boldsymbol{\eta}\, f(\boldsymbol{\eta}) \int \mathrm{d}^2\boldsymbol{\mu}\, \delta(\boldsymbol{\eta} - \boldsymbol{\mu})\, g(\boldsymbol{\mu}) && \text{(A.49c)} \\
&\quad = \int \mathrm{d}^2\boldsymbol{\eta}\, f(\boldsymbol{\eta})\, g(\boldsymbol{\eta}). && \text{(A.49d)}
\end{aligned}$$

□

A.3 Displacement Operators and Coherent States

We now turn to systems of n fermionic modes, with creation and annihilation operators $c_i^\dagger$ and c_i obeying the usual anti-commutation relations

$$\{c_i, c_j\} = \{c_i^\dagger, c_i^\dagger\} = 0, \tag{A.50a}$$

$$\{c_i, c_j^\dagger\} = \delta_{ij}. \tag{A.50b}$$

We define the vacuum $|0\rangle$ of the set of modes to be the state annihilated by all annihilation operators

$$c_i |0\rangle = 0 \quad \forall i. \tag{A.51}$$

As it will often be convenient to work in the "real" representation, we also define Majorana operators

$$x_i = c_i^\dagger + c_i, \quad p_i = \frac{c_i^\dagger - c_i}{i}, \tag{A.52}$$

which obey the equivalent anti-commutation relations

$$\{r_i, r_j\} = 2\delta_{ij}. \tag{A.53}$$

For convenience, we will often use the notation

$$r_{2i-1} = x_i, \quad r_{2i} = p_i. \tag{A.54}$$

A.3.1 Definitions and Basic Properties

In analogy with the bosonic phase-space formalism, we start by defining fermionic displacement operators.

Definition A.11 (Fermionic Displacement Operator)
Fermionic displacement operators, parameterized by a vector $\boldsymbol{\eta}$ of Grassmann numbers η_i and their conjugates $\bar{\eta}_i$, are defined by

$$D(\boldsymbol{\eta}) = \exp\left(-\sum_i \eta_i c_i^\dagger + \bar{\eta}_i c_i\right). \tag{A.55}$$

Note that all summands with different indices i commute with one another, as they consist of products of a Grassmann with a fermionic operator, and the operators in each summand act on different modes; the anti-commutativity of the Grassmann numbers "cancels" that of the fermionic operators. Also,

the commutator of the two terms within each summand $[\eta_i c_i^\dagger, \bar{\eta}_i c_i] = \bar{\eta}_i \eta_i$ is product of two Grassmanns, so commutes with everything. Thus, using the Cambell-Baker-Hausdorf relation for exponents that commute with their commutator,

$$e^{A+B} = e^A e^B e^{-\frac{1}{2}[A,B]}, \tag{A.56}$$

we can decompose the displacement operator in a number of ways:

$$D(\boldsymbol{\eta}) = e^{-\sum_i \eta_i c_i^\dagger} e^{-\sum_i \bar{\eta}_i c_i} e^{-\frac{1}{2}\sum_i \bar{\eta}_i \eta_i} \tag{A.57}$$

$$= \prod_i \left(1 - \eta_i c_i^\dagger - \bar{\eta}_i c_i + \tfrac{1}{2}\eta_i c_i^\dagger \bar{\eta}_i c_i + \tfrac{1}{2}\bar{\eta}_i c_i \eta_i c_i^\dagger\right). \tag{A.58}$$

Equivalently, we can write the displacement operator in the "real" representation,

$$D(\boldsymbol{\alpha}) = \exp\left(-\sum_i \alpha_{2i-1} x_i + \alpha_{2i} p_i\right) = \exp\left(-\sum_i \alpha_i r_i\right), \tag{A.59}$$

where $\boldsymbol{\alpha}$ is a vector of "real" Grassmann numbers

$$\alpha_{2i-1} = \frac{\bar{\eta}_i + \eta_i}{2}, \quad \alpha_{2i} = \frac{\bar{\eta}_i - \eta_i}{2i}. \tag{A.60}$$

If we now apply an annihilation operator to the state produced by displacing the fermionic vacuum,

$$c_j D(\boldsymbol{\eta}) |0\rangle = c_j \prod_i \left(1 - \eta_i c_i^\dagger - \bar{\eta}_i c_i + \tfrac{1}{2}\eta_i c_i^\dagger \bar{\eta}_i c_i + \tfrac{1}{2}\bar{\eta}_i c_i \eta_i c_i^\dagger\right) |0\rangle \tag{A.61a}$$

$$= \prod_{i\neq j} \left(1 - \eta_i c_i^\dagger - \tfrac{1}{2}\eta_i \bar{\eta}_i\right) c_j \left(1 - \eta_j c_j^\dagger - \tfrac{1}{2}\eta_j \bar{\eta}_j\right) |0\rangle \tag{A.61b}$$

$$= \prod_{i\neq j} \left(1 - \eta_i c_i^\dagger - \tfrac{1}{2}\eta_i \bar{\eta}_i\right) \eta_j |0\rangle \tag{A.61c}$$

$$= \prod_i \left(1 - \eta_i c_i^\dagger - \tfrac{1}{2}\eta_i \bar{\eta}_i\right) \left(1 - \eta_j c_j^\dagger - \tfrac{1}{2}\eta_j \bar{\eta}_j\right) \eta_j |0\rangle \tag{A.61d}$$

since $\eta_j^2 = 0$,

$$= \eta_j \exp\left(-\sum_i \eta_i c_i^\dagger + \bar{\eta}_i c_i\right) |0\rangle \tag{A.61e}$$

$$= \eta_j D(\boldsymbol{\eta}) |0\rangle. \tag{A.61f}$$

This shows that the displaced vacuum is a simultaneous eigenstate of all the annihilation operators. That is, it is a fermionic coherent state. This gives a concrete second-quantization description of fermionic coherent states, introduced in a more abstract way in Section A.1.

Definition A.12 (Fermionic Coherent State)
A fermionic coherent state $|\boldsymbol{\eta}\rangle$ of a collection of fermionic modes with creation and annihilation operators $c_i^\dagger$ and c_i is defined to be a simultaneous eigenstate of all the annihilation operators,

$$c_i |\boldsymbol{\eta}\rangle = \eta_i |\boldsymbol{\eta}\rangle , \tag{A.62}$$

and is generated by the action of a displacement operator on the vacuum state:

$$|\boldsymbol{\eta}\rangle = D(\boldsymbol{\eta}) |0\rangle = \exp\left(-\frac{1}{2}\sum_i \bar{\eta}_i \eta_i\right) \exp\left(-\sum_i \eta_i c_i^\dagger\right) |0\rangle . \tag{A.63}$$

The duals $\langle\boldsymbol{\eta}|$ of these coherent states are obtained by taking the Hermitian conjugate of Eq. (A.63),

$$\langle\boldsymbol{\eta}| = \langle 0| D^\dagger(\boldsymbol{\eta}) = \langle 0| \exp\left(\sum_i \bar{\eta}_i c_i\right) \exp\left(-\frac{1}{2}\sum_i \bar{\eta}_i \eta_i\right). \tag{A.64}$$

They are clearly right-eigenstates of all creation operators:

$$\langle\boldsymbol{\eta}| c_i^\dagger = \langle\boldsymbol{\eta}| \bar{\eta}_i. \tag{A.65}$$

We will frequently need to know the overlap between two coherent states.

Result A.13 (Coherent State Overlap)
The overlap of two fermionic coherent states $|\boldsymbol{\eta}\rangle$ and $|\boldsymbol{\xi}\rangle$ is given by

$$\langle\boldsymbol{\eta}|\boldsymbol{\xi}\rangle = \exp\left(\sum_i \bar{\eta}_i \xi_i - \tfrac{1}{2}\bar{\eta}_i \eta_i - \tfrac{1}{2}\bar{\xi}_i \xi_i\right). \tag{A.66}$$

Proof Writing the coherent states as displaced vacuum states,

$$\begin{aligned}
\langle\boldsymbol{\eta}|\boldsymbol{\xi}\rangle &= \langle 0| D^\dagger(\boldsymbol{\eta}) D(\boldsymbol{\xi}) |0\rangle && \text{(A.67a)}\\
&= \langle 0| \prod_i \left(1 - c_i\bar{\eta}_i - c_i^\dagger \eta_i + \tfrac{1}{2} c_i^\dagger \eta_i c_i \bar{\eta}_i + \tfrac{1}{2} c_i \bar{\eta}_i c_i^\dagger \eta_i\right) \times \\
&\qquad\qquad \left(1 - \xi_i c_i^\dagger - \bar{\xi}_i c_i + \tfrac{1}{2}\xi_i c_i^\dagger \bar{\xi}_i c_i + \tfrac{1}{2}\bar{\xi}_i c_i \xi_i c_i^\dagger\right) |0\rangle && \text{(A.67b)}\\
&= \prod_i \langle 0_i| \left(1 - c_i\bar{\eta}_i - \tfrac{1}{2}\bar{\eta}_i \eta_i c_i c_i^\dagger\right) \left(1 - \xi_i c_i^\dagger - \tfrac{1}{2}\bar{\xi}_i \xi_i c_i c_i^\dagger\right) |0_i\rangle && \text{(A.67c)}\\
&= \prod_i \left(1 + \bar{\eta}_i \xi_i - \tfrac{1}{2}\bar{\eta}_i \eta_i - \tfrac{1}{2}\bar{\xi}_i \xi_i + \tfrac{1}{4}\bar{\eta}_i \eta_i \tfrac{1}{2}\bar{\xi}_i \xi_i\right) && \text{(A.67d)}\\
&= \exp\left(\sum_i \bar{\eta}_i \xi_i - \tfrac{1}{2}(\bar{\eta}_i \eta_i + \bar{\xi}_i \xi_i)\right). && \text{(A.67e)}
\end{aligned}$$

□

Finally, two successive displacements are, up to Grassmann number factors, equivalent to a displacement by the sum.

