Properties and Generalisations of Daemonic Ergotropy

Masterarbeit
zur Erlangung des akademischen Grades
Master of Science
(M.Sc.)

der Universität Siegen

Department Physik

vorgelegt von
Fabian Bernards

Juni 2018

Abstract

Daemonic Ergotropy quantifies the extractable energy from a state of a system S in an idealised laboratory with respect to a Hamiltonian if an ancilla A that may be correlated with S can be measured first [G. Francica et al., NPJ Quantum Information 3, 12 (2017)].

We extend the original definition of daemonic ergotropy to allow for generalised measurements and provide a multipartite generalisation. We propose a see-saw algorithm to find an optimal measurement and give analytical results for some classes of states.

Contents

Preface

This work is about energy extraction from quantum systems. Unlike other works [1] that are closer related to classical thermodynamics the main object under study is not a system in a heat bath. Instead of dealing with thermalisation, the system is placed in an idealised laboratory, in which an experimenter has full control over the unitary evolution of a given state. This sets the frame in which energy extraction is discussed here and the fundamental quantity is *ergotropy* [2], which is the maximal energy decrease a state can experience with respect to a reference Hamiltonian undergoing any arbitrary unitary evolution [2].

As always in this kind of issues, it is for a daemon to infuse them with just the right zest. Suppose, a daemon has access to an ancilla, which may be correlated with the system. Then, the energy it can extract from the system – the daemonic ergotropy – will be larger than the ergotropy because the daemon can perform a measurement on the ancilla and in that way learn something about the system before starting to extract energy [3].

In this work, we focus on the question what the optimal measurement is and how to find it. We also suggest to allow for generalised measurements in order to exploit the daemon's full potential and reinforce this suggestion by providing an example where the daemon can extract more energy using a generalised measurement than it could using a projective one. We however also show for two classes of states, pure states and quantum-classical states, that projective measurements are optimal.

Before addressing these questions, we begin with a brief overview on operational quantum mechanics in chapter one. Since correlations are key for the daemons ability to increase energy extraction, we discuss this topic in the second chapter, focusing on entanglement and discord. In the third chapter, we discuss the two works that introduce ergotropy and daemonic ergotropy, which provides the basis for our own considerations in the remaining chapters. In chapter four, we discuss daemonic gain for three-qubit systems before we address questions about the optimal measurement in the fifth chapter. We close with discussing a multipartite generalisation of daemonic ergotropy.

The appendix offers a short discussion of semidefinite programming, which we used in our algorithm to find the optimal measurement.

Chapter 1

States, Channels and Measurements

The basic framework in quantum mechanics to describe experiments is given by mainly three concepts. States describe the objects under study, channels describe how states can evolve or be manipulated and measurements describe how a measurement result is obtained from a given state. In this section we are going to discuss these concepts one after another.

1.1 States

In quantum mechanics, two different mathematical objects can serve as states. The first one, the *state vector*, is a unit-length vector in a complete inner product space, the *Hilbert space* [4]. Conversely, every unit vector in the Hilbert space is a state of the system, which gives rise to the superposition of states. This is a concept that is unknown in classical physics and leads to non-classic features such as entanglement and discord. If we consider two systems, A and B, with Hilbert spaces H_A and H_B, then the joint system has a state in $H_A \otimes H_B$, which is the Hilbert space spanned by the basis $\{ |i\rangle_A \otimes |j\rangle_B \}$, where $\{ |i\rangle_A \}$ and $\{ |i\rangle_B \}$ form the bases of H_A and H_B. If the joint system is in a product state $|\Psi\rangle_{AB}$, which means that there are vectors $|\phi\rangle_A$ and $|\psi\rangle_B$ such that $|\Psi\rangle_{AB} = |\phi\rangle_A |\psi\rangle_B$, then one can interpret this as system A being in state $|\phi\rangle_A$ and system B being in the state $|\psi\rangle_B$. However, if system AB is not in a product state, this interpretation fails. One can then not find vectors in the individual Hilbert spaces, that would describe the respective subsystem correctly. Thus, the state $|\Psi\rangle_{AB}$ is said to be *entangled*. Since one still needs to find a valid description for subsystems, the concept of a state needs to be extended.

This extension is found in density operators. Those are self adjoined, positive semidefinite operators with trace one. From now on, we will refer to density operators as states and denote the set of states acting on Hilbert space H as $S(H)$. States that can be described with a state vector $|\phi\rangle$ are called pure states and their density operator is the projector on the state vector, $|\phi\rangle\langle\phi|$. Any non-pure state is a mixed

state. The motivation to call states pure or mixed is the fact that any density matrix ϱ can be diagonalised and as such is a convex combination of pure states $\varrho = \sum_i p_i |i\rangle\langle i|$ which provides a correct description if the system is in one of the pure states $|i\rangle$, but one does not know in which one and therefore assigns probabilities p_i to the different cases. Anyway, this interpretation is not always possible, because mixed states are also used to describe parts of larger, entangled system in a pure state. If a system consists of two parts A and B with state spaces $S(H_A)$ and $S(H_B)$, one can compute the partial trace to yield the reduced state of system A as $\varrho_A = \mathrm{Tr}_B(\varrho_{AB})$, where ϱ_{AB} is the joint state. In fact, any mixed state $\varrho_A = \sum_i p_i |i\rangle\langle i|$ can be expressed as a reduced state from a pure state $|\psi\rangle_{AB} = \sum_i \sqrt{p_i}\, |i\rangle_A \otimes |i\rangle_B$, the *purification* of ϱ_A.

We will later return to discussing state properties in the section about entanglement right after covering channels and measurements.

1.2 Channels

Channels describe state transformations. Therefore, a channel is a linear map

$$\Lambda : S(H) \to S(K). \tag{1.1}$$

Channels may also capture changes in the dimension of the state space which is why the Hilbert spaces H and K do not necessarily coincide. Examples for this would be attaching or detaching an ancilla. Since every output of the mapping is a state again, channels are positive (map positive operators to positive operators) and trace-preserving. Mappings that are otherwise identical to channels but map states to subnormalised states are called *operations*. In this case, the process described by the operation only happens with some probability.

In general, a state $\varrho = \sum_i p_i \varrho_i$ can be decomposed into convex combinations in different ways and the evolution of the state should be determined by the evolution of its constituents such that

$$\Lambda(\varrho) = \sum_i p_i \Lambda(\varrho_i). \tag{1.2}$$

Here, the linearity of the channel ensures that the resulting state after applying the channel does not depend on the chosen convex decomposition of the input state.

If a channel Λ is applied only to part of a bigger system, the whole system will still be in a state after applying the channel. Hence, channels are not only positive but *completely positive* which means that any finite dimensional extension $\Lambda \otimes \mathbb{1}_d$ of the channel is still positive.

These considerations lead to the definition of a channel in the following way.

Definition 1.1 (Channel [4])**.** A channel is a linear, completely positive and trace preserving mapping $\Lambda : S(H) \to S(K)$.

So far, our notion of a channel is rather abstract. Therefore, we will now present the Choi-Jamiolkowski isomorphism. To any linear map between a n-dimensional and a

m-dimensional Hilbert space, for example a channel, this isomorphism assigns a linear operator on an $(n \times m)$-dimensional Hilbert space, called the *Choi matrix*[4]. Choi's theorem will then allow us to derive the two most popular representations of channels: The *Stinespring dilation* and the *Kraus form*. In the following, we denote the set of bounded linear operators on a Hilbert space H as $L(H)$.

Definition 1.2 (Choi-Jamiolkowski isomorphism [4])**.** The Choi-Jamiolkowski isomorphism $J : \phi \mapsto C_\phi$ maps a linear map $\phi : L(H_n) \to L(H_m)$ to the Choi matrix

$$C_\phi = (\mathbb{1}_n \otimes \phi) \sum_{i,j}^{n} |ii\rangle\langle jj| = \sum_{i,j} |i\rangle\langle j| \otimes \phi(|i\rangle\langle j|) \tag{1.3}$$

for a fixed basis $\{\,|i\rangle\}$ on H_n.

The inverse mapping $J^{-1} : C_\phi \mapsto \phi$ is [4]

$$J^{-1}(C_\phi)(X) = \mathrm{Tr}_1\left((X^T \otimes \mathbb{1})C_\phi\right) \tag{1.4}$$

where X is the operator the mapping ϕ would act on and $\mathrm{Tr}_1\,()$ the trace over the first system. This can be easily shown.

$$\begin{aligned}
J^{-1}(C_\phi)(X) &= \mathrm{Tr}_1\left((X^T \otimes \mathbb{1})C_\phi\right) && (1.5)\\
&= \mathrm{Tr}_1\left(\sum_{i,j} X^T|i\rangle\langle j| \otimes \phi(|i\rangle\langle j|)\right) && (1.6)\\
&= \sum_{n,i,j} \langle n|\, X^T\, |i\rangle|j\rangle\langle n|\phi(|i\rangle\langle j|) && (1.7)\\
&= \sum_{i,j} \langle j|\, X^T\, |i\rangle\phi(|i\rangle\langle j|) && (1.8)\\
&= \sum_{i,j} X_{ij}\phi(|i\rangle\langle j|) && (1.9)\\
&= \phi(\sum_{i,j} X_{ij}|i\rangle\langle j|) = \phi(X). && (1.10)
\end{aligned}$$

Theorem 1.1 (Choi's Theorem [5])**.** The following statements concerning a linear map $\phi : L(H_n) \to L(H_m)$ are equivalent:

1. $\mathbb{1}_n \otimes \phi$ is a positive map. (ϕ is n-positive)
2. The Choi matrix of ϕ, C_ϕ is positive semidefinite.
3. ϕ is completely positive.

Proof.

1. ⇒ 2. The Choi matrix of ϕ is positive, because ϕ is n-positive and $\sum_{ij} |ii\rangle\langle jj|$ up to a normalisation is a (positive semidefinite) pure state, so it is positive semidefinite as well.