Result A.14 (Displacement Operator Product)

$$D(\boldsymbol{\mu})D(\boldsymbol{\nu}) = \exp\left(-\frac{1}{2}\sum_i \bar{\mu}_i\nu_i + \mu_i\bar{\nu}_i\right) D(\boldsymbol{\mu}+\boldsymbol{\nu}). \tag{A.68}$$

Proof The only terms in the exponents of the two displacement operators (Definition A.11) that do *not* commute are those acting on the same mode. The commutator of these is

$$[(\mu_i c_i^\dagger + \bar{\mu}_i c_i), (\nu_i c_i^\dagger + \bar{\nu}_i c_i)] = -\mu_i\bar{\nu}_i\{c_i^\dagger, c_i\} - \bar{\mu}_i\nu_i\{c_i, c_i^\dagger\} = -\mu_i\bar{\nu}_i - \bar{\mu}_i\nu_i. \tag{A.69}$$

Since the commutator contains only pairs of Grassmann numbers, it commutes with any combination of Grassmanns and operators, so we can use the Cambell-Baker-Hausdorf relation Eq. (A.56) again to give

$$D(\boldsymbol{\mu})D(\boldsymbol{\nu}) = \exp\left(-\sum_i \mu_i c_i^\dagger + \bar{\mu}_i c_i\right)\exp\left(-\sum_i \nu_i c_i^\dagger + \bar{\nu}_i c_i\right) \tag{A.70a}$$

$$= \exp\left(-\frac{1}{2}\sum_i [(\mu_i c_i^\dagger + \bar{\mu}_i c_i), (\nu_i c_i^\dagger + \bar{\nu}_i c_i)]\right)\times \exp\left(-\sum_i (\mu_i+\nu_i)c_i^\dagger + (\bar{\mu}_i+\bar{\nu}_i)c_i\right) \tag{A.70b}$$

$$= \exp\left(-\frac{1}{2}\sum_i \bar{\mu}_i\nu_i + \mu_i\bar{\nu}_i\right) D(\boldsymbol{\mu}+\boldsymbol{\nu}). \tag{A.70c}$$

□

A.3.2 Completeness Properties

The coherent states and displacement operators introduced in the previous section respectively form complete bases for states and operators in the fermionic Fock space. Therefore, any state or operator can be decomposed in terms of them. (In fact, both the coherent states and displacement operators form *over-complete* bases.) We now prove this.

Theorem A.15 (Completeness of Coherent States)
The fermionic coherent states form an over-complete basis for the fermionic Fock space $\mathcal{F}$, so that any state $|\psi\rangle$ can be written as

$$|\psi\rangle = \int \mathrm{d}^2\boldsymbol{\eta}\, \langle\boldsymbol{\eta}|\psi\rangle\, |\boldsymbol{\eta}\rangle \tag{A.71}$$

Proof We will prove that the coherent states give rise to a resolution of identity,

$$\int \mathrm{d}^2\boldsymbol{\eta}\, |\boldsymbol{\eta}\rangle\langle\boldsymbol{\eta}| = \mathbb{1}_{\mathcal{F}} \tag{A.72}$$

(where $\mathbb{1}_{\mathcal{F}}$ is the identity operator on the Fock space), as this implies that they must span the entire Fock space, and leads directly to Eq. (A.71).

The easiest way to prove this is via Schur's lemma: if we can show that the left hand side of Eq. (A.72) commutes with all operators in the Fock space algebra, then Schur's lemma implies that it must be proportional to the identity. First, we show it commutes with the annihilation operators:

$$\begin{aligned}
c_i \int \mathrm{d}^2\boldsymbol{\eta}\, |\boldsymbol{\eta}\rangle\langle\boldsymbol{\eta}| &= \int \mathrm{d}^2\boldsymbol{\eta}\, c_i\, |\boldsymbol{\eta}\rangle\langle\boldsymbol{\eta}| && \text{(A.73a)}\\
&= \int \mathrm{d}^2\boldsymbol{\eta}\, \eta_i D(\boldsymbol{\eta})\, |0\rangle\langle 0|\, D^\dagger(\boldsymbol{\eta}) && \text{(A.73b)}\\
&= \int \mathrm{d}^2\boldsymbol{\eta}\, \eta_i e^{-\frac{1}{2}\sum_i \bar{\eta}_i\eta_i} e^{-\sum_i \eta_i c_i^\dagger}\, |0\rangle\langle 0|\, e^{-\frac{1}{2}\sum_i \bar{\eta}_i\eta_i} e^{-\sum_i c_i\bar{\eta}_i} && \text{(A.73c)}\\
&= \int \mathrm{d}^2\boldsymbol{\eta}\, \left(-\partial_{\bar{\eta}_i} e^{-\sum_i \bar{\eta}_i\eta_i} e^{\sum_i \eta_i c_i^\dagger}\, |0\rangle\right)\left(\langle 0|\, e^{-\sum_i c_i\bar{\eta}_i}\right) && \text{(A.73d)}\\
&= \int \mathrm{d}^2\boldsymbol{\eta}\, \left(e^{-\sum_i \bar{\eta}_i\eta_i} e^{\sum_i \eta_i c_i^\dagger}\, |0\rangle\right)\left(\partial_{\bar{\eta}_i}\, \langle 0|\, e^{-\sum_i c_i\bar{\eta}_i}\right) && \text{(A.73e)}
\end{aligned}$$

integrating by parts using Result A.4, and noting that the function in the left hand bracket commutes with all Grassmann variables,

$$\begin{aligned}
&= \int \mathrm{d}^2\boldsymbol{\eta}\, \left(e^{-\sum_i \bar{\eta}_i\eta_i} e^{\sum_i \eta_i c_i^\dagger}\, |0\rangle\right)\left(\langle 0|\, e^{-\sum_i c_i\bar{\eta}_i} c_i\right) && \text{(A.73f)}\\
&= \int \mathrm{d}^2\boldsymbol{\eta}\, |\boldsymbol{\eta}\rangle\langle\boldsymbol{\eta}|\, c_i. && \text{(A.73g)}
\end{aligned}$$

Taking the Hermitian conjugate of both sides shows that the left hand side of Eq. (A.72) also commutes with all creation operators. Since the creation and annihilation operators generate the complete Fock space algebra, it therefore commutes with all elements of the algebra, which concludes the proof. □

If we expand the exponential containing the creation operators in Definition A.12 of the fermionic coherent states, we see that a coherent state consists of a superposition of all possible number states, with Grassmann number coefficients. These coefficients will pick up a sign factor when commuted

past another Grassmann number, or past a product of an odd number of creation and annihilation operators. Therefore, if we use the coherent state resolution of identity (Eq. (A.72)) inside a trace over an operator A, and want to cyclically permute the bra to the front, we again have to be careful about minus signs; the bra will pick up an extra sign factor when commuted past the ket.

$$\mathrm{Tr}[A] = \mathrm{Tr}\left[A \int \mathrm{d}^2\boldsymbol{\eta}\, |\boldsymbol{\eta}\rangle\langle\boldsymbol{\eta}|\right] = \int \mathrm{d}^2\boldsymbol{\eta}\, \langle -\boldsymbol{\eta}|\, A\, |\boldsymbol{\eta}\rangle \tag{A.74}$$

Completeness of the coherent states allows us to prove the following Lemma, which we will need to prove completeness of the displacement operators.

Lemma A.16 (Coherent State Diadic)
For any diadic $|\boldsymbol{\mu}\rangle\langle\boldsymbol{\nu}|$ formed from coherent states $|\boldsymbol{\mu}\rangle$ and $|\boldsymbol{\nu}\rangle$, the following operator identity holds:

$$|\boldsymbol{\mu}\rangle\langle\boldsymbol{\nu}| = \int \mathrm{d}^2\boldsymbol{\eta}\, \left\langle\boldsymbol{\nu}\right| D^\dagger(\boldsymbol{\eta}) \left|\boldsymbol{\mu}\right\rangle D(\boldsymbol{\eta}). \tag{A.75}$$

Proof Introducing two more coherent states $|\boldsymbol{\gamma}\rangle$ and $|\boldsymbol{\delta}\rangle$,

$$\int \mathrm{d}^2\boldsymbol{\eta}\, \left\langle\boldsymbol{\nu}\right| D^\dagger(\boldsymbol{\eta}) \left|\boldsymbol{\mu}\right\rangle \langle\boldsymbol{\gamma}|\, D(\boldsymbol{\eta})\, |\boldsymbol{\delta}\rangle$$

$$= \int \mathrm{d}^2\boldsymbol{\eta}\, \langle\boldsymbol{\nu}|\, e^{\sum_i \eta_i c_i^\dagger} e^{-\frac{1}{2}\sum_i \bar{\eta}_i\eta_i} e^{\sum_i \bar{\eta}_i c_i} |\boldsymbol{\mu}\rangle \langle\boldsymbol{\gamma}|\, e^{-\sum_i \eta_i c_i^\dagger} e^{-\frac{1}{2}\sum_i \bar{\eta}_i\eta_i} e^{-\sum_i \bar{\eta}_i c_i} |\boldsymbol{\delta}\rangle \tag{A.76a}$$

$$= \langle\boldsymbol{\nu}|\boldsymbol{\mu}\rangle \langle\boldsymbol{\gamma}|\boldsymbol{\delta}\rangle \int \mathrm{d}^2\boldsymbol{\eta}\, \exp\left(-\sum_i \bar{\eta}_i\eta_i - (\bar{\gamma}_i - \bar{\nu}_i)\eta_i - \bar{\eta}_i(\mu_i - \delta_i)\right) \tag{A.76b}$$

$$= \langle\boldsymbol{\nu}|\boldsymbol{\mu}\rangle \langle\boldsymbol{\gamma}|\boldsymbol{\delta}\rangle \exp\left(\sum_i (\bar{\gamma}_i - \bar{\nu}_i)(\mu_i - \delta_i)\right) \tag{A.76c}$$

evaluating the Gaussian integral using Result A.7,

$$= \langle\boldsymbol{\gamma}|\boldsymbol{\mu}\rangle \langle\boldsymbol{\nu}|\boldsymbol{\delta}\rangle\,. \tag{A.76d}$$

The final equality follows from expanding the inner products using Result A.13 for the overlap of two coherent states, and noticing that the terms in the exponential effectively reorder the coherent states in the inner products.

Since this integral identity holds for all $|\boldsymbol{\gamma}\rangle$ and $|\boldsymbol{\delta}\rangle$ and the coherent states form a complete basis, Eq. (A.75) is true for all matrix elements of the operators on each side of the equality, therefore holds for the operators themselves. □

Completeness of the displacement operators is now easy to prove from completeness of the coherent states.

Theorem A.17 (Completeness of Displacement Operators)
The fermionic displacement operators form an over-complete basis for operators on the fermionic Fock space $\mathcal{F}$, so that any operator A can be decomposed as

$$A = \int \mathrm{d}^2\boldsymbol{\eta}\mathrm{d}^2\boldsymbol{\mu}\ \langle\boldsymbol{\mu}|\, AD^\dagger(\boldsymbol{\eta})\,|\boldsymbol{\mu}\rangle\, D(\boldsymbol{\eta}). \tag{A.77}$$

Proof Inserting the coherent state resolution of identity (Eq. (A.72)) twice, we have

$$A = \left(\int \mathrm{d}^2\boldsymbol{\mu}\,|\boldsymbol{\mu}\rangle\langle\boldsymbol{\mu}|\right) A \left(\int \mathrm{d}^2\boldsymbol{\nu}\,|\boldsymbol{\nu}\rangle\langle\boldsymbol{\nu}|\right) \tag{A.78a}$$

$$= \int \mathrm{d}^2\boldsymbol{\mu}\,\mathrm{d}^2\boldsymbol{\nu}\,\langle\boldsymbol{\mu}|\, A\,|\boldsymbol{\nu}\rangle\,|\boldsymbol{\mu}\rangle\langle\boldsymbol{\nu}| \tag{A.78b}$$

$$= \int \mathrm{d}^2\boldsymbol{\mu}\,\mathrm{d}^2\boldsymbol{\nu}\,\langle\boldsymbol{\mu}|\, A\,|\boldsymbol{\nu}\rangle \int \mathrm{d}^2\boldsymbol{\eta}\,\big\langle\boldsymbol{\nu}\big|\, D^\dagger(\boldsymbol{\eta})\,\big|\boldsymbol{\mu}\big\rangle\, D(\boldsymbol{\eta}) \tag{A.78c}$$

using Lemma A.16 to expand the coherent state diadic,

$$= \int \mathrm{d}^2\boldsymbol{\eta}\mathrm{d}^2\boldsymbol{\mu}\ \langle\boldsymbol{\mu}|\, A\left(\int \mathrm{d}^2\boldsymbol{\nu}\,|\boldsymbol{\nu}\rangle\langle\boldsymbol{\nu}|\right) D^\dagger(\boldsymbol{\eta})\,|\boldsymbol{\mu}\rangle\, D(\boldsymbol{\eta}) \tag{A.78d}$$

$$= \int \mathrm{d}^2\boldsymbol{\eta}\mathrm{d}^2\boldsymbol{\mu}\ \langle\boldsymbol{\mu}|\, AD^\dagger(\boldsymbol{\eta})\,|\boldsymbol{\mu}\rangle\, D(\boldsymbol{\eta}). \tag{A.78e}$$

□

A.4 Fermionic Phase Space

Theorem A.17, in particular Eq. (A.77), shows that there is a one-to-one correspondence between operators A on the fermionic Fock space and Grassmann functions $\int \mathrm{d}^2\boldsymbol{\mu}\,\langle\boldsymbol{\mu}|\, AD(\boldsymbol{\eta})\,|\boldsymbol{\mu}\rangle$. Therefore, these functions provide faithful representations of operators. In particular, they can be used to represent fermionic density operators ρ, leading to "phase-space" representations of fermionic states analogous to the complex functions that provide phase-space representations of bosonic states.

To establish our first phase-space representation, we will need a relation between the Fourier transform of a coherent state diadic and the displacement operators.

Lemma A.18 (Fourier Transform of Coherent State Diadic)
The following operator identity holds for the Fourier transform of the coherent state diadic $|\mu\rangle\langle-\mu|$:

$$\int \mathrm{d}^2\boldsymbol{\mu} e^{-\sum_i \bar{\mu}_i\eta_i+\mu_i\bar{\eta}_i} |\mu\rangle\langle-\mu| = e^{-\frac{1}{2}\sum_i \bar{\eta}_i\eta_i} \int \mathrm{d}^2\boldsymbol{\mu} D^\dagger(\boldsymbol{\eta}) |\mu\rangle\langle-\mu| . \tag{A.79}$$

Proof Introducing two additional coherent states $|\boldsymbol{\delta}\rangle$ and $|\boldsymbol{\gamma}\rangle$, we can calculate the matrix element of the Fourier transformed diadic,

$$\langle\boldsymbol{\delta}| \int \mathrm{d}^2\boldsymbol{\mu}\, e^{-\sum_i \bar{\mu}_i\eta_i+\mu_i\bar{\eta}_i} |\mu\rangle \langle-\mu|\gamma\rangle \tag{A.80a}$$

$$= \int \mathrm{d}^2\boldsymbol{\mu}\, e^{-\sum_i \bar{\mu}_i\eta_i+\mu_i\bar{\eta}_i} e^{\sum_i \bar{\delta}_i\mu_i-\frac{1}{2}\bar{\delta}_i\delta_i-\frac{1}{2}\bar{\mu}_i\mu_i} e^{\sum_i -\bar{\mu}_i\gamma_i-\frac{1}{2}\bar{\mu}_i\mu_i-\frac{1}{2}\bar{\gamma}_i\gamma_i} \tag{A.80b}$$

using Result A.13 for the overlap of two coherent states,

$$= e^{-\frac{1}{2}\sum_i \bar{\delta}_i\delta_i+\bar{\gamma}_i\gamma_i} \int \mathrm{d}^2\boldsymbol{\mu}\, e^{-\sum_i \bar{\mu}_i\mu_i-(\bar{\eta}_i+\bar{\delta}_i)\mu_i+\bar{\mu}_i(\eta_i+\gamma_i)} \tag{A.80c}$$

$$= e^{-\frac{1}{2}\sum_i \bar{\delta}_i\delta_i+\bar{\gamma}_i\gamma_i+\bar{\eta}_i\eta_i} e^{-\sum_i \bar{\delta}_i\eta_i} \int \mathrm{d}^2\boldsymbol{\mu'}\, e^{-\sum_i \bar{\mu'}_i\mu'_i-(\bar{\eta}_i+\bar{\delta}_i)\mu'_i+\bar{\mu'}_i(\eta_i+\gamma_i)} \tag{A.80d}$$

by "shifting" the integration variable $\mu_i = \mu'_i + \eta_i$ (but *without* shifting the conjugate variable $\bar{\mu}_i = \bar{\mu'}_i$),

$$= e^{-\sum_i \bar{\eta}_i\eta_i} \int \mathrm{d}^2\boldsymbol{\mu'}\, e^{\sum_i \eta_i\bar{\delta}_i} e^{\sum_i \bar{\eta}_i\mu_i} \langle\boldsymbol{\delta}|\boldsymbol{\mu}\rangle \langle-\boldsymbol{\mu}|\boldsymbol{\gamma}\rangle \tag{A.80e}$$

$$= e^{-\frac{1}{2}\sum_i \bar{\eta}_i\eta_i} \int \mathrm{d}^2\boldsymbol{\mu'}\, e^{-\frac{1}{2}\sum_i \bar{\eta}_i\eta_i} \left\langle\boldsymbol{\delta}\right| e^{\sum_i \eta_i c_i^\dagger} e^{\sum_i \bar{\eta}_i c_i} \left|\boldsymbol{\mu}\right\rangle \langle-\boldsymbol{\mu}|\boldsymbol{\gamma}\rangle \tag{A.80f}$$

$$= e^{-\frac{1}{2}\sum_i \bar{\eta}_i\eta_i} \int \mathrm{d}^2\boldsymbol{\mu'}\, \left\langle\boldsymbol{\delta}\right| D^\dagger(\boldsymbol{\eta}) \left|\boldsymbol{\mu}\right\rangle \langle-\boldsymbol{\mu}|\boldsymbol{\gamma}\rangle , \tag{A.80g}$$

using the decomposition of the displacement operator from Eq. (A.77) in the final equality.

Since this holds for any $|\boldsymbol{\delta}\rangle$ and $|\boldsymbol{\gamma}\rangle$ and the coherent states form a complete basis (Theorem A.15), Eq. (A.79) is true for all matrix elements of the operators on each side of the equality, therefore holds for the operators themselves. □

We are now ready to derive the characteristic function phase-space representation of fermionic states. We first define the characteristic function, then show that it does indeed provide a faithful representation of any fermionic state.