2. ⇒ 3. We can diagonalise C_ϕ to write

$$C_\phi = \sum_i^{nm} |\psi_i\rangle\langle\psi_i| \tag{1.11}$$

in terms of the non-normalised states $|\psi_i\rangle$. Note, that we could only absorb the eigenvalues of C_ϕ into the definitions of the ψ_i because C_ϕ is positive by assumption. If we now define

$$P_k = \langle k|_n \otimes \mathbb{1}_m, \tag{1.12}$$

we can compute

$$P_k C_\phi P_l^\dagger = \phi(|k\rangle_n\langle l|_n) = \sum_i^{nm} P_k|\psi_i\rangle\langle\psi_i|P_l^\dagger. \tag{1.13}$$

Defining operators K_i such that

$$K_i |k\rangle_n = P_k |\phi_i\rangle \tag{1.14}$$

then allows us to write

$$\phi(|k\rangle_n\langle l|_n) = \sum_i^{nm} K_i |k\rangle_n\langle l|_n K_i^\dagger. \tag{1.15}$$

Because ϕ is linear, we can represent ϕ in the above way for any $A \in L(H_n)$ it could act on:

$$\begin{aligned}
\phi(A) &= \phi(\sum_{k,l} A_{kl}|k\rangle\langle l|) && (1.16)\\
&= \sum_{k,l} A_{kl}\phi(|k\rangle\langle l|) && (1.17)\\
&= \sum_{k,l} A_{kl} \sum_i K_i|k\rangle\langle l|K_i^\dagger && (1.18)\\
&= \sum_i K_i A K_i^\dagger. && (1.19)
\end{aligned}$$

This is known as the Kraus form of ϕ and the operators K_i are called *Kraus operators*. From the Kraus form, we can easily deduce that ϕ is completely positive. Therefore, we first note that for any positive operator A also $(\mathbb{1} \otimes K)A(\mathbb{1} \otimes K^\dagger)$ is positive as we can write

$$\langle\psi|(\mathbb{1} \otimes K)A(\mathbb{1} \otimes K^\dagger)|\psi\rangle = \langle\tilde{\psi}|A|\tilde{\psi}\rangle = \langle\tilde{\psi}|A|\tilde{\psi}\rangle \geq 0 \tag{1.20}$$

with $|\tilde{\psi}\rangle$ being some not necessarily normalised vector. Trivially, sums of positive operators are positive again, hence any map with a Kraus form is completely positive.

3. $\Rightarrow$ **1.** This is trivially true. □

For a channel, we additionally require the trace to be preserved. Therefore, a map is a channel, if and only if it has a Kraus form and is trace preserving,

$$\mathrm{Tr}\left(\phi(A)\right) = \mathrm{Tr}\left(\sum_i K_i A K_i^\dagger\right) = \mathrm{Tr}\left(A \sum_i K_i^\dagger K_i\right) = \mathrm{Tr}\left(A\right) \tag{1.21}$$

implying

$$\sum_i K_i^\dagger K_i = \mathbb{1}. \tag{1.22}$$

If the trace is not preserved but reduced, the map is an operation.

Another common way of representing a channel is provided by Stinespring's dilation theorem. It states that any channel can be modeled by a (joint) unitary transformation of the system and an environment, which initially are in a product state $\varrho \otimes \xi$. The new state of the system after action of the channel is then recovered by tracing out the environment.

Theorem 1.2 (Stinespring's dilation [4])**.** To every channel Λ acting on a system in some state $\varrho \in S(H)$, there exists a Hilbert space E with a pure state $\xi \in S(E)$ and a unitary acting on $H \otimes E$, such that

$$\Lambda(\varrho) = \mathrm{Tr}_E\left(U(\varrho \otimes \xi)U^\dagger\right). \tag{1.23}$$

Proof. A direct proof of a even slightly more general version of the theorem can be found for example in [6]. We are instead going to use a derivation in [7] and start with the Kraus form of Λ,

$$\Lambda(\varrho) = \sum_i K_i \varrho K_i^\dagger. \tag{1.24}$$

Given the Kraus operators, one can always define a unitary U such that

$$U\left|\phi\right\rangle_H \left|i\right\rangle_E = \sum_m (K_m \otimes \mathbb{1})\left|\phi\right\rangle_H \left|i+m\right\rangle_E. \tag{1.25}$$

One can see that U is indeed unitary by checking that it preserves scalar products [7]

$$\begin{aligned}
\left\langle\phi,0\right| U^\dagger U \left|\phi,0\right\rangle &= \sum_{m,n} \left\langle\phi,m\right| (K_m^\dagger \otimes \mathbb{1})(K_n \otimes \mathbb{1}) \left|\phi,n\right\rangle && (1.26)\\
&= \sum_{m,n} \left\langle\phi\right| K_m^\dagger K_n \left|\phi\right\rangle \otimes \left\langle m|n\right\rangle && (1.27)\\
&= \left\langle\phi\right| \sum_n K_n^\dagger K_n \left|\phi\right\rangle && (1.28)\\
&= \left\langle\phi\right| \mathbb{1} \left|\phi\right\rangle && (1.29)\\
&= 1. && (1.30)
\end{aligned}$$

Plugging U into equation (1.23) and choosing $\xi = |0\rangle\langle 0|$ yields

$$\mathrm{Tr}_E\left(U(\varrho \otimes |0\rangle\langle 0|)U^\dagger\right) = \mathrm{Tr}_E\left(\sum_{n,l} K_n \varrho K_l^\dagger \otimes |n\rangle\langle l|\right) \tag{1.31}$$

$$= \mathrm{Tr}_E\left(\sum_{n,l} K_n \varrho K_l^\dagger \otimes |n\rangle\langle l|\right) \tag{1.32}$$

$$= \sum_n K_n \varrho K_n^\dagger, \tag{1.33}$$

which ends the proof. □

One interpretation of Stinespring's dilation theorem is that channels describe transformations on subsystems of systems that in whole evolve unitarily.

1.3 Measurements

A description of a measurement in quantum mechanics should provide two pieces of information. First, it should define the probabilities for every measurement outcome $i \in \Omega$ in the set of outcomes Ω for a given state. Second, it should define the state of the system after it has been measured, depending on the measurement outcome. In this sense, a measurement is a mapping that maps the state that shall be measured to the post-measurement state. Therefore, if the set of outcomes Ω is discrete, one can define a *measurement operator* M_i for each outcome i, such that the state of the system in state ϱ will be proportional to $M_i \varrho M_i^\dagger$ and the probability p_i of the outcome i will be encoded in the normalisation of $M_i \varrho M_i^\dagger$ such that $p_i = \mathrm{Tr}\left(\varrho M_i^\dagger M_i\right)$ [7].

A measurement is called *complete*, if one obtains an outcome every time the measurement is performed. Then, the probabilities p_i have to sum up to one. Demanding completeness is not a restriction on the measurement, since one can always treat the case in which no outcome was obtained as a measurement result by simply introducing one more outcome label.

If we are only interested in the probabilities $p_i = \mathrm{Tr}\left(\varrho M_i^\dagger M_i\right)$ for the various outcomes but not in the post-measurement states, we do not need to know the measurement operators and it suffices to deal with the effects $E_i = M_i^\dagger M_i$ instead. From the completeness of the measurement it follows that $\sum_i E_i = \mathbb{1}$ because $\sum_i p_i = \sum_i \mathrm{Tr}\left(\varrho E_i\right) = \mathrm{Tr}\left(\varrho \sum_i E_i\right) = 1$. An operator is an effect, if it is positive-semidefinite with eigenvalues lower or equal to one which is fulfilled by the effects E_i.

1.3.1 Positive Operator Valued Measures

If the set of measurement outcomes Ω is not discrete, one can not simply define a measurement by explicitly assigning an effect to every possible outcome. In this case, we need a mapping that maps a subset of outcomes to an effect which will then allow

us to compute the probability for the outcome to be in the respective subset. Such a mapping is called *positive operator valued measure* (POVM). It is defined as follows:

Definition 1.3 (POVM [4])**.** A positive operator-valued measure (POVM) is a mapping $P : F \to E(H)$ such that

1. $P(\emptyset) = 0$,
2. $P(\Omega) = \mathbb{1}$,
3. $P(\cup_i X_i) = \sum_i P(X_i)$, for any sequence $\{X_i\}$ of disjoint sets in F,

where F is a sigma-algebra on Ω (collection of subsets of Ω including the empty set and Ω itself that is closed under countable unions and complement) and $E(H)$ is the set of effects on the Hilbert-space H.

In the definition, choosing F to be a sigma-algebra ensures that for outcome subsets of defined probability also the probabilities of their complements and countable unions are well defined. Lastly, the measurement is required to be complete, so the probability for the event 'any outcome' should be one and the probability for the event 'no outcome at all' should be zero.

In future discussions when dealing with a discrete set of outcomes we will identify a POVM P with the sequence $\{E_i := P(i)\}$ and refer to the effects E_i as *elements* of the POVM [4].

Sometimes, it can be of interest to find the optimal POVM for a task. If the objective function of the POVM is convex, it is helpful to note that POVMs form a convex set and in d dimensions an extremal POVM contains at most d^2 elements [8].

1.3.2 Projection Valued Measures

An important class of POVMs are *projection valued measures* (PVMs). These are POVMs with elements that are all projectors. For PVMs the measurement operators coincide with the effects. This implies that measurements described by PVMs are repeatable, which means that if the same PVM measurement is performed a second time, the post measurement states will not change, since the measurement operators are projectors.

1.3.3 The Naimark Extension

POVMs and PVMs are related to each other by the following fact: Any POVM can be completed to a PVM on an extended Hilbert space by a procedure called Naimark extension [9]. This means that given a projective measurement on a system, the measurement behaves like a POVM on a reduced system.

Consider a PVM with rank-one elements P_i acting on Hilbert space H with dimension d_H. Then, d_H orthonormal vectors $|w_i\rangle \in H$ define such a PVM if one sets $P_i := |w_i\rangle\langle w_i|$.

Now consider a POVM with N rank-one elements $A_i = |a_i\rangle\langle a_i|$ acting on H and $N \geq d_H$. Since $N \geq d_H$, the POVM does not refer to a projective measurement. The

Naimark extension now works in the following way: We introduce a second Hilbert space K with dimension $N - d_H$ and then find N orthonormal vectors $|w_i\rangle$ in the extended Hilbert space $H \oplus K$ of dimension N, where the projection of any $|w_i\rangle$ onto H shall be $|a_i\rangle$. Hence, we have to find N vectors $|b_i\rangle$ on K, such that the vectors

$$|w_i\rangle = |a_i\rangle \oplus |b_i\rangle \tag{1.34}$$

form an orthonormal basis of $H \oplus K$.

If we choose bases for the Hilbert spaces H and K, we can define a matrix α with column vectors $|a_i\rangle$ and a matrix β with column vectors $|b_i\rangle$. Demanding that the vectors $|w_i\rangle$ are orthonormal is then equivalent to demanding that the matrix

$$M = \begin{pmatrix} \alpha \\ \beta \end{pmatrix} \tag{1.35}$$

is unitary. This implies

$$MM^\dagger = \begin{pmatrix} \alpha \\ \beta \end{pmatrix} \begin{pmatrix} \alpha^\dagger & \beta^\dagger \end{pmatrix} = \begin{pmatrix} \alpha\alpha^\dagger & \alpha\beta^\dagger \\ \beta\alpha^\dagger & \beta\beta^\dagger \end{pmatrix} \stackrel{!}{=} \mathbb{1}_N \tag{1.36}$$

$$\Leftrightarrow \quad \alpha\alpha^\dagger = \mathbb{1}_{d_H},\ \alpha\beta^\dagger = \beta\alpha^\dagger = 0,\ \beta\beta^\dagger = \mathbb{1}_{N-d_H}. \tag{1.37}$$

The first equation (1.37) is automatically satisfied because the elements of a POVM sum up to identity. The second equation means that the d_H row vectors of α are orthogonal to the $N - d_H$ row vectors of β in an N-dimensional space. Such a β can be found, since the row vectors of α only span a d_H-dimensional subspace. One can even find the row vectors of β such that they are mutually orthonormal, such that the third equation is satisfied as well. The vectors b_i we aimed to in the beginning are then defined by the column vectors of β and a projective measurement on the extended space is found.