Definition A.19 (Characteristic Function)
The fermionic characteristic function $\chi(\boldsymbol{\eta})$ for a state ρ of a collection of fermionic modes is defined by

$$\chi(\boldsymbol{\eta}) = \mathrm{Tr}[\rho D(\boldsymbol{\eta})] \tag{A.81}$$

where $\boldsymbol{\eta}$ is a vector of Grassmann variables η_i and their conjugates $\bar{\eta}_i$, and the displacement operator $D(\boldsymbol{\eta})$ is defined by Definition A.11.

If the displacement operator is written in the "real" representation $D(\boldsymbol{\alpha})$ of Eq. (A.59), where $\boldsymbol{\alpha}$ is a vector of "real" Grassmann variables, we obtain a "real" representation of the characteristic function, which is nothing more than a trivial change of variables:

$$\chi(\boldsymbol{\alpha}) = \mathrm{Tr}[\rho D(\alpha)] \tag{A.82}$$

Theorem A.20 (Characteristic Function Decomposition)
The fermionic characteristic function provides a faithful representation of fermionic density operators, and the expansion of any state ρ in terms of the characteristic function is given by

$$\rho = \int \mathrm{d}^2\boldsymbol{\eta}\, \chi(\boldsymbol{\eta}) E(\boldsymbol{\eta}) e^{\frac{1}{2}\sum_i \bar{\eta}_i \eta_i} \tag{A.83}$$

where the operator $E(\boldsymbol{\eta})$ is the Fourier transform of a coherent state diadic,

$$E(\boldsymbol{\eta}) = \int \mathrm{d}^2\boldsymbol{\mu}\, e^{-\sum_i \bar{\mu}_i \eta_i + \mu_i \bar{\eta}_i} \left|\boldsymbol{\mu}\right\rangle\!\left\langle -\boldsymbol{\mu}\right| . \tag{A.84}$$

Proof From Theorem A.17, we know that any operator A can be expanded in terms of displacement operators. Setting $A = \rho$ in Eq. (A.77) we have

$$\rho = \int \mathrm{d}^2\boldsymbol{\eta}\, \mathrm{d}^2\boldsymbol{\mu}\, \left\langle \boldsymbol{\mu}\right| \rho D^\dagger(\boldsymbol{\eta}) \left|\boldsymbol{\mu}\right\rangle D(\boldsymbol{\eta}) \tag{A.85a}$$

$$= \int \mathrm{d}^2\boldsymbol{\eta}\, \mathrm{d}^2\boldsymbol{\mu}\, \mathrm{d}^2\boldsymbol{\gamma}\, \left\langle \boldsymbol{\mu}\right| \rho \left|\boldsymbol{\gamma}\right\rangle \left\langle \boldsymbol{\gamma}\right| D^\dagger(\boldsymbol{\eta}) \left|\boldsymbol{\mu}\right\rangle D(\boldsymbol{\eta}) \tag{A.85b}$$

inserting a coherent state resolution of identity (Eq. (A.72)),

$$= \int \mathrm{d}^2\boldsymbol{\eta}\, \mathrm{d}^2\boldsymbol{\mu}\, \mathrm{d}^2\boldsymbol{\gamma}\, \mathrm{Tr}\Big[\rho \left|-\boldsymbol{\gamma}\right\rangle\!\left\langle \boldsymbol{\mu}\right|\Big] \left\langle \boldsymbol{\gamma}\right| D^\dagger(\boldsymbol{\eta}) \left|\boldsymbol{\mu}\right\rangle D(\boldsymbol{\eta}) \tag{A.85c}$$

$$= \int \mathrm{d}^2\boldsymbol{\xi}\, \mathrm{Tr}\big[\rho D(\boldsymbol{\xi})\big] \int \mathrm{d}^2\boldsymbol{\gamma}\, \mathrm{d}^2\boldsymbol{\eta}\, \mathrm{d}^2\boldsymbol{\mu}\, \left\langle \boldsymbol{\mu}\right| D^\dagger(\boldsymbol{\xi}) \left|-\boldsymbol{\gamma}\right\rangle \left\langle \boldsymbol{\gamma}\right| D^\dagger(\boldsymbol{\eta}) \left|\boldsymbol{\mu}\right\rangle D(\boldsymbol{\eta}) \tag{A.85d}$$

using Lemma A.16 to expand the coherent state diadic,

$$= \int \mathrm{d}^2\boldsymbol{\xi}\ \mathrm{Tr}\big[\rho D(\boldsymbol{\xi})\big] \int \mathrm{d}^2\boldsymbol{\gamma}\, D^\dagger(\xi)\, |-\boldsymbol{\gamma}\rangle\langle\boldsymbol{\gamma}| \tag{A.85e}$$

noticing that the second integral is just an expansion in terms of displacement operators (Theorem A.17),

$$= \int \mathrm{d}^2\boldsymbol{\xi}\, \chi(\boldsymbol{\xi}) \int \mathrm{d}^2\boldsymbol{\gamma}\, e^{\frac{1}{2}\sum_i \bar{\xi}_i\xi_i} e^{-\sum_i \bar{\gamma}_i\xi_i + \gamma_i\bar{\xi}_i}\, |\boldsymbol{\gamma}\rangle\langle-\boldsymbol{\gamma}|\,. \tag{A.85f}$$

using Lemma A.18 for the Fourier transform of the coherent state diadic,

$$= \int \mathrm{d}^2\boldsymbol{\xi}\, \chi(\boldsymbol{\xi}) e^{\frac{1}{2}\sum_i \bar{\eta}_i\eta_i} E(\boldsymbol{\xi}). \tag{A.85g}$$

□

Successive derivatives of the characteristic function about 0 give successively higher-order correlation functions. For example, working in the "real" representation (recall that the r_i are the Majorana operators $r_{2i-1} = x_i$ and $r_{2i} = p_i$, as defined in Eq. (A.54)),

$$\begin{aligned} \partial_{\alpha_i}\chi(\boldsymbol{\alpha}) &= \partial_{\alpha_i} \mathrm{Tr}\big[\rho D(\boldsymbol{\alpha})\big] && \text{(A.86a)}\\ &= \mathrm{Tr}\left[\rho\, \partial_{\alpha_i}(1-\alpha_i r_i) e^{-\sum_{j\neq i}\alpha_j r_j}\right] && \text{(A.86b)}\\ &= -\mathrm{Tr}\big[\rho\, r_i e^{-\sum_{j\neq i}\alpha_j r_j}\big] && \text{(A.86c)} \end{aligned}$$

Similarly, taking second derivatives,

$$\partial_{\alpha_i}\partial_{\alpha_j}\chi(\boldsymbol{\alpha}) = -\partial_{\alpha_i}\mathrm{Tr}\big[\rho\, r_j\, e^{-\sum_{k\neq j}\alpha_k r_k}\big] = \begin{cases} \mathrm{Tr}\big[\rho\, r_i r_j\, e^{-\sum_{k\neq i,j}\alpha_k r_k}\big] & i\neq j \\ 0 & i=j, \end{cases} \tag{A.87}$$

and so on for higher derivatives. From these, we have

$$\partial_{\alpha_i}\chi(\boldsymbol{\alpha})\Big|_{\boldsymbol{\alpha}=0} = -\mathrm{Tr}[\rho\, r_i] = \langle r_i\rangle\,, \tag{A.88}$$

$$\partial_{\alpha_i}\partial_{\alpha_j}\chi(\boldsymbol{\alpha})\Big|_{\boldsymbol{\alpha}=0} = \left\{ \begin{array}{ll} \mathrm{Tr}\big[\rho\, r_i r_j\big] & i\neq j \\ 0 & i=j \end{array} \right\} = \frac{1}{2}\,\langle [r_i, r_j]\rangle\,, \tag{A.89}$$

and so on. Thus fermionic correlation functions can be obtained by taking derivatives of the characteristic function $\chi(\boldsymbol{\alpha})$ at $\boldsymbol{\alpha} = 0$.

It is possible to define a number of other fermionic phase-space representations (for an exhaustive analysis, see Cahill and Glauber [1999]), but

a particularly useful one is the fermionic P–representation, as it gives a decomposition of fermionic states in terms of coherent states. In analogy with the bosonic P–representation, it is defined as Grassmann Fourier transform of the characteristic function.

Definition A.21 (P–Representation)
The P–representation for a fermionic state with characteristic function $\chi(\boldsymbol{\xi})$ (in the "complex" representation) is given by

$$P(\boldsymbol{\xi}) = \int \mathrm{d}^2\boldsymbol{\eta} \exp\left(-\sum_i \bar{\eta}_i\xi_i + \eta_i\bar{\xi}_i\right) \exp\left(\frac{1}{2}\sum_i \bar{\eta}_i\eta_i\right) \chi(\boldsymbol{\eta}), \tag{A.90}$$

where $\boldsymbol{\eta}$ and $\boldsymbol{\xi}$ are both vectors of Grassmann variables along with their conjugates.

That the P–representation is indeed a faithful representation of fermionic states is demonstrated by the following theorem.