For every POVM P that does not have rank-one elements exclusively, one can define a second POVM P' for which every element of the original POVM has been split up into rank-one effects. Then, a projective measurement on an extended space can be found and the original POVM is obtained by assigning the same label to the elements of P' that constitute one element of P.

1.3.4 Ozawa's Theorem

Ozawa's theorem offers another way of reducing a generalised measurement to a projective measurement on an extended space. It states that for any generalised measurement with measurement operators M_m, one can append an ancilla in a pure state $|0\rangle\langle 0|_A$ to the system and find a unitary U acting on system and ancilla such that the post-measurement states

$$M_m \varrho M_m^\dagger = \mathrm{Tr}_A \left(P_m U(\varrho \otimes |0\rangle\langle 0|) U^\dagger \right) \tag{1.38}$$

are the same as if one had performed a projective measurement $P_m = \mathbb{1}_S \otimes |m\rangle\langle m|_A$ after the joint unitary evolution U. One can prove this statement in exactly the same

way as Stinespring's dilation theorem by showing that the above equation holds for the unitary

$$U\left|\phi\right\rangle_H\left|i\right\rangle_A = \sum_m (M_m \otimes \mathbb{1})\left|\phi\right\rangle_H\left|i+m\right\rangle_A. \tag{1.39}$$

In case we are only interested in the probabilities $p_m = \mathrm{Tr}\left(M_m \varrho M_m^\dagger\right)$ of the outcomes but not in the post-measurement states, the measurement on the system can be represented by a joint projective measurement on system and ancilla with projectors $U^\dagger P_m U$. This can be seen easily by applying the trace in equation (1.38) and exploiting that the trace is cyclic afterwards.

Chapter 2

Multipartite Systems, Entanglement and Classicality

2.1 Entanglement

A pure state is entangled, if it is not a product of two states. Mixtures of product states

$$\varrho_{\text{sep}} = \sum_i p_i \varrho_i^A \otimes \varrho_i^B. \tag{2.1}$$

are called *separable*. If a state is not separable, it is entangled [10]. In multipartite systems one distinguishes further between fully separable states, biseparable states and genuinely multipartite entangled states. Fully separable states are convex combinations of pure states that are product in every bipartition. Pure states that are a product regarding one bipartition and convex combinations of such pure states are called biseparable. Note, that the composing states do not need to be products with respect to the same bipartition. If a state is neither fully separable nor biseparable it is called genuinely multipartite entangled [10].

In the following we will discuss some ways to find out whether a state is separable or entangled.

2.1.1 The Positive Partial Transpose Criterion

There are many criteria for separability [10] but we will focus on the two that are used in this work.

The PPT, or *positive partial transpose* criterion is based on the observation that for any separable state

$$\varrho = \sum_i p_i \varrho_i^A \otimes \varrho_i^B \tag{2.2}$$

also the partial transpose

$$\varrho^{T_A} = \sum_i p_i (\varrho_i^A)^T \otimes \varrho_i^B \tag{2.3}$$

refers to a state, since the transpose is a positive and trace preserving map. However, the partial transpose of an entangled state is not necessarily a state anymore, as the partial transpose of the Bell state $|\psi\rangle = 1/2(|00\rangle + |11\rangle)$ is not positive. Although the partial transpose of a state depends on the chosen basis, its spectrum does not, which makes the PPT criterion basis independent [10]. For qubit-qubit and qubit-qutrit systems, the PPT criterion is not only a necessary but also a sufficient criterion for separability [11].

2.1.2 Bell Inequalities

Another way of proving a state is entangled is the violation of a Bell inequality. A Bell inequality establishes bounds on a function of expectation values that are obeyed by any system that allows for a description with a *local hidden variable* model (LHV). This is a model in which two assumptions, namely *realism* and *no-signaling*, are met. In the following, we discuss these assumptions and derive the general form of a bipartite LHV model. With this, we can reproduce the upper bound of the most famous Bell inequality, the Clauser-Horne-Shimony-Holt inequality. This inequality can be violated in quantum mechanics. However, separable states can never violate a Bell inequality, since they can always be described with a local hidden variable model, which is why Bell inequalities can be used to detect entanglement.

In the end, we discuss two Bell inequalities for the tripartite case.

Local Hidden Variable Models

Consider a setting with two parties in which Alice and Bob can both choose a measurement from their respective set of local measurements $\{A_1, \ldots, A_I\}$ and $\{B_1, \ldots, B_J\}$. Alice and Bob perform measurements at the same time and they evaluate their results together, so they record one outcome $m_{ij} = (a_i, b_j)$ if Alice obtained a_i after performing measurement A_i and Bob obtained b_j after performing measurement B_j. Alice and Bob can now extrapolate probability distributions $p(m_{ij})$ for their outcomes m_{ij} for all possible values of i and j. If Alice and Bob believe in *realism* however, they believe that in every instance of the experiment also combinations of measurements that were not performed would have obtained a value and thus it should be possible to find a joint probability distribution $p(m_{11}, m_{12}, \ldots)$ that has the probability distributions $p(m_{ij})$ as marginals [12]. This is equivalent to the belief that for any pair of measurements the outcome is predetermined by some hidden variable λ, since one can always write

$$p(m_{11}, m_{12}, \ldots) = \int \mathrm{d}\lambda p(\lambda)\chi(m_{11}, m_{12}, \ldots|\lambda) \tag{2.4}$$

where $\chi : \mathbb{R}^{IJ+1} \mapsto \{0, 1\}$ is a deterministic function and we assumed a discrete set of outcomes. It can be easily understood that this is possible by considering λ uniformly distributed on the interval $[0, 1]$ and defining

$$\chi(m_{11}, m_{12}, \ldots|\lambda) = \begin{cases} 1, \lambda \leq p(m_{11}, m_{12}, \ldots) \\ 0, \text{ otherwise.} \end{cases} \tag{2.5}$$

We can now write down a model for the probability distributions $p(m_{ij})$ by computing the marginal

$$p(m_{ij}) = \sum_{m_{kl}|k\neq i,l\neq j} p(m_{11}, m_{12}, \ldots) \tag{2.6}$$

$$= \int \mathrm{d}\lambda p(\lambda) \sum_{m_{kl}|k\neq i,l\neq j} \chi(m_{11}, m_{12}, \ldots|\lambda) \tag{2.7}$$

$$= \int \mathrm{d}\lambda p(\lambda)\tilde{\chi}(m_{ij}|\lambda), \tag{2.8}$$

where $\tilde{\chi}(m_{ij}|\lambda) = \tilde{\chi}((a_i, b_j)|\lambda)$ again only takes values zero and one. Since $\tilde{\chi}$ is a valid probability distribution, it will only take the value 1 for exactly one value of a_i and b_j for any given λ. Hence, it can be factorised as

$$\tilde{\chi}((a_i, b_j)|\lambda) = \tilde{\chi}^A_{ij}(a_i|\lambda)\tilde{\chi}^B_{ij}(b_j|\lambda), \tag{2.9}$$

where the factors χ^A_{ij} and χ^B_{ij} may still depend on the considered pair of measurements indicated by the indices i, j. We can now plug this into equation (2.8) to yield a realistic model for the experimentally quantifiable probability distributions

$$p(a_i, b_j) = \int \mathrm{d}\lambda p(\lambda)\tilde{\chi}^A_{ij}(a_i|\lambda)\tilde{\chi}^B_{ij}(b_j|\lambda). \tag{2.10}$$

Since Alice and Bob perform their measurements always at the same time, they will also assume *no-signaling*. This means, that Bob's choice of the measurement setting cannot affect Alice's outcome and vice versa such that

$$p(a_i) = \sum_{b_j} p(a_i, b_j) \quad \forall j \tag{2.11}$$

and

$$p(b_j) = \sum_{a_i} p(a_i, b_j) \quad \forall i \tag{2.12}$$

holds. If we now impose this additional restriction on our realistic model stated in equation (2.10), we directly obtain

$$\tilde{\chi}^A_{ij} = \tilde{\chi}^A_{ik} \quad \forall j, k \tag{2.13}$$

and

$$\tilde{\chi}^B_{ij} = \tilde{\chi}^B_{nj} \quad \forall i, n \tag{2.14}$$

which allows us to write down a realistic, non-signaling model

$$p(a_i, b_j) = \int p(\lambda)\tilde{\chi}^A(a_i|\lambda)\tilde{\chi}^B(b_j|\lambda)\ \mathrm{d}\lambda. \tag{2.15}$$

This is called a *local hidden variable* model and is oftentimes stated in the slightly different notation

$$p(a,b|i,j) = \int p(\lambda)\tilde{\chi}^A(a|i,\lambda)\tilde{\chi}^B(b|j,\lambda)\ \mathrm{d}\lambda. \tag{2.16}$$

Sometimes, the functions $\tilde{\chi}^A(a|i,\lambda)$ and $\tilde{\chi}^B(b|j,\lambda)$ are not assumed to only take values zero and one. However, this approach is completely equivalent, since one apply the same argument as in equation (2.5) on both functions. This generates two additional hidden variables that can later on be unified to one new hidden variable $\tilde{\lambda}$, which results in an equation of form (2.16) again.

The CHSH Inequality

For a bipartite system the most common Bell inequality is the Clauser-Horne-Shimony-Holt (CHSH) inequality

$$\langle A_1B_1\rangle + \langle A_2B_1\rangle + \langle A_1B_2\rangle - \langle A_2B_2\rangle \leq 2, \tag{2.17}$$

where A_1, A_2 are local measurements on the first, B_1, B_2 are local measurements on the second system and the outcomes a, b can take on values ± 1. In order to verify the upper bound of 2, we first note that since the probability distribution $p(a,b|i,j)$ is a convex combination in λ, the left hand side of the CHSH-inequality can never exceed the maximal value for a fixed λ. In this case, the joint expectation values factorise and the local expectation values are measurement results, that are obtained with certainty. We denote these outcomes of the measurements A_1, A_2, B_1, B_2 as $\alpha_1, \alpha_2, \beta_1, \beta_2$ respectively. For a fixed local hidden variable, the CHSH inequality then reads

$$\alpha_1\beta_1 + \alpha_2\beta_1 + \alpha_1\beta_2 - \alpha_2\beta_2 \leq 2. \tag{2.18}$$

After factorising the bound of the CHSH inequality becomes obvious.