Theorem A.22 (P–Representation Decomposition)
Any state ρ of a set of fermionic modes can be decomposed in terms of the P–representation,

$$\rho = \int \mathrm{d}^2\boldsymbol{\xi}\, P(\boldsymbol{\xi})\, |\boldsymbol{\xi}\rangle\langle -\boldsymbol{\xi}|\,. \tag{A.91}$$

Proof From Eq. (A.83) from Theorem A.20 for the decomposition of any operator A in terms of the characteristic function $\chi(\boldsymbol{\eta})$, we have

$$\rho = \int \mathrm{d}^2\boldsymbol{\xi}\, \left(\chi(\boldsymbol{\xi})\, e^{\frac{1}{2}\sum_i \bar{\xi}_i\xi_i}\right) E(\boldsymbol{\xi}) \tag{A.92a}$$

$$= \int \mathrm{d}^2\boldsymbol{\xi} \int \mathrm{d}^2\boldsymbol{\eta} \exp\left(-\sum_i \bar{\eta}_i\xi_i + \eta_i\bar{\xi}_i\right) \exp\left(\frac{1}{2}\sum_i \bar{\eta}_i\eta_i\right) \chi(\boldsymbol{\eta}) \;\times \int \mathrm{d}^2\boldsymbol{\zeta}\, \exp\left(-\sum_i \bar{\zeta}_i\xi_i + \zeta_i\bar{\xi}_i\right) E(-\boldsymbol{\zeta}) \tag{A.92b}$$

applying the Grassmann Parseval theorem of Lemma A.10,

$$= \int \mathrm{d}^2\boldsymbol{\xi}\, P(\boldsymbol{\xi}) \int \mathrm{d}^2\boldsymbol{\zeta}\, \exp\left(\sum_i \bar{\zeta}_i\xi_i + \zeta_i\bar{\xi}_i\right) E(\boldsymbol{\zeta}) \tag{A.92c}$$

$$= \int \mathrm{d}^2\boldsymbol{\xi}\, P(\boldsymbol{\xi})\, |\boldsymbol{\xi}\rangle\langle -\boldsymbol{\xi}|\,, \tag{A.92d}$$

noticing that the integral is just the inverse Fourier transform of $E(\boldsymbol{\zeta})$, which was defined in Eq. (A.84) to be the Fourier transform of the diadic $|\boldsymbol{\xi}\rangle\langle -\boldsymbol{\xi}|$.□

A.5 Fermionic Gaussian States

We are finally in a position to define fermionic Gaussian states based on their phase-space representation: they are simply states whose phase-space representations are Gaussian functions of Grassmann variables. In this section, we will emphasize the "real" representation, as it is mathematically slightly more elegant.

A.5.1 Fermionic Gaussian State Definition

Definition A.23 (Fermionic Gaussian State)
A fermionic Gaussian state is one whose characteristic function in the "real" representation is a Gaussian Grassmann function of the form

$$\chi(\boldsymbol{\alpha}) = \exp\left(-\tfrac{1}{2}\boldsymbol{\alpha}^T\Gamma\boldsymbol{\alpha} + 2\boldsymbol{\nu}^T\boldsymbol{\alpha}\right), \tag{A.93}$$

where $\boldsymbol{\alpha}$ is a vector of "real" Grassmann variables, $\boldsymbol{\nu}$ is a vector of "real" Grassmann numbers, and Γ is a real, anti-symmetric matrix.

Alternatively, the definition can be written in the "complex" representation:

$$\chi(\boldsymbol{\eta}) = \exp\left(-\tfrac{1}{2}\boldsymbol{\eta}^\dagger M\boldsymbol{\eta} + \boldsymbol{\mu}^\dagger\boldsymbol{\eta}\right), \tag{A.94}$$

where $\boldsymbol{\eta}$ is a vector of Grassmann variables and their conjugates, and the vector $\boldsymbol{\mu}$ and matrix M are related to the vector $\boldsymbol{\nu}$ and the real, anti-symmetric matrix Γ appearing in Definition A.23 by

$$\boldsymbol{\mu} = \Omega^{-1}\boldsymbol{\nu}, \quad M = \Omega^{-1}\Gamma\Omega, \tag{A.95a}$$

$$\Omega = \bigoplus \frac{1}{2}\begin{pmatrix} 1 & 1 \\ i & -i \end{pmatrix}. \tag{A.95b}$$

Taking derivatives of $\chi(\boldsymbol{\alpha})$ with respect to the variables α_i about $\boldsymbol{\alpha} = 0$, as we did in Section A.4, gives

$$\partial_{\alpha_i}\chi(\boldsymbol{\alpha})\Big|_{\boldsymbol{\alpha}=0} = -2\nu_i \tag{A.96}$$

$$\partial_{\alpha_i}\partial_{\alpha_j}\chi(\boldsymbol{\alpha})\Big|_{\boldsymbol{\alpha}=0} = \begin{cases} \frac{1}{2}(\Gamma_{ij} - \Gamma_{ji}) + \nu_i\nu_j & i \neq j \\ 0 & i = j. \end{cases} \tag{A.97}$$

Comparing with the corresponding Eq. (A.88) and Eq. (A.89) from Section A.4, the $i = j$ cases are consistent, and we must have

$$-2\nu_i = -\langle r_i \rangle \tag{A.98a}$$

$$\frac{1}{2}(\Gamma_{ij} - \Gamma_{ji}) + 4\nu_i\nu_j = \frac{1}{2}\langle [r_i, r_j] \rangle \quad i \neq j. \tag{A.98b}$$

There is some freedom in the choice of Γ, but we can choose the components so that the matrix is anti-symmetric. As its components are commutators of Majorana operators, the matrix Γ appearing in Definition A.23 is in fact the covariance matrix, or matrix of two-point correlation functions. We summarize this in the following result.

Result A.24 (Gaussian Characteristic Function)
The displacement coefficients and covariance matrix in the characteristic function representation of a fermionic Gaussian state (Definition A.23) can be chosen to be

$$\nu_i = \frac{1}{2}\langle r_i \rangle , \tag{A.99a}$$

$$\Gamma_{ij} = \frac{1}{2}\langle [r_i, r_j] \rangle - \langle r_i \rangle \langle r_j \rangle . \tag{A.99b}$$

A.5.2 Examples of Fermionic Gaussian States

The simplest examples of fermionic Gaussian states are the coherent states themselves.

Result A.25 (Coherent States as Gaussian States)
A fermionic coherent state $|\boldsymbol{\mu}\rangle\langle\boldsymbol{\mu}|$ is a Gaussian state, with "real" characteristic function

$$\chi(\alpha) = e^{-\frac{1}{2}\boldsymbol{\alpha}^T\Gamma\boldsymbol{\alpha}+2\boldsymbol{\nu}^T\boldsymbol{\alpha}}, \tag{A.100a}$$

$$\nu_{2i-1} = \frac{\bar{\mu}_i + \mu_i}{2}, \quad \nu_{2i} = \frac{\bar{\eta}_i - \eta_i}{2i}, \quad \Gamma = \bigoplus \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} . \tag{A.100b}$$

Proof From Definition A.19 of the characteristic function in the "complex" representation, we have for a coherent state $\rho = |\boldsymbol{\mu}\rangle\langle\boldsymbol{\mu}|$

$$\chi(\boldsymbol{\eta}) = \mathrm{Tr}[|\boldsymbol{\mu}\rangle\langle\boldsymbol{\mu}| \, D(\boldsymbol{\eta})] \tag{A.101a}$$

$$= \int \mathrm{d}^2\boldsymbol{\xi} \, \langle\boldsymbol{\xi}|\boldsymbol{\mu}\rangle \, \langle\boldsymbol{\mu}| \, D(\boldsymbol{\eta}) \, |-\boldsymbol{\mu}\rangle \tag{A.101b}$$

$$= \int \mathrm{d}^2\boldsymbol{\xi} e^{-\sum_i \bar{\xi}_i\mu_i + \frac{1}{2}\bar{\xi}_i\xi_i + \frac{1}{2}\bar{\mu}_i\mu_i} e^{-\frac{1}{2}\sum_i \bar{\nu}_i\xi_i + \nu_i\bar{\xi}_i} \, \langle\boldsymbol{\mu}|\boldsymbol{\nu} + \boldsymbol{\xi}\rangle \tag{A.101c}$$

using Result A.13 for the coherent state overlap, and expanding $|\boldsymbol{\xi}\rangle = D(\boldsymbol{\xi})\,|0\rangle$ before using Result A.14 for the product of two displacement operators,

$$= e^{\sum_i \bar{\mu}_i\eta_i - \bar{\mu}_i\mu_i - \frac{1}{2}\eta\bar{u}_i\nu_i} \int \mathrm{d}^2\boldsymbol{\xi} e^{-\sum_i \bar{\xi}_i\xi_i + (\bar{\eta}_i\bar{\mu}_i)\xi_i - \bar{\xi}_i\mu_i} \tag{A.101d}$$

$$= e^{\sum_i \bar{\mu}_i\eta_i - \bar{\mu}_i\mu_i} e^{-\sum_i (\bar{\eta}_i - \bar{\mu}_i)\mu_i} \tag{A.101e}$$

using Result A.7 to carry out the Gaussian integral,

$$= \exp\left(-\tfrac{1}{2}\boldsymbol{\eta}^\dagger M \boldsymbol{\eta} + \boldsymbol{\mu}^\dagger \boldsymbol{\eta}\right), \quad M = \bigoplus \frac{1}{2}\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \tag{A.101f}$$

Thus the characteristic function indeed has the Gaussian form of Eq. (A.94). Transforming to the "real" representation of Eq. (A.93), we have

$$\boldsymbol{\nu} = \Omega \boldsymbol{\mu}, \quad \nu_{2i-1} = \frac{\bar{\mu}_i + \mu_i}{2}, \quad \nu_{2i} = \frac{\bar{\mu}_i - \mu_i}{2i}, \tag{A.102}$$

$$\Gamma = (\Omega^\dagger)^{-1} M\, \Omega^{-1} = \bigoplus \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}. \tag{A.103}$$

□

As noted in Section A.1, fermionic coherent states can not be physical states, since the expectation values of the observables r_i are the Grassmann numbers ν_i. Indeed, no Gaussian state with non-zero displacement $\boldsymbol{\nu}$ in Eq. (A.93) can be a physical state, for the same reason.