When trying to prove a violation of a Bell inequality experimentally, one also assumes that the joint expectation values, which are given through the LHV model as

$$\langle A_iB_j\rangle = \int\ \mathrm{d}\lambda p(\lambda)\sum_{a,b}\tilde{\chi}^A(a|i,\lambda)\tilde{\chi}^B(b|j,\lambda)ab \tag{2.19}$$

can be estimated by computing the average value of the outcomes for the A_iB_j measurement. This can fail, if the performed measurement is influenced by the value of the local hidden variable (*freedom of choice loophole* [13]), or if the detected events are not a fair sample (*detection loophole*).

LHV Model for Separable States

In quantum mechanics, there is always a local hidden variable model for fully separable states. We can easily see this by evaluating equation (2.16) in a quantum mechanical

framework. For any measurements with labels i, j we have POVMs A^i, B^i that assign effects A^i_a, B^j_b to any outcomes a, b. We can then write

$$P(a,b|i,j) = \mathrm{Tr}\left(\sum_k p_k \varrho^A_k \otimes \varrho^B_k (A^i_a \otimes B^j_b)\right) \tag{2.20}$$

$$= \sum_k p_k \underbrace{\mathrm{Tr}\left(\varrho^A_k A^i_a\right)}_{p_A(a|i,k)} \underbrace{\mathrm{Tr}\left(\varrho^B_k B^j_b\right)}_{p_B(b|j,k)}, \tag{2.21}$$

so the joint probabilities behave according to a local hidden variable model with local hidden variable k. If any kind of states are allowed, the CHSH inequality can be violated and the left hand side of the inequality can take on values up to $2\sqrt{2}$, the Tsirelson bound. This maximal violation is obtained for the Bell state $|\psi\rangle = 1/\sqrt{2}(|00\rangle + |11\rangle)$.

Note, that entanglement is no sufficient condition to violate a Bell inequality and there are in fact entangled states for which a local hidden variable model can be found [10].

Mermin's Inequality and Svetlichny's Inequality

In the discussion of the CHSH inequality, we saw that is sufficient to consider the outcomes for a optimal value of the local hidden variable. We therefore adopt the notation where $\alpha_i, \beta_j, \gamma_k$ denote the outcomes for the measurements A_i, B_j, C_k for the fixed value of the local hidden variable to derive tripartite Bell inequalities from the CHSH inequality. One can write down two versions of the CHSH inequality in this notation:

$$S_2 := \alpha_1\beta_1 + \alpha_1\beta_2 + \alpha_2\beta_1 - \alpha_2\beta_2 \leq 2 \tag{2.22}$$

$$S_2' := \alpha_2\beta_2 + \alpha_2\beta_1 + \alpha_1\beta_2 - \alpha_1\beta_1 \leq 2 \tag{2.23}$$

where both bounds can be saturated simultaneously. Now we introduce the third system and write

$$S_3 := S_2(\gamma_1 + \gamma_2) + S_2'(\gamma_1 - \gamma_2) = 2(\alpha_2\beta_1\gamma_1 + \alpha_1\beta_2\gamma_1 + \alpha_1\beta_1\gamma_2 - \alpha_2\beta_2\gamma_2) \leq 4, \tag{2.24}$$

which directly yields *Mermin's* inequality [14]

$$\langle A_2B_1C_1\rangle + \langle A_1B_2C_1\rangle + \langle A_1B_1C_2\rangle - \langle A_2B_2C_2\rangle \leq 2. \tag{2.25}$$

The bound will be obtained for extremal values (± 2) of S_2 and S_2'. Mermin's inequality in principle captures all kinds of non-local correlations including two party correlations. In contrast, for deriving *Svetlichny's* inequality, one only employs a bipartite local hidden variable model, treating Alice and Bob as one party and Charlie as the other. Consequently, for a fixed hidden variable the pairs outcomes $(\alpha_i\beta_j)$ are fixed as well as the outcomes γ_k for Charlie. However, there are not necessarily fixed values for the local outcomes of Alice and Bob, since these two systems are not described by a local hidden variable model. Thus, non-local correlations between Alice and Bob only will not lead to a violation of Svetlichny's Inequality.

One can now construct Svetlichny by calculating $S_2\gamma_1 - S_2'\gamma_2$. This yields

$$\begin{aligned}(\alpha_1\beta_1)\gamma_1 + (\alpha_1\beta_1)\gamma_2 + (\alpha_1\beta_2)\gamma_1 - (\alpha_1\beta_2)\gamma_2 \\ +(\alpha_2\beta_1)\gamma_1 - (\alpha_2\beta_1)\gamma_2 - (\alpha_2\beta_2)\gamma_1 - (\alpha_2\beta_2)\gamma_2 \leq 4\end{aligned} \tag{2.26}$$

and the upper bound is easily verified by rewriting the left-hand side as

$$((\alpha_1\beta_1) - (\alpha_2\beta_2))(\gamma_1 + \gamma_2) + ((\alpha_1\beta_2) + (\alpha_2\beta_1))(\gamma_1 - \gamma_2). \tag{2.27}$$

Svetlichny's equality then reads [14]

$$\begin{aligned}\langle A_1B_1C_1\rangle + \langle A_1B_1C_2\rangle + \langle A_1B_2C_1\rangle - \langle A_1B_2C_2\rangle \\ +\langle A_2B_1C_1\rangle - \langle A_2B_1C_2\rangle - \langle A_2B_2C_1\rangle - \langle A_2B_2C_2\rangle \leq 4\end{aligned} \tag{2.28}$$

and since it is symmetric with respect to all three parties any nonlocal correlations that only concern two of the parties will not lead to a violation of Svetlichny's equality. Hence, a violation of Svetlichny's inequality indicates a genuinely multipartite entangled state.

2.1.3 Entanglement Measures and Maximally Entangled States

Entanglement measures quantify entanglement. There is no unique way to do this, as many different measures have been proposed, but in order to qualify as an entanglement measure, some requirements have to be met. First, the entanglement of separable states is zero. Second, local operations and classical communication (LOCC) should not increase the entanglement of a state. In particular, this implies that entanglement measures are invariant under local unitary transformations [10].

Due to these conditions, there is only one possible candidate as a unique maximally entangled state for two qubits. Since entanglement measures are invariant under local unitaries, the Schmidt decomposition of any state is equivalent to

$$|\psi\rangle = \sqrt{\lambda}\,|00\rangle + \sqrt{1-\lambda}\,|11\rangle. \tag{2.29}$$

The invariance under local unitaries ensures symmetry under exchange of λ and $1-\lambda$. Thus, if there is one maximally entangled state, it must be the Bell state

$$|\psi\rangle = \frac{1}{\sqrt{2}}(\,|00\rangle + |11\rangle). \tag{2.30}$$

This is motivation to capture the entanglement of a pure state by quantifying the minimal amount of superpositions for any decomposition of the state in a product basis. One way of doing this in the bipartite case is calculating the entropy of the Schmidt coefficients. For pure states, this is equivalent to the entropy of the reduced density matrix. This entanglement measure is the entanglement of formation [10]. For mixed states it is extended using the convex roof construction. Given a state $\varrho = \sum_i p_i|\phi_i\rangle\langle\phi_i|$ the entanglement is then defined as

$$E(\varrho) = \inf_{p_i,\,|\phi_i\rangle} \sum_i p_i E(\,|\phi_i\rangle), \tag{2.31}$$

where all possible decompositions of ϱ are considered. This infimum is generally hard to find but for two qubits an analytical solution involving the concurrence is known. Concurrence is an entanglement measure, defined as

$$C(|\phi\rangle) = \sqrt{2(1 - \mathrm{Tr}\,(\varrho_A^2))} \tag{2.32}$$

for pure states and its convex roof for two qubits is

$$C(\varrho) = \max\{0, \lambda_1 - \lambda_2 - \lambda_3 - \lambda_4\}. \tag{2.33}$$

The λ_i are the eigenvalues of the matrix $X = \sqrt{\sqrt{\varrho}(\sigma_y \otimes \sigma_y)\varrho^*(\sigma_y \otimes \sigma_y)\sqrt{\varrho}}$ in decreasing order. Now, the entanglement of formation is

$$E_F(\varrho) = h(\frac{1}{2}(1 + \sqrt{1 - C^2(\varrho)})) \tag{2.34}$$

with the binary entropy function $h(p) = -p\log(p) - (1-p)\log(1-p)$ [10].

Another way of creating an entanglement measure is utilising the PPT criterion to define the negativity as

$$N(\varrho) = \frac{1}{2}(\|\varrho^{T_B}\|_1 - 1) \tag{2.35}$$

with the trace norm $\|\varrho^{T_B}\|_1 = \mathrm{Tr}\left(\sqrt{(\varrho^{T_B})^\dagger \varrho^{T_B}}\right)$. Diagonalising ϱ^{T_B} and using that ϱ^{T_B} has trace one shows that one can rewrite the negativity as

$$N(\varrho) = \frac{1}{2}\sum_i |\lambda_i| - \lambda_i \tag{2.36}$$

where λ_i are the eigenvalues of ϱ^{T_B}. This means that the negativity is equal to the sum of the negative eigenvalues of the partial transpose matrix. Later, when dealing with three qubits, we will calculate the geometric average of the negativities with respect to every bipartition and simply refer to this quantity as negativity again.

2.1.4 Three Qubits

We have argued in the previous section, why

$$|\psi\rangle = \frac{1}{\sqrt{2}}(|00\rangle + |11\rangle) \tag{2.37}$$

is the maximally entangled two qubit state. In higher dimensions one cannot recognise one state as maximally entangled. Instead, entangled states split up into several classes, so-called SLOCC classes, since states within each class are related to each other via invertible stochastic local operations and classical communication (SLOCC). It has been shown in [15] that invertible SLOCCs act as invertible local operators such as

$$|\phi\rangle = A \otimes B \otimes \ldots \otimes N\,|\psi\rangle \tag{2.38}$$

with invertible operators $A, B, \ldots, N$ on a state. For three qubits pure states can be divided into states that are completely product, bipartite entangled states and two classes of genuinely tripartite entangled states: The GHZ-class and the W-class. These classes are named after their most famous members, the Greenberger-Horne-Zeilinger state, short GHZ-state

$$|GHZ\rangle = \frac{1}{\sqrt{2}}(\,|000\rangle + |111\rangle) \tag{2.39}$$

and the W-state

$$|W\rangle = \frac{1}{\sqrt{3}}(\,|001\rangle + |010\rangle + |100\rangle). \tag{2.40}$$

These states are maximally entangled in a sense that any three-qubit state that is neither in the GHZ-class nor the W-class can be reached via SLOCC operations from these states. Under local unitaries, any state can be brought into the form

$$|\psi\rangle = \lambda_0\,|000\rangle + \lambda_1 e^{i\phi}\,|100\rangle + \lambda_2\,|101\rangle + \lambda_3\,|110\rangle + \lambda_4\,|111\rangle. \tag{2.41}$$

Alternatively, also a decomposition in the form

$$|\psi\rangle = \kappa_0 e^{i\theta}\,|000\rangle + \kappa_1\,|001\rangle + \kappa_2\,|010\rangle + \kappa_3\,|100\rangle + \kappa_4\,|111\rangle \tag{2.42}$$

which is a superposition of a GHZ-class and a W-class state is also possible [16]. However, the coefficients κ_i are not always unique in the second decomposition (equation 2.42).

2.2 State Discrimination

The general task in state discrimination is to find out in which state a system is prepared, given an ensemble of possible states ϱ_k that occur with a priori probabilities p_k [17]. For non-orthogonal states, this task can not be done perfectly [7]. If one can discriminate between two states, say ϱ_0 and ϱ_1 perfectly, then there needs to be an effect E, such that $\mathrm{Tr}\,(\varrho_0 E) = 0$ and $\mathrm{Tr}\,(\varrho_1 E) = 1$. Since $E \leq \mathbb{1}$, $\mathrm{Tr}\,(\varrho_1 E) = 1$ implies that E is a sum of a projector on the range of ϱ_1 and an effect acting on the complementary space. From $\mathrm{Tr}\,(\varrho_0 E) = 0$ we can then conclude that ϱ_0 is a mixture of states that lie in the complementary space exclusively and so ϱ_0 is orthogonal to ϱ_1. For orthogonal states, an optimal POVM can be chosen to be a PVM and its elements are the projectors acting on the orthogonal subspaces defined by the ranges of the states.

2.2.1 Unambiguous State Discrimination

If the ensemble contains non-orthogonal states, an unambiguous distinction is in general not possible. In (optimal) *unambiguous state discrimination* one therefore seeks a measurement that allows unambiguous state discrimination for some outcomes while the probability for ambiguous outcomes is as low as possible [18].

Optimal unambiguous state discrimination was an early example for a task that can be better performed with generalised measurements rather than projective measurements. We therefore discuss the following problem which was first understood by Ivanovic and Peres [19] [20]. Consider an ensemble consisting of two non-orthogonal pure states $|\psi_0\rangle$ and $|\psi_1\rangle$ occurring with probabilities $p_0 \leq p_1$ without loss of generality. Then the optimal projective measurement has the elements $\Pi_0 = |\psi_0\rangle\langle\psi_0|$ and $\Pi_1 = \mathbb{1} - \Pi_0$, where outcome 1 implies that the state $|\psi_1\rangle$ was present and outcome 0 is inconclusive. The probability of this outcome is

$$p = \mathrm{Tr}\left(|\psi_0\rangle\langle\psi_0|[p_0|\psi_0\rangle\langle\psi_0| + p_1|\psi_1\rangle\langle\psi_1|]\right) = p_0 + p_1|\langle\psi_0|\psi_1\rangle|^2. \tag{2.43}$$

However, if we use a POVM, we can have two conclusive results corresponding to the elements

$$E_0 = \alpha_0(\mathbb{1} - |\psi_1\rangle\langle\psi_1|) \tag{2.44}$$
$$E_1 = \alpha_1(\mathbb{1} - |\psi_0\rangle\langle\psi_0|) \tag{2.45}$$

that are completed by an element corresponding to an inconclusive result

$$E_2 = \mathbb{1} - E_0 - E_1 = (1 - \alpha_0 - \alpha_1)\mathbb{1} + \alpha_0|\psi_1\rangle\langle\psi_1| + \alpha_1|\psi_0\rangle\langle\psi_0|. \tag{2.46}$$

The factors $\alpha_0, \alpha_1 \in [0, 1]$ are necessary to ensure $E_2 \succeq 0$. Now the probability for an inconclusive outcome is

$$p = \mathrm{Tr}\left(E_2(p_0|\psi_0\rangle\langle\psi_0| + p_1|\psi_1\rangle\langle\psi_1|\right) \tag{2.47}$$
$$= 1 - (1 - |\langle\psi_0|\psi_1\rangle|^2)(\alpha_0 p_0 + \alpha_1 p_1). \tag{2.48}$$

This is better than what one could achieve with PVMs. To see this, consider the simple case $p_0 = p_1 = |\langle\psi_0|\psi_1\rangle|^2 = 1/2$. For symmetry reasons assume that $\alpha_0 = \alpha_1$. We now want to choose α_0, α_1 optimal. Therefore, we first calculate the eigenvalues of $E_0 + E_1$

$$\lambda_{1,2} = \frac{1}{2}(\alpha_0 + \alpha_1) \pm \sqrt{\frac{(\alpha_0 + \alpha_1)^2}{4} - \alpha_0\alpha_1(1 - |\langle\psi_0|\psi_1\rangle|^2)} \tag{2.49}$$

and note that for an optimal measurement α_0 and α_1 will be chosen such that E_2 becomes rank-one and therefore the larger eigenvalue will be 1. Then, $E_0 + E_1 \preceq \mathbb{1}$ holds, so E_2 is positive semidefinite and $\alpha_0 = 1/(1+\sqrt{2}/2)$. We can now compute the probabilities of receiving an inconclusive outcome for the projective measurement and the generalised measurement to see that using a POVM is indeed advantageous

$$p_{\mathrm{PVM}} = \frac{3}{4} > p_{\mathrm{POVM}} = \frac{1}{\sqrt{2}}. \tag{2.50}$$

2.2.2 Minimum Error Discrimination

The problem of minimum error discrimination is to determine the state of a system that occurs in one state ϱ_n of a set of states $\{\varrho_i\}$ with an associated a priori probability p_n by performing a measurement while minimising the probability of making a wrong guess. The strategy is to perform a measurement represented by a POVM P with elements E_i and guessing that the system was in state ϱ_i whenever the measurement outcome was i. The task is then to find the POVM that minimises the expectation value of the error

$$P_{\text{err}}^{\min} = \min_P \sum_i \sum_{j \neq i} p_j \text{Tr}\left(\varrho_j E_i\right) \tag{2.51}$$

when considering all possible measurement outcomes. This is equivalent to maximising the probability of getting a correct result [17]

$$P_{\text{corr}} = 1 - P_{\text{err}}^{\min} = \max_P \sum_i p_i \text{Tr}\left(\varrho_i E_i\right). \tag{2.52}$$

We now follow [17] and present the minimum error condition. If a POVM with elements E_i is optimal, then the respective success probability is higher than for any other POVM with elements F_i

$$\sum_i p_i \text{Tr}\left(\varrho_i E_i\right) \geq \sum_j p_j \text{Tr}\left(\varrho_j F_j\right) \tag{2.53}$$

One can now insert $\sum F_j = \mathbb{1}$ to obtain

$$\sum_{ij} p_i \text{Tr}\left(\varrho_i E_i F_j\right) - p_j \text{Tr}\left(\varrho_j F_j\right) \geq 0 \tag{2.54}$$

$$\Leftrightarrow \quad \sum_j \text{Tr}\left(\left(\sum_i p_i \varrho_i E_i - p_j \varrho_j\right) F_j\right) \geq 0 \tag{2.55}$$

which implies that the optimal POVM has been found if

$$\sum_i p_i \varrho_i E_i - p_j \varrho_j \geq 0, \quad \forall j \tag{2.56}$$

because the operators F_j are positive semidefinite. On a computer, one can find the optimal POVM with software for convex optimisation since equation (2.52) defines a semidefinite programme. For semidefinite programms, solutions can be found efficiently and certified [21]. We discuss this in more detail in the appendix.

2.3 Classicality and Discord

Classical theories are fundamentally deterministic and realistic which allows to identify the state of a system with its properties, i.e. the outcomes that would certainly be obtained by performing perfect measurements on the system. In classical physics, a

measurement is perfect if it reveals the measured property without an error and without changing the state. If two systems are in different states, they differ in at least one property. The possibility of sharply measuring these properties then makes any two different classical states perfectly distinguishable.

The quantum mechanical equivalent of this situation is as follows. Given a state $|i\rangle$ of the system, the perfect measurement has to have one outcome i that unambiguously and with certainty indicates that the system is in state $|i\rangle$. Hence, the POVM E representing the measurement has an element E_i such that $E_i\,|j\rangle = \delta_{ij}\,|j\rangle$. Consequently, the states of the system have to be orthogonal and the perfect measurement for this system is a rank-one projective measurement. If the state (in the classical meaning) is not known, we can account for this by describing the system with the density matrix $\varrho = \sum_i p_i |i\rangle\langle i|$, which here means that the system is in one of the states $|i\rangle$ and one assigns probabilities p_i to the different cases. While the assumptions of determinism and realism automatically evoke the ignorance interpretation of mixed states, this interpretation is not generally viable in a more general quantum mechanical setting, since the mixed state could be a local description of a larger, entangled system in a pure state.

As all states in this classical model have to be orthogonal, there is no superposition and thus also no entanglement in multipartite systems. In bipartite systems, one makes the following distinction: A classical-classical state is a state of the form

$$\varrho = \sum_{ij} c_{ij} |i_A j_B\rangle\langle i_A j_B|, \tag{2.57}$$

where the states $|i_A\rangle$ and $|j_B\rangle$ form an orthonormal basis of their respective Hilbert spaces, which are the measurement bases. Such a state describes a bipartite classical system in the above sense and its density matrix is diagonal in the measurement basis. The lack of off-diagonal terms in a classical model has provoked interest and they have henceforth been studied under the name *coherences* [22] [23].

A state of the form

$$\varrho = \sum_{ij} q_{ij} |a_i j_B\rangle\langle a_i j_B| \tag{2.58}$$

$$= \sum_j p_j\, \varrho_j^A \otimes |j_B\rangle\langle j_B|, \tag{2.59}$$

where $|j_B\rangle$ is the preferred basis of system B again and $|a_i\rangle$ are non-orthogonal states is called *quantum-classical*. If we now go on to also drop any restriction on the states of system B, we obtain the general form of a separable bipartite state.

A measure for the quantumness of the correlations in a bipartite state is *quantum discord* [24] [25]. The idea of discord is to quantify the total correlations in the state and subtract the classical correlations from it [24]. As a measure for the total correlations mutual information is used. The mutual information of a state is the information of the state as a whole minus the information that can be inferred from the state about its

subsystems. In quantum mechanics, the mutual information is

$$I_{\text{mut}}(\varrho_{AB}) = I(\varrho_{AB}) - I(\varrho_A) - I(\varrho_B) \tag{2.60}$$

where the information $I(\varrho) = -S(\varrho) = \text{Tr}\,(\varrho \log \varrho)$ is the negative von Neumann entropy. Mutual information can easily be shown to be non-negative and the difference between the information in the joint state and the information in the reduced states is just the information about the correlations which is missing in the latter.