The fermionic vacuum can be written as a coherent state $|\boldsymbol{\xi}\rangle\langle\boldsymbol{\xi}|$ with $\boldsymbol{\xi} = 0$, so it is also a Gaussian state.

Result A.26 (Fermionic Vacuum as Gaussian State)
The fermionic vacuum is a Gaussian state with "real" characteristic function

$$\chi(\boldsymbol{\alpha}) = e^{-\frac{1}{2}\boldsymbol{\alpha}^T \Gamma_{\text{vac}} \boldsymbol{\alpha}}, \quad \Gamma_{\text{vac}} = \bigoplus \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}. \tag{A.104}$$

Naturally, the vacuum state is a valid, physical state of a set of fermionic modes; it is the only coherent state with all displacement coefficients $\boldsymbol{\nu} = 0$.

A.5.3 P–Representation of a Gaussian State

We can obtain the form of the P–representation for Gaussian states by transforming the characteristic function Eq. (A.93) according to Eq. (A.90) from Definition A.21 of the P–representation.

Theorem A.27 (Gaussian State P–Representation)
The "real" P–representation of a Gaussian state with covariance matrix Γ and displacement coefficients $\boldsymbol{\nu}$ is

$$P(\boldsymbol{\alpha}) = (2i)^N \operatorname{Pf}(\Gamma - \Gamma_{\text{vac}})\, e^{-2(\boldsymbol{\alpha}+\boldsymbol{\nu})(\Gamma - \Gamma_{\text{vac}})^{-1}(\boldsymbol{\alpha}+\boldsymbol{\nu})}, \tag{A.105}$$

where $\boldsymbol{\alpha}$ is a vector of "real" Grassmann numbers, and N is the number of fermionic modes.

Proof We can change variables in the integral of Eq. (A.90) defining the P–representation, from Grassmann variables $\boldsymbol{\eta}$ to the "real" Grassmann variables $\boldsymbol{\beta} = \Omega\boldsymbol{\eta}$. The Jacobian is $|J| = \det\Omega^{-1} = (2i)^N$, where N is the number of fermionic modes. This gives the following expression for the "real" P–representation:

$$P(\boldsymbol{\alpha}) = (2i)^N \int \mathrm{d}^2\boldsymbol{\beta} \exp\left(\sum_i \alpha_i\beta_i\right) \exp\left(-i\sum_i \beta_{2i-1}\beta_{2i}\right) \chi(\boldsymbol{\beta}), \quad \text{(A.106)}$$

(where $\boldsymbol{\alpha}$ is related to $\boldsymbol{\xi}$ from Eq. (A.90) by $\boldsymbol{\alpha} = \Omega\,\boldsymbol{\xi}$).

Writing all the sums in vector notation and substituting the Gaussian characteristic function $\chi(\boldsymbol{\beta})$ of Eq. (A.93), we obtain

$$\begin{aligned} P(\boldsymbol{\alpha}) &= (2i)^N \int \mathrm{d}\boldsymbol{\beta} \exp\left(2\boldsymbol{\alpha}^T\boldsymbol{\beta}\right) \times \\ &\qquad \exp\left(\frac{1}{2}\,\boldsymbol{\beta}^T\Gamma_{\text{vac}}\boldsymbol{\beta}\right) \exp\left(-\frac{1}{2}\boldsymbol{\beta}^T\Gamma\boldsymbol{\beta} + 2\boldsymbol{\nu}^T\boldsymbol{\beta}\right) && \text{(A.107a)} \\ &= (2i)^N \int \mathrm{d}\boldsymbol{\beta} \exp\left(-\frac{1}{2}\,\boldsymbol{\beta}^2(\Gamma - \Gamma_{\text{vac}})\boldsymbol{\beta} + 2(\boldsymbol{\alpha}+\boldsymbol{\nu})^T\boldsymbol{\beta}\right) && \text{(A.107b)} \\ &= (2i)^N \operatorname{Pf}(\Gamma - \Gamma_{\text{vac}}) \exp\left(-2(\boldsymbol{\alpha}+\boldsymbol{\nu})(\Gamma - \Gamma_{\text{vac}})^{-1}(\boldsymbol{\alpha}+\boldsymbol{\nu})\right), && \text{(A.107c)} \end{aligned}$$

carrying out the Grassmann Gaussian integral using Result A.5. □

If we transform to Grassmann variables η_i, μ_i and their conjugates using $\boldsymbol{\eta} = \Omega^{-1}\boldsymbol{\alpha}$ and $\boldsymbol{\mu} = \Omega^{-1}\boldsymbol{\nu}$, we obtain the "complex" P–representation:

$$P(\boldsymbol{\eta}) = (2i)^N \operatorname{Pf}(\Gamma - \Gamma_{\text{vac}})\, e^{-2(\boldsymbol{\eta}^\dagger + \boldsymbol{\mu}^\dagger)\Omega^\dagger(\Gamma - \Gamma_{\text{vac}})^{-1}\Omega(\boldsymbol{\eta}+\boldsymbol{\mu})}. \quad \text{(A.108)}$$

Since the vector $\boldsymbol{\eta}$ contains both Grassmann variables *and* their conjugates, as does μ, we can permute the entries of M so that the exponent can be rewritten as $(\boldsymbol{\eta}^T + \boldsymbol{\mu}^T)M'(\boldsymbol{\eta} + \boldsymbol{\mu})$. Noting that the diagonal entries of M' are irrelevant since the square of any Grassmann is zero, and re-balancing the off-diagonal entries that give the coefficients of the $\bar{\eta}_i\eta_i$ and $\eta_i\bar{\eta}_i$ terms, we can rewrite this in the form

$$P(\boldsymbol{\eta}) = (2i)^N \operatorname{Pf}(\Gamma - \Gamma_{\text{vac}}) \exp\left(-\frac{1}{2}(\boldsymbol{\eta}^T + \boldsymbol{\mu}^T)A(\boldsymbol{\eta} + \boldsymbol{\mu}),\right) \quad \text{(A.109)}$$

where A is some anti-symmetric matrix. (The factor of $1/2$ is included for convenience.) This form will be useful later.

A.5.4 Wick's Theorem

Wick's theorem can be stated in a number of ways, for example as an operator identity, or in terms of time-ordered products in field theories. Writing fermionic states in their phase-space representation, it is clear that the fermionic Gaussian states are precisely those fermionic states for which Wick's theorem is valid, and the theorem becomes an application of Grassmann Gaussian integration. For simplicity, we state and prove Wick's theorem only for *physical* Gaussian states, that is those whose displacement coefficients are all zero.

Theorem A.28 (Wick's Theorem)
The expectation value of a product of fermionic operators with respect to a physical fermionic Gaussian state can be expressed in terms of expectation values of pairs of operators:

$$\left\langle \prod_{i=1}^{n} a_i \right\rangle = \begin{cases} 0 & n \text{ odd} \\ \left(\frac{1}{2}\right)^{n/2} \sum_{\pi} \operatorname{sgn}(\pi) \prod_{i}^{n/2} \left\langle a_{\pi(2i-1)} a_{\pi(2i)} \right\rangle & n \text{ even}, \end{cases} \tag{A.110}$$

where the sum is over all permutations π of the n indices i, and each a_i is either a creation or an annihilation operator $c_{s_i}^\dagger$ or c_{s_i} acting on the s_i^{th} mode.

Proof Define a product of fermionic operators to be normal-ordered if all creation operators stand to the left of all annihilation operators. If the operators in $\prod_i a_i$ are *not* normal-ordered, we can move all operators that act on different modes past each other until they are. Each interchange of two of the operators gives rise to a sign factor from the anti-commutation relations, but this is already accounted for on the right hand side of Eq. (A.110), since swapping the order of two operators changes the signum of the permutation. Therefore, if the theorem holds for normal-ordered products of operators that each act on a different mode, then it holds for any product of the same operators.