In order to discriminate between classical and non-classical correlations, one can utilise that the perfect measurement does not disturb the classical state and consequently also does not affect its inherent correlations. Thus, for a bipartite system with parties A and B that behaves classically on subsystem B such as

$$\varrho = \sum_i p_i\, \varrho_i \otimes |i\rangle\langle i| \tag{2.61}$$

the mutual information equals the additional accessible information about system A when measuring system B

$$J_{A:B}(\varrho_{AB}) = \sum_i p_i I(\varrho_i) - I(\varrho_A), \tag{2.62}$$

where $\varrho_A = \text{Tr}_B\,(\varrho) = \sum_i p_i \varrho_i$ and ϱ_i are the conditional states of system A if the outcome of the measurement on B was i. This can be easily seen by calculating $I(\varrho)$ for a state as given in equation (2.61) and diagonalising the conditional states $\varrho_i = \sum_j c_{ij} |j^{(i)}\rangle\langle j^{(i)}|$.

$$\begin{aligned}
I(\varrho) &= \text{Tr}\left(\sum_i p_i \varrho_i \otimes |i\rangle\langle i| \log(\sum_l p_l \varrho_l \otimes |l\rangle\langle l|)\right) && (2.63)\\
&= \text{Tr}\left(\sum_{ij} p_i c_{ij} |j^{(i)}\rangle\langle j^{(i)}| \otimes |i\rangle\langle i| \log(\sum_{lm} p_l c_{lm} |m^{(l)}\rangle\langle m^{(l)}| \otimes |l\rangle\langle l|)\right) && (2.64)\\
&= \sum_{ij} p_i c_{ij} \log(p_i c_{ij}) && (2.65)\\
&= \sum_i p_i \log p_i + \sum_{ij} p_i c_{ij} \log c_{ij} && (2.66)\\
&= I(\varrho_B) + \sum_i p_i I(\varrho_i). && (2.67)
\end{aligned}$$

Plugging this into equation (2.60) then leads to the same expression as in equation (2.62).

It has been shown [25] that for general states the mutual information $I_{\text{mut}}(\varrho_{AB})$ is always larger or equal to the accessible mutual information $J_{A:B}$ for any measurement, defined as

$$J_{A:B}(\varrho_{AB}, \mathbf{P}) = \sum_i p_i I(\varrho_{A|i}) - I(\varrho_A) \tag{2.68}$$

with $\varrho_{A|i} = {}^{\mathrm{Tr}_B(\varrho_{AB}(\mathbb{1}\otimes P_i))}/_{\mathrm{Tr}(\varrho_{AB}(\mathbb{1}\otimes P_i))}$ and $p_i = \mathrm{Tr}\left(\varrho_{AB}(\mathbb{1}\otimes P_i)\right)$. For the optimal measurement, the difference between the two is the quantum discord

$$\delta_{A:B}(\varrho_{AB}) = I_{\mathrm{mut}}(\varrho_{SA}) - \max_{\mathbf{P}} J_{A:B}(\varrho_{AB}, \mathbf{P}). \tag{2.69}$$

While Olliver and Zurek [25] restrict the possible measurements to rank-one projective measurements, Henderson and Vedral [24] allow for any generalised measurements. However, it has been shown that at least for two qubits, projective measurements are optimal for rank-two states and they give the same discord for rank-three and rank-four states up to small corrections [26]. We have already seen in the above discussion that a quantum-classical state has zero discord $\delta(A:B)$. Indeed, Olliver Zurek additionally to this showed the converse statement: Vanishing discord $\delta(A:B)$ implies that the state is quantum-classical [25].

Another way to introduce discord is to state a set of rules any measure for classical correlations should adhere to and claiming that the accessible mutual information qualifies as such a measure [24][1].

If the discord vanishes for a state, this implies that the density matrix has a block-diagonal shape in the measurement basis [25]. Hence, demanding vanishing discord for both parties implies that the state is classical-classical. On the other hand, for two qubits discord takes its maximum for the Bell state and decreases monotonically if the Bell state is mixed with white noise. However, the maximal discord for a given purity can also increase for lower purity, as has been shown in [27] for two qubits. We reproduce this result in figure (2.1), which shows a plot of discord versus linear entropy $S_L = {}^4/_3(1 - \mathrm{Tr}\left(\varrho^2\right))$ including the upper bound. Maximal discord for a given purity can be achieved with Werner states $\varrho = p|\phi^+\rangle\langle\phi^+| + (1-p)\mathbb{1}$ in a low purity regime and with a subset of so-called two-parameter states

$$\varrho(a,b) = \frac{1}{2}\begin{pmatrix} a & 0 & 0 & a \\ 0 & 1-a-b & 0 & 0 \\ 0 & 0 & 1-a+b & 0 \\ a & 0 & 0 & a \end{pmatrix} \tag{2.70}$$

for higher purities. Two-parameter states with $b = 0$ are referred to as α states. In the transition point for $S_L \approx 0.9$ states can entail higher discord than more impure but also slightly purer states.

[1] The accessible mutual information does in fact not meet the suggested criteria for a measure of classical correlations as it is not symmetric under exchange of the subsystems.

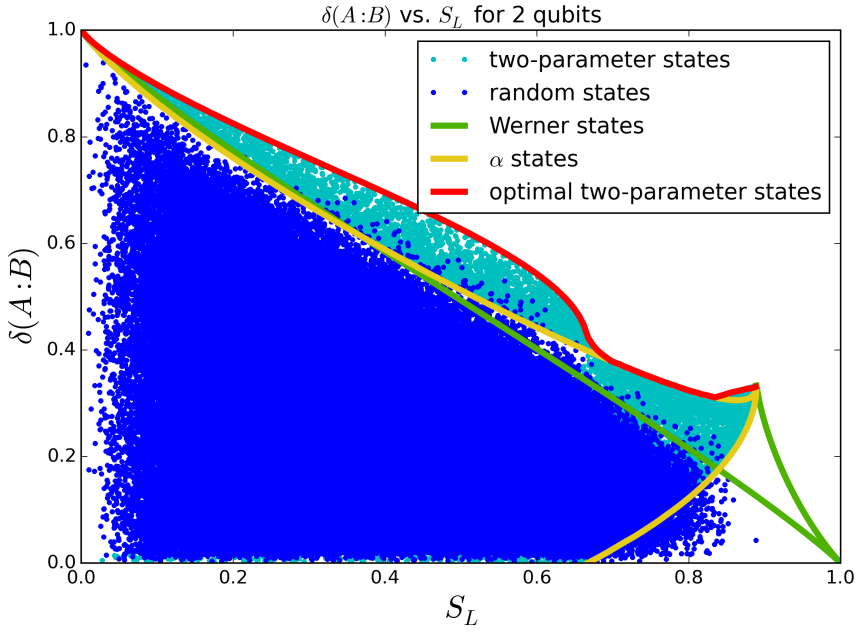

Figure 2.1: Discord versus linear entropy $S_L = 4/3(1 - \mathrm{Tr}\,(\varrho^2))$ for two qubits. The plot is similar to the one provided in [27].

Chapter 3

Ergotropy, Daemonic Ergotropy and Daemonic Gain

In this chapter, we discuss energy extraction from states while focusing on three quantities: ergotropy, daemonic ergotropy and daemonic gain. Ergotropy hereby is the maximal amount of energy that can be extracted via a unitary channel [2]. Daemonic Ergotropy is a recent extension to the concept in a bipartite setting, quantifying the extractable energy on one party after performing a measurement on the other [3]. The maximal gain from performing a measurement is the daemonic gain [3]. In the following, we will discuss each quantity in more detail, stating the definition first and collecting some useful and previously established properties.

3.1 Ergotropy

Consider a system S in a given state ϱ_S and a Hamiltonian H acting on this system. This defines the energy of the state. Now assume, the system is exposed to an additional, time dependent external potential $V(t)$ for some time, which may in principle induce any unitary time evolution U [2]. By how much can the energy of the system have decreased at most through that process? Comparing the energy of the system before the action of the potential $V(t)$ to the minimal possible energy afterwards directly leads to the definition of ergotropy.

Definition 3.1 (Ergotropy [2])**.** The ergotropy for a given density matrix ϱ and Hamiltonian H is given by

$$W(\varrho, H) = \mathrm{Tr}\left(\varrho H\right) - \min_{U} \mathrm{Tr}\left(U\varrho U^\dagger H\right), \tag{3.1}$$

where the minimisation runs over all unitary matrices.

Restricting oneself to unitary channels makes sense, as allowing for any type of channel would reduce the problem to absurdity. Considering the Stinespring dilation

$$\Lambda(\varrho) = \mathrm{Tr}_E\left(U(\varrho \otimes |0\rangle\langle 0|)U^\dagger\right) \tag{3.2}$$

clarifies this. One way to minimise the energy of the state would be equivalent to exchanging it for a pure state by performing a swap

$$U_{\text{swap}} = \sum_{ij} |i\rangle\langle j| \otimes |j\rangle\langle i| \tag{3.3}$$

and mapping the pure state to the ground state of the Hamiltonian afterwards, which both are unitary operations.

3.1.1 Ergotropy Is Sublinear And Convex

For the sake of a later argument, we treat ergotropy as a function of unnormalised states and note that ergotropy is sublinear in both the state and the Hamiltonian. Since ergotropy is symmetric in its arguments, it suffices to show the sublinearity in the state.

$$\begin{aligned} W(\varrho_1 + \varrho_2, H) &= \operatorname{Tr}\left((\varrho_1 + \varrho_2)H\right) - \min_U \operatorname{Tr}\left(U(\varrho_1 + \varrho_2)U^\dagger H\right) && (3.4)\\ &\leq \operatorname{Tr}\left(\varrho_1 H\right) + \operatorname{Tr}\left(\varrho_2 H\right) && (3.5)\\ &\quad - \min_U \operatorname{Tr}\left(U\varrho_1 U^\dagger H\right) - \min_{\tilde{U}} \operatorname{Tr}\left(\tilde{U}\varrho_2 \tilde{U}^\dagger H\right) \\ &= W(\varrho_1, H) + W(\varrho_2, H) && (3.6)\\ W(\lambda\varrho, H) &= \lambda W(\varrho, H). && (3.7) \end{aligned}$$

This implies, that ergotropy is also convex in both arguments.

3.1.2 Calculating Ergotropy

Although the definition of ergotropy includes a minimisation, it is actually easy to compute as the minimisation effectively comes down to diagonalising the state and the Hamiltonian, as was argued in [2]. In the following however, we give a detailed proof.