The only non-zero product of two operators on the *same* mode that is not normal-ordered is $c_i c_i^\dagger$. If the product $\prod_i a_i$ contains c_i and $c_i^\dagger$ in that order, we can move them to the beginning of the product by the preceding argument, so it is sufficient to consider products of the form $c_i c_i^\dagger \prod_{i \geq 3} a_i$. Divide the permutations π into two sets, π_0 being those that leave $c_i c_i^\dagger$ at the beginning of the product, and π_1 being those that do not. Then, assuming the theorem

is true for normal-ordered products and for $\langle \prod_{i\geq 3} a_i \rangle$,

$$\left\langle c_i c_i^\dagger \prod_{i=1}^{n} a_i \right\rangle = \left\langle (1 - c_i^\dagger c_i) \prod_{i=3}^{n} a_i \right\rangle \tag{A.111a}$$

$$= \left\langle \prod_{i=3}^{n} a_i \right\rangle - \left\langle c_i^\dagger c_i \prod_{i=3}^{n} a_i \right\rangle \tag{A.111b}$$

$$= \sum_{\pi_0} \operatorname{sgn}(\pi_0) \prod_{i\geq 3} \langle a_{\pi_0(2i-1)} a_{\pi_0(2i)} \rangle - \langle c_i^\dagger c_i \rangle \sum_{\pi_0} \operatorname{sgn}(\pi_0) \prod_{i=3}^{n} \langle a_{\pi_0(2i-1)} a_{\pi_0(2i)} \rangle$$
$$+ \sum_{\pi_1} \operatorname{sgn}(\pi_1) \prod_{i=1}^{n} \langle a_{\pi_1(2i-1)} a_{\pi_1(2i)} \rangle \tag{A.111c}$$

where the extra sign factor in the third term of comes from interchanging the order of c_i and $c_i^\dagger$, which changes the signum of the permutation π_1 in each term of the sum,

$$= \langle c_i c_i^\dagger \rangle \sum_{\pi_0} \operatorname{sgn}(\pi_0) \prod_{i=3} \langle a_{\pi_0(2i-1)} a_{\pi_0(2i)} \rangle + \sum_{\pi_1} \operatorname{sgn}(\pi_1) \prod_{i=3}^{n} \langle a_{\pi_1(2i-1)} a_{\pi_1(2i)} \rangle \tag{A.111d}$$

$$= \sum_{\pi} \operatorname{sgn}(\pi) \prod_{i=1}^{n} \langle a_{\pi(2i-1)} a_{\pi(2i)} \rangle \,. \tag{A.111e}$$

If the operators in the product $\prod_{i\geq 3} a_i$ are not normally-ordered, we can apply this result recursively until the entire operator product $\prod_i a_i$ is normal-ordered. Therefore, if the theorem is true for normal-ordered products, it is true for any product of operators.

Writing the state in its "complex" P–representation and assuming the operator product is normal-ordered so that $\prod_{i=1}^{n} a_i = \prod_j c_{s_j}^\dagger \prod_k c_{s_k}$, we can re-express the expectation value as a Grassmann Gaussian integral:

$$\left\langle \prod_i c_{s_i}^\dagger \prod_j c_{s_j} \right\rangle = \operatorname{Tr}\Big[\prod_i c_{s_i}^\dagger \prod_j c_{s_j} \int \mathrm{d}^2\boldsymbol{\eta} P(\boldsymbol{\eta}) \, |\boldsymbol{\eta}\rangle\langle -\boldsymbol{\eta}| \Big] \tag{A.112a}$$

$$= \int \mathrm{d}^2\boldsymbol{\eta} P(\boldsymbol{\eta}) \, \langle \boldsymbol{\eta}| \prod_i c_{s_i}^\dagger \prod_j c_{s_j} \, |\boldsymbol{\eta}\rangle \tag{A.112b}$$

$$= \int \mathrm{d}^2\boldsymbol{\eta} P(\boldsymbol{\eta}) \prod_i \bar{\eta}_{s_i} \prod_j \eta_{s_j}. \tag{A.112c}$$

Now, from Eq. (A.109), the "complex" P–representation of a physical Gaussian state can be written in the form

$$P(\boldsymbol{\xi}) = K \exp\left(-\tfrac{1}{2}\boldsymbol{\xi}^T A^{-1} \boldsymbol{\xi}\right). \tag{A.113}$$

We have absorbed all normalization factors into K, A is an anti-symmetric matrix related to the state's covariance matrix, the displacement coefficients must be zero if the state is to be physical (see Section A.5.2), and $\xi_{2i-1} = \eta_i$, $\xi_{2i} = \bar{\eta}_i$ is a convenient relabelling of $\boldsymbol{\eta}$. The expectation value of a normal-ordered product of two creation or annihilation operators is then given by

$$\langle a_i a_j \rangle = K \int \mathrm{d}\boldsymbol{\xi} \exp\left(-\frac{1}{2}\boldsymbol{\xi}^T A^{-1} \boldsymbol{\xi}\right) \xi_{s_i} \xi_{s_j} \tag{A.114a}$$

$$= K \operatorname{Pf}(A^{-1}) \frac{1}{2}(A_{s_i,s_j} - A_{s_j,s_i}) \tag{A.114b}$$

$$= K \operatorname{Pf}(A^{-1}) A_{s_i,s_j} \tag{A.114c}$$

carrying out the integration using Result A.6. By the above argument, this also holds for any product of two fermionic operators, even if they are not normal-ordered.

Finally, the expectation value of the normal-ordered product $\prod_{i=1}^{n} a_i = \prod_j c_{s_j}^\dagger \prod_k c_{s_k}$ is given by

$$\left\langle \prod_j c_{s_j}^\dagger \prod_k c_{s_k} \right\rangle$$

$$= \int \mathrm{d}\boldsymbol{\xi} \exp\left(-\frac{1}{2}\boldsymbol{\xi}^T A^{-1} \boldsymbol{\xi}\right) \prod_j \xi_{2s_j} \prod_k \xi_{2s_k - 1} \tag{A.115a}$$

$$= \int \mathrm{d}\boldsymbol{\xi} \exp\left(-\frac{1}{2}\boldsymbol{\xi}^T A^{-1} \boldsymbol{\xi}\right) \prod_{i=1}^{n} \xi_{s_i} \tag{A.115b}$$

$$= \begin{cases} 0 & n \text{ odd} \\ K \operatorname{Pf}(A^{-1}) \left(\frac{1}{2}\right)^{n/2} \sum_\pi \operatorname{sgn}(\pi) \prod_i A_{\pi(2s_i-1),\pi(2s_i)} & n \text{ even} \end{cases} \tag{A.115c}$$

$$= \begin{cases} 0 & n \text{ odd} \\ \left(\frac{1}{2}\right)^{n/2} \sum_\pi \operatorname{sgn}(\pi) \prod_i \langle a_{\pi(2s_i-1)} a_{\pi(2s_i)} \rangle & n \text{ even,} \end{cases} \tag{A.115d}$$

again using Result A.6 to carry out the integration, and using Eq. (A.112) in the final equality. This proves Wick's theorem for normal-ordered products. But, by the argument given above, this implies that the theorem is true for any product of operators, thereby concluding the proof. □

Note that since the operators x_i and p_i are linear combinations of creation and annihilation operators, Wick's theorem also holds for expectation values of products of these Majorana operators.

A.5.5 Canonical Transformations and Time Evolution

Orthogonal transformations of the Majorana operators r_i preserve their anti-commutation relations:

$$r'_i = O_{ij}\, r_j, \quad O^T O = O O^T = \mathbb{1} \tag{A.116a}$$

$$\{r'_i, r'_j\} = \{O_{im} r_m, O_{jn} r_n\} = O_{im} O_{jn} \{r_m, r_n\} = O_{im} O^T_{nj} 2\delta_{mn} = 2\delta_{ij}. \tag{A.116b}$$

Since the covariance matrix elements are given by expectation values of commutators of Majorana operators, and the displacement coefficients are given by expectation values of single operators, these *canonical transformations* of the Majorana operators will induce transformations of the covariance matrix and displacement coefficients:

$$\boldsymbol{\nu'} = O\boldsymbol{\nu}, \tag{A.117}$$

$$\Gamma' = O\Gamma O^T, \tag{A.118}$$

but will leave the forms of all phase-space representations unchanged. In particular, Gaussian states will remain Gaussian. (Γ' is still anti-symmetric, since orthogonal transformations preserve anti-symmetry).

One consequence of this is that, since all anti-symmetric matrices can be brought into the block-diagonal form of Eq. (A.30) by orthogonal transformations, all fermionic Gaussian states can be brought by canonical transformations of the Majorana operators into a normal form analogous to the Wilson normal form for bosonic Gaussian states.

Result A.29 (Wilson Normal Form)
The covariance matrix Γ *of any fermionic Gaussian state can be brought into a normal form*

$$\Gamma = \bigoplus_i \begin{pmatrix} 0 & \lambda_i \\ -\lambda_i & 0 \end{pmatrix}. \tag{A.119}$$

A particularly important class of transformations are those induced by evolution under a Hamiltonian. If the Hamiltonian is quadratic in terms of the Majorana operators,

$$H = \sum_{i,j} h_{ij}\, r_i r_j, \tag{A.120}$$

then from the Heisenberg evolution equations, the Majorana operators evolve

according to

$$\frac{\mathrm{d}r_n}{\mathrm{d}t} = i[H, r_n] = i\sum_{i,j} h_{ij}[r_i r_j, r_n] = i\sum_{i,j} h_{ij}(r_i\delta_{nj} - r_j\delta_{ni}) \tag{A.121a}$$

$$= -i\sum_{i}(h_{ni} - h_{in})r_n. \tag{A.121b}$$

Integrating this, we we obtain the transformation

$$r_n(t) = e^{-iAt}r_n, \quad A_{ij} = h_{ij} - h_{ji}. \tag{A.122}$$

Since A is anti-symmetric, $\exp(-iAt)$ is orthogonal. Therefore, if a fermionic system evolves under a quadratic Hamiltonian, the fermionic operators undergo a canonical transformation. Thus Gaussian initial states will remain Gaussian at all times, and the evolution can be completely described by the corresponding transformations of the covariance matrix and displacement coefficients.

Result A.30 (Fermionic Gaussian State Evolution)
The evolution of a Gaussian state under a quadratic Hamiltonian

$$H = \sum_{i,j} h_{ij}\, r_i r_j \tag{A.123}$$

can be completely described by an orthogonal transformation of the initial displacement coefficients $\boldsymbol{\nu}$ *and covariance matrix* Γ*:*

$$\boldsymbol{\nu}(t) = O\,\boldsymbol{\nu}, \tag{A.124a}$$

$$\Gamma(t) = O\Gamma O^T, \tag{A.124b}$$

$$O = e^{-iAt}, \quad A_{ij} = h_{ij} - h_{ji}. \tag{A.124c}$$

This concludes our survey of the fermionic Gaussian state formalism. Examples of its application to non-trivial calculations can be found in Chapter 4.