Theorem 3.1. The ergotropy for a state $\varrho = \sum_{i=0}^{d-1} r_i |r_i\rangle\langle r_i|$ with eigenvalues $r_{i+1} \leq r_i$ and a Hamiltonian $H = \sum_{i=0}^{d-1} e_i |e_i\rangle\langle e_i|$ with eigenvalues $e_{i+1} \geq e_i$ is

$$W = \operatorname{Tr}\left(\varrho H\right) - \sum_{i=0}^{d-1} r_i e_i \tag{3.8}$$

and the optimal unitary is

$$V = \sum_{i=0}^{d-1} |e_i\rangle\langle r_i|. \tag{3.9}$$

Proof. The theorem directly follows from the trace inequality of John von Neumann [28]. Alternatively, one can exploit Birkhoff and von Neumann's theorem about doubly

stochastic matrices to prove the theorem [29], which we will show here.

$$\min_{U\in\mathbf{U}} \mathrm{Tr}\left(U\varrho U^\dagger H\right) = \min_{U\in\mathbf{U}} \mathrm{Tr}\left(UV\varrho V^\dagger U^\dagger H\right) \tag{3.10}$$

$$= \min_{U\in\mathbf{U}} \mathrm{Tr}\left(U\sum_j r_j|e_j\rangle\langle e_j|U^\dagger \sum_i e_i|e_i\rangle\langle e_i|\right) \tag{3.11}$$

$$= \min_{U\in\mathbf{U}} \mathrm{Tr}\left(\sum_{k,l} u_{kl}|e_k\rangle\langle e_l| \sum_j r_j|e_j\rangle\langle e_j| \sum_{m,n} u^*_{mn}|e_n\rangle\langle e_m| \sum_i e_i|e_i\rangle\langle e_i|\right) \tag{3.12}$$

$$= \sum_{i,j} |u_{ij}|^2 e_i r_j, \tag{3.13}$$

with $U = \sum u_{ij}|i\rangle\langle j|$. Now, consider the matrix S with elements $S_{ij} = |u_{ij}|^2$. Such a matrix with elements $0 \leq S_{ij} \leq 1 \forall i,j$ and $\sum_i S_{ij} = \sum_j S_{ij} = 1 \forall i,j$ is called doubly stochastic matrix and according to Birkhoff and von Neumann's theorem is a convex combination of permutation matrices. Therefore, we can write $|u_{ij}|^2 = \sum_n p_n \sigma^n_{ij}$, with $\sum_n p_n = 1$ and σ^n being permutation matrices. With this, we can write

$$\min_U \mathrm{Tr}\left(UV\varrho V^\dagger U^\dagger H\right) = \min_U \sum_{i,j} |u_{ij}|^2 e_i r_j \tag{3.14}$$

$$= \min_U \sum_n p_n \sum_{i,j} \sigma^n_{ij} e_i r_j \tag{3.15}$$

$$= \min_\sigma \sum_{i,j} \sigma_{ij} e_i r_j \tag{3.16}$$

$$= \sum_i e_i r_i \tag{3.17}$$

We now show the validity of the last equality. Let $\epsilon_j = \sum_i \sigma_{ij} e_i$ and σ be a permutation matrix. If there would exist n, m such that $\epsilon_n - \epsilon_m =: \epsilon > 0$ and $r_n - r_m =: r > 0$ then

$$\sum_i \epsilon_i r_i = \sum_{i\neq n,m} \epsilon_i r_i + \epsilon_n r_n + \epsilon_m r_m \tag{3.18}$$

$$= \sum_{i\neq n,m} \epsilon_i r_i + 2\epsilon_m r_m + r\epsilon_m + \epsilon r_m + r\epsilon \tag{3.19}$$

$$> \sum_{i\neq n,m} \epsilon_i r_i + 2\epsilon_m r_m + r\epsilon_m + \epsilon > r_m. \tag{3.20}$$

$$= \sum_{i\neq n,m} \epsilon_i r_i + r_n\epsilon_m + r_m\epsilon_n \tag{3.21}$$

Consequently, σ cannot be optimal in this case and for the optimal permutation matrix the following must be true for all n, m: If $r_n \geq r_m$ then also $\epsilon_n \leq \epsilon_m$. Since we defined the r_i to be decreasing with their index, the above statement is equivalent to: For all n, m with $n \geq m$ it must hold true that $\epsilon_n \leq \epsilon_m$. With this, we can conclude that $\epsilon_i = e_i \forall i$ and therefore $\sigma^{\mathrm{opt}}_{ij} = \delta_{ij}$, which proves the theorem. □

3.2 Daemonic Ergotropy

Recently, the concept of ergotropy has been extended for systems with an ancilla that can be measured before extracting the work from the system [3]. The idea is that a measurement on the ancilla, which in general will be correlated with the system, will increase the knowledge of the system and thus enhance work extraction. This scenario is reminiscent of Maxwell's demon and the ergotropy of the conditional states of the system – averaged over the measurement outcomes – is called *daemonic ergotropy*.

Definition 3.2 (Daemonic Ergotropy [3])**.** Let ϱ_{SA} be a bipartite density matrix consisting of system S and ancilla A, H be a Hamiltonian acting on S and Π a PVM [1] with elements Π_i acting on A. Then, the daemonic ergotropy is defined as

$$W_D(\varrho_{SA}, H, \Pi) = \mathrm{Tr}\left(\varrho_S H\right) - \sum_i \min_{U_i} \mathrm{Tr}\left(U_i \varrho_{S|i} U_i^\dagger H\right), \tag{3.22}$$

where $\varrho_{S|i} = \mathrm{Tr}_A\left(\varrho_{SA}(\mathbb{1} \otimes \Pi_i)\right)$ is the unnormalised conditional state of the system for outcome i, $p_i = \mathrm{Tr}_A\left(\varrho_{S|i}\right)$ is the probability for this outcome and $\varrho_S = \mathrm{Tr}_A\left(\varrho_{SA}\right)$ is the reduced state of the system. The optimisation runs over all unitary matrices.

As Francica et al. show [3], the daemonic ergotropy of a state ϱ_{SA} is always larger or equal to the ergotropy of the corresponding reduced state $\varrho_S = \mathrm{Tr}_A\left(\varrho_{SA}\right)$. We now show this in a different way by exploiting the subadditivity of ergotropy and writing

$$\begin{aligned} W_D(\varrho_{SA}, H, \Pi) &= \mathrm{Tr}\left(\varrho_S H\right) - \sum_i \min_{U_i} \mathrm{Tr}\left(U_i \varrho_{S|i} U_i^\dagger H\right) && (3.23)\\ &= \sum_i W(\mathrm{Tr}_A\left((\mathbb{1} \otimes \Pi_i)\varrho_{SA}\right), H) && (3.24)\\ &\geq W(\mathrm{Tr}_A(\sum_i (\mathbb{1} \otimes \Pi_i)\varrho_{SA}), H) && (3.25)\\ &= W(\varrho_S, H). && (3.26) \end{aligned}$$

It is natural to ask how much one can gain by measuring the ancilla before work extraction compared to just ignoring the ancilla all together. This quantity is the *daemonic gain*. It is the difference between maximal daemonic ergotropy and the ergotropy for a given state and Hamiltonian.

Definition 3.3 (Maximal Daemonic Ergotropy [3])**.** The maximal daemonic ergotropy is the daemonic ergotropy for the optimal PVM .

$$W_{MD}(\varrho_{SA}, H) = \max_{\Pi} W_D(\varrho_{SA}, H, \Pi) \tag{3.27}$$

[1] The restriction to performing only PVMs on the ancilla and a possible extension of the definition to POVMs is going to be discussed later.

Definition 3.4 (Daemonic Gain [3])**.** The daemonic gain of a density matrix ϱ_{SA} and a Hamiltonian H is defined as the difference between the maximal daemonic ergotropy and the ergotropy of the reduced state $\varrho_S = \mathrm{Tr}_A(\varrho_{SA})$

$$\delta W(\varrho_{SA}, H) = W_{MD}(\varrho_{SA}, H) - W(\varrho_S, H). \tag{3.28}$$

We now summarise some previously established results on daemonic ergotropy [3]. For pure states, the daemonic ergotropy takes its maximum for any rank-one projective measurement, because the conditional states are always pure in this case. Thus, the daemonic ergotropy for a pure state $|\psi\rangle$ is

$$W_D(|\psi\rangle\langle\psi|) = \mathrm{Tr}(\varrho_S H) - \epsilon_0 \tag{3.29}$$

with $\varrho_S = \mathrm{Tr}_A(|\psi\rangle\langle\psi|)$ and a Hamiltonian with non-degenerate lowest energy eigenvalue ϵ_0. Then, the daemonic gain

$$\delta W(|\psi\rangle\langle\psi|) = \min_U \left(\mathrm{Tr}\left(U^\dagger \varrho_S U H\right)\right) - \epsilon_0 \tag{3.30}$$

can only be zero, if the reduced state ϱ_S is pure, which implies that $|\psi\rangle$ is a product state. Obviously, the daemonic gain vanishes for product states, so the daemonic gain vanishes for pure states if and only if they are product states [3].

Investigating connections between daemonic gain and correlations in the state further, Francica et al. prove that vanishing daemonic gain implies vanishing discord $\delta_{A:S}$, where the demon measures system A to enhance work extraction on system S but when calculating discord, system S is the one that is measured. This can be proved by contradiction. Suppose, the discord $\delta_{A:S}$ of a state

$$\varrho_{SA} = \sum c_{klab} |k\rangle\langle l| \otimes |a\rangle\langle b| \tag{3.31}$$

is unequal zero. Then, as we discussed earlier, ϱ_{SA} cannot be a classical-quantum state and thus there must exist indices α, β such that

$$c_{kl\alpha\beta} \neq c_{kl\alpha\beta}\delta_{kl} \tag{3.32}$$

for any basis on the ancilla system. We now want to show that this implies non-vanishing daemonic gain and therefore consider the conditional states corresponding to the projective measurement Π^A with elements $|a\rangle\langle a|$ acting on the second system, A. The daemonic ergotropy will only equal the ergotropy, if all conditional states are diagonal in the same basis for all projective measurements. Therefore, if this is not the case for Π^A, then daemonic ergotropy and ergotropy are not the same and hence the daemonic gain does not vanish. If $\alpha = \beta$, this follows directly from equation (3.32), when choosing the computational basis of the system to be the one in which all conditional states when performing Π^A shall be diagonal. Then, equation (3.32) implies that in this case $\varrho_{S|\alpha} \propto \mathrm{Tr}_A(\varrho_{SA}(\mathbb{1} \otimes |\alpha\rangle\langle\alpha|))$ is not diagonal. The remaining case to discuss is the one

if $\alpha \neq \beta$ and all conditional states for Π^A are simultaneously diagonal. In this case, one can construct another projective measurement $\tilde{\Pi}^A$ with elements

$$\tilde{\Pi}^A_a = \begin{cases} |a\rangle\langle a|, & a \neq \alpha \text{ and } a \neq \beta \\ \frac{1}{2}(|\alpha\rangle + |\beta\rangle)(\text{h.c.}), & a = \alpha \\ \frac{1}{2}(|\alpha\rangle - |\beta\rangle)(\text{h.c.}), & a = \beta \end{cases} \tag{3.33}$$

and show with help of equation (3.32) that the conditional states with respect to this measurement are not simultaneously diagonal [3].