Bibliography

I. Affleck, T. Kennedy, E. H. Lieb, and H. Tasaki. Valence bond ground states in isotropic quantum antiferromagnets. *Comm. Math. Phys.*, 115(3): 477–528, 1988.

A. Ambainis. Quantum lower bounds by quantum arguments. *J. Comput. Syst. Sci.*, 64(4):750, 2002.

K. Audenaert, F. Verstraete, and B. De Moor. Variational characterizations of separability and entanglement of formation. *Phys. Rev. A.*, 64(5):052304, Oct 2001.

C. H. Bennett, G. Brassard, C. Crépeau, R. Jozsa, A. Peres, and W. K. Wootters. Teleporting an unknown quantum state via dual classical and Einstein-Podolsky-Rosen channels. *Phys. Rev. Lett.*, 70(13):1895–1899, Mar 1993.

F. A. Berezin. *The method of second quantisation*. Academic Press, 1966.

R. Bhatia. *Matrix Analysis*. Graduate Texts in Mathematics. Springer, 1997.

S. L. Braunstein, C. M. Caves, R. Jozsa, N. Linden, S. Popescu, and R. Schack. Separability of very noisy mixed states and implications for NMR quantum computing. *Phys. Rev. Lett.*, 83(5):1054–1057, Aug 1999.

S. Bravyi, M. B. Hastings, and F. Verstraete. Lieb-Robinson bounds and the generation of correlations and topological quantum order. *quant-ph/0603121*, 2006.

H. J. Briegel and R. Raussendorf. Persistent entanglement in arrays of interacting particles. *Phys. Rev. Lett.*, 86(5):910–913, Jan 2001.

H.-J. Briegel, W. Dür, J. I. Cirac, and P. Zoller. Quantum repeaters: the role of imperfect local operations in quantum communication. *Phys. Rev. Lett.*, 81(26):5932–5935, Dec 1998.

K. E. Cahill and R. J. Glauber. Density operators for fermions. *Phys. Rev. A.*, 59(2):1538–1555, Feb 1999.

M. Christandl, N. Datta, A. Ekert, and A. J. Landahl. Perfect state transfer in quantum spin networks. *Phys. Rev. Lett.*, 92(18):187902, 2004.

J. I. Cirac and P. Zoller. Quantum computations with cold trapped ions. *Phys. Rev. Lett.*, 74(20):4091–4094, May 1995.

J. I. Cirac, P. Zoller, H. J. Kimble, and H. Mabuchi. Quantum state transfer and entanglement distribution among distant nodes in a quantum network. *Phys. Rev. Lett.*, 78(16):3221–3224, Apr 1997.

J. I. Cirac, W. Dür, B. Kraus, and M. Lewenstein. Entangling operations and their implementation using a small amount of entanglement. *Phys. Rev. Lett.*, 86(3):544–547, Jan 2001.

T. S. Cubitt and J. I. Cirac. Engineering correlation and entanglement dynamics in spin systems, 2006a. In preparation.

T. S. Cubitt and J. I. Cirac. Engineering the correlation and entanglement dynamics of the XY–model, 2006b. In preparation.

T. S. Cubitt, F. Verstraete, W. Dur, and J. I. Cirac. Separable states can be used to distribute entanglement. *Phys. Rev. Lett.*, 91(3):037902, 2003.

T. S. Cubitt, F. Verstraete, and J. I. Cirac. Entanglement flow in multipartite systems. *Phys. Rev. A.*, 71(5):052308, 2005.

W. Dür, H.-J. Briegel, J. I. Cirac, and P. Zoller. Quantum repeaters based on entanglement purification. *Phys. Rev. A.*, 59(1):169–181, Jan 1999a.

W. Dür, J. I. Cirac, and R. Tarrach. Separability and distillability of multiparticle quantum systems. *Phys. Rev. Lett.*, 83(17):3562–3565, Oct 1999b.

W. Dür, G. Vidal, J. I. Cirac, N. Linden, and S. Popescu. Entanglement capabilities of nonlocal Hamiltonians. *Phys. Rev. Lett.*, 87(13):137901, Sep 2001.

A. Einstein, B. Podolsky, and N. Rosen. Can quantum-mechanical description of physical reality be considered complete? *Phys. Rev.*, 47(10):777–780, May 1935.

A. K. Ekert. Quantum cryptography based on Bell's theorem. *Phys. Rev. Lett.*, 67(6):661–663, Aug 1991.

R. P. Feynman. Simulating physics with computers. *Int. J. Theor. Physics*, 21(6/7):467–488, 1982.

F. R. Gantmacher. *Matrix Theory.* American Mathematical Society, English translation edition, 2000.

N. Gisin, G. Ribordy, W. Tittel, and H. Zbinden. Quantum cryptography. *Rev. Mod. Phys.*, 74(1):145–195, Mar 2002.

G. H. Golub and C. F. van Loan. *Matrix Computations.* Johns Hopkins University Press, third edition, 1996.

R. A. Horn and C. R. Johnson. *Matrix Analysis.* Cambridge University Press, 1985.

M. Horodecki, P. Horodecki, and R. Horodecki. Inseparable two spin–$\frac{1}{2}$ density matrices can be distilled to a singlet form. *Phys. Rev. Lett.*, 78(4):574–577, Jan 1997.

M. Horodecki, P. Horodecki, and R. Horodecki. General teleportation channel, singlet fraction, and quasidistillation. *Phys. Rev. A.*, 60(3):1888–1898, Sep 1999.

A. Jamiołkowski. Linear transformations which preserve trace and positive semidefiniteness of operators. *Rep. on Math. Phys.*, 3(4):275–278, Dec 1972.

R. Jozsa. Fidelity for mixed quantum states. *J. Mod. Opt.*, 41(12):2315–2323, Dec 1994.

E. Kamke. *Differentialgleichungen: Lösungsmethoden und Lösungen*, volume 1. Akademische Verlagsgesellschaft, Geest & Portig K.-G., 7^{th} edition, 1961.

N. Khaneja and S. J. Glaser. Efficient transfer of coherence through Ising spin chains. *Phys. Rev. A.*, 66(6):060301, Dec 2002.

B. Kraus and J. I. Cirac. Discrete entanglement distribution with squeezed light. *Phys. Rev. Lett.*, 92(1):013602, 2004.

E. Lieb, T. Schultz, and D. Mattis. Two soluble models of an antiferromagnetic chain. *Annals of Phys.*, 16(3):407–466, Dec 1961.

M. A. Nielsen and I. L. Chuang. *Quantum Computation and Quantum Information.* CUP, 2000.

M. A. Nielsen, M. J. Bremner, J. L. Dodd, A. M. Childs, and C. M. Dawson. Universal simulation of Hamiltonian dynamics for quantum systems with finite-dimensional state spaces. *Phys. Rev. A.*, 66(2):022317, Aug 2002.

T. J. Osborne and N. Linden. The propagation of quantum information through a spin system. *quant-ph/0312141*, 2003.

A. Peres. Separability criterion for density matrices. *Phys. Rev. Lett.*, 77(8): 1413–1415, Aug 1996.

M. Popp, F. Verstraete, M. A. Martin-Delgado, and J. I. Cirac. Localizable entanglement. *Phys. Rev. A.*, 71(4):042306, 2005.

E. Schrödinger. Discussion of probability distributions between separated systems. *Proc. Cam. Phil. Soc.*, 31:555–563, 1935.

E. Schrödinger. Probability relations between separated systems. *Proc. Cam. Phil. Soc.*, 32:446–451, 1936.

P. W. Shor. Polynomial-time algorithms for prime factorization and discrete logarithms on a quantum computer. *SIAM Journal on Computing*, 26(5): 1484–1509, 1997.

B. Simons. Concepts in theoretical physics: lecture notes. `www.tcm.phy.cam.ac.uk/~bds10/tp3.html`, 2001.

A. Uhlmann. The "transition probability" in the state space of a *-algebra. *Rep. Math. Phys.*, 9(2):273–279, Apr 1976.

S. J. van Enk, J. I. Cirac, and P. Zoller. Photonic channels for quantum communication. *Science*, 279(5348):205–208, Jan 1998.

F. Verstraete and H. Verschelde. Optimal teleportation with a mixed state of two qubits. *Phys. Rev. Lett.*, 90(9):097901, 2003.

F. Verstraete, M. Popp, and J. I. Cirac. Entanglement versus correlations in spin systems. *Phys. Rev. Lett.*, 92(2):027901, 2004.

G. Vidal. Efficient classical simulation of slightly entangled quantum computations. *Phys. Rev. Lett.*, 91(14):147902, 2003.

G. Vidal and R. Tarrach. Robustness of entanglement. *Phys. Rev. A.*, 59(1): 141–155, Jan 1999.

P. Wocjan, M. Rotteler, D. Janzing, and T. Beth. Universal simulation of Hamiltonians using a finite set of control operations. *Quantum Inf. Comput.*, 2(2):133–150, Feb 2002.

W. K. Wootters. Entanglement of formation of an arbitrary state of two qubits. *Phys. Rev. Lett.*, 80(10):2245–2248, Mar 1998.

M.-H. Yung, D. W. Leung, and S. Bose. An exact effective two-qubit gate in a chain of three spins. *Quantum Inf. Comput.*, 4(3):174–185, May 2003.

M. Żukowski, A. Zeilinger, M. A. Horne, and A. K. Ekert. "Event-ready-detectors" Bell experiment via entanglement swapping. *Phys. Rev. Lett.*, 71(26):4287–4290, Dec 1993.