In the two-qubit case, zero daemonic gain implies both zero entanglement and zero discord and maximal discord and maximal entanglement both are sufficient conditions for maximal daemonic gain [3]. An example for states that minimise daemonic gain at a given concurrence C are the ones locally equivalent to

$$\varrho = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & x & C/2 & 0 \\ 0 & C/2 & 1-x & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} \tag{3.34}$$

with $x = (1 \pm \sqrt{1-C^2})/2$. Maximal daemonic gain can be achieved for any concurrence for the states locally equivalent to

$$\varrho = \begin{pmatrix} 1/2 & 0 & 0 & C/2 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ C/2 & 0 & 0 & 1/2 \end{pmatrix}, \tag{3.35}$$

which all have $\delta W = 1$ if the Hamiltonian is chosen $H = -\sigma_z$.

In the next chapter, we will start with a similar investigation of the three-qubit case.

Chapter 4

A Case Study in Three Qubits

In the previous chapter, we discussed the work of Francica et al.[3], who invented daemonic gain and established connections to entanglement and discord. They also did a more detailed analysis of the two qubit case. Very much in the same spirit, we will now investigate how daemonic gain is related to genuine multipartite entanglement for pure, three qubit states, while keeping the bipartite setting of one system and one ancilla. This gives rise to two scenarios in which the ancilla either consists of one or two qubits. We will now first consider the case in which the ancilla consists of a single qubit. We set the Hamiltonian to be $H = \mathrm{diag}\{-3, -1, 1, 3\}$. As a measure for genuine multipartite entanglement we use the geometric average of the bipartite negativities with respect to every bipartition and refer to this quantity simply as negativity [30]. Figure (4.1) shows a plot of daemonic gain versus negativity.

We will now first compute the daemonic gain for pure, three qubit states and a one-qubit ancilla, because this will be helpful for our further analysis.

4.1 Daemonic Gain for Pure, Three Qubit States

Since the states are pure, the daemonic ergotropy takes the simple form (see eq.(3.30))

$$W_D = \mathrm{Tr}\left(\varrho_S H\right) - \epsilon_0, \tag{4.1}$$

where ϵ_0 is the lowest eigenvalue of the Hamiltonian and the daemonic gain becomes

$$\delta W = \min_U \mathrm{Tr}\left(U^\dagger \varrho_S U H\right) - \epsilon_0. \tag{4.2}$$

According to Acin et al. [16], any three qubit state is up to local unitaries equivalent to

$$|k\rangle = k_0 e^{i\theta} |000\rangle + k_1 |001\rangle + k_2 |010\rangle + k_3 |100\rangle + k_4 |111\rangle, \quad k_i \in \mathbb{R}. \tag{4.3}$$

We will from now on omit the complex phase of the first coefficient, as we will later argue that it is irrelevant for our purpose. With this, we calculate the reduced state of

the first two qubits

$$\varrho_S = \begin{pmatrix} k_0^2 + k_1^2 & k_0 k_2 & k_0 k_3 & k_1 k_4 \\ k_0 k_2 & k_2^2 & k_2 k_3 & 0 \\ k_0 k_3 & k_2 k_3 & k_3^2 & 0 \\ k_1 k_4 & 0 & 0 & k_4^2 \end{pmatrix} \tag{4.4}$$

and its eigenvalues

$$\mathrm{eig}\,(\varrho_S) = (\frac{1}{2} + \frac{1}{2}\sqrt{\sigma}, \frac{1}{2} - \frac{1}{2}\sqrt{\sigma}, 0, 0) \tag{4.5}$$

with

$$\sigma = \sum_i k_i^4 + 2k_0^2(k_1^2 + k_2^2 + k_3^2 - k_4^2) \tag{4.6}$$

$$+ 2k_1^2(-k_2^2 - k_3^2 + k_4^2) + 2k_2^2(k_3^2 - k_4^2) - 2k_3^2 k_4^2. \tag{4.7}$$

The daemonic gain is thus

$$\delta W = \left(\frac{1}{2} + \frac{1}{2}\sqrt{\sigma}\right)\epsilon_0 + \left(\frac{1}{2} - \frac{1}{2}\sqrt{\sigma}\right)\epsilon_1 - \epsilon_0 \tag{4.8}$$

$$= \left(\frac{1}{2} - \frac{1}{2}\sqrt{\sigma}\right)(\epsilon_1 - \epsilon_0) \tag{4.9}$$

with a maximum of $\delta W = \frac{1}{2}(\epsilon_1 - \epsilon_0)$.

4.2 Daemonic Gain and Entanglement

In order to relate daemonic gain and entanglement, we will in the following discuss the plot shown in figure (4.1) of daemonic gain versus negativity in the one-qubit ancilla case. We thereby set our focus to understanding the bounds that confine the region in the plot in which states may lie. Especially, we show how to derive an analytic approximation to the lower right boundary, which is the green line in the plot.

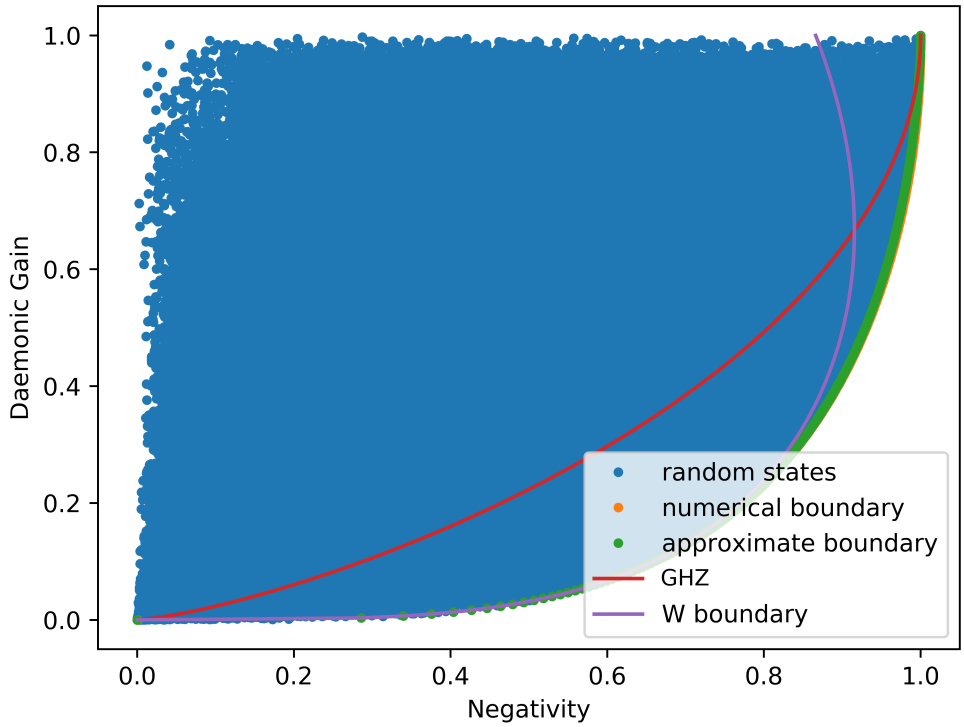

Figure 4.1: Daemonic gain and negativity for pure states. GHZ-like states $|GHZ\rangle = \sqrt{p}\,|000\rangle + \sqrt{1-p}\,|111\rangle$ lie on the red line. W-like states $|W\rangle = a\,|100\rangle + b\,|010\rangle + c\,|001\rangle$ lie in the region that is confined by the purple line to the right. The blue dots refer to random pure states and lie in a region that is bounded by the orange line which is approximated by the green line.

In the following, we discuss the states that lie on the boundaries in the plot in figure 4.1, which we refer to as top boundary, left boundary and right and lower boundary.

Top boundary

As can be seen easily in equation (4.9) states with maximum daemonic gain are the ones with $\sigma = 0$. States that lie on the top boundary are for example

$$|\Psi\rangle = \frac{1}{\sqrt{2}}(\,|0\alpha 1\rangle + |100\rangle), \text{ with} \tag{4.10}$$

$$|\alpha\rangle = \cos\alpha\,|0\rangle + \sin\alpha\,|1\rangle. \tag{4.11}$$

For $\alpha = \frac{\pi}{2}$ the state $|\Psi\rangle$ is maximally entangled, since it is locally equivalent to the GHZ-state, for $\alpha = 0$ the negativity vanishes because the second qubit can be factorised, so

there is no entanglement between the second qubit and the other two. At the same time the reduced density matrix has a doubly degenerate eigenvalue 1/2, which ensures $\sigma = 0$ and maximal daemonic gain.

Left boundary

States that lie on the line defined by the negativity being zero are for example

$$|\Psi\rangle = \sqrt{p}\,|001\rangle + \sqrt{1-p}\,|100\rangle. \tag{4.12}$$

The second qubit can be factorised which ensures the negativity to be zero. At the same time, the reduced density matrix has two eigenvalues p and $1-p$, so for $p \leq \frac{1}{2}$ the daemonic gain is $\delta W = p(\epsilon_1 - \epsilon_0)$, reaching its maximum for $p = \frac{1}{2}$.

Right and lower boundary

Finding an analytic solution for the right, lower boundary in the plot in figure (4.1) by calculating the negativity for general pure states (equation 4.3) and optimising over the parameters shows to be impractical. Therefore, instead of using negativity as entanglement measure, we use the average entanglement of formation with respect to all bipartitions. For this, we compute the Schmidt coefficients for every bipartition of a state. For a tripartite state we gain three sets of Schmidt coefficients $\{\lambda_i^A\}, \{\lambda_i^B\}, \{\lambda_i^C\}$. Then the average entanglement of formation is

$$E = -\frac{1}{3}\Big(\sum_i \lambda_i^A \log(\lambda_i^A) + \sum_j \lambda_j^B \log(\lambda_j^B) + \sum_k \lambda_k^C \log(\lambda_k^C)\Big). \tag{4.13}$$

This entanglement measure is easier to compute for pure states and satisfies the requirement that states which maximise negativity at a given daemonic gain will at least approximately maximise this entanglement measure, too. This is demonstrated in figure (4.2) and non-trivial as the used version of negativity is a entanglement measure that captures genuine multipartite entanglement only while the average entanglement of formation is also sensitive to bipartite entanglement.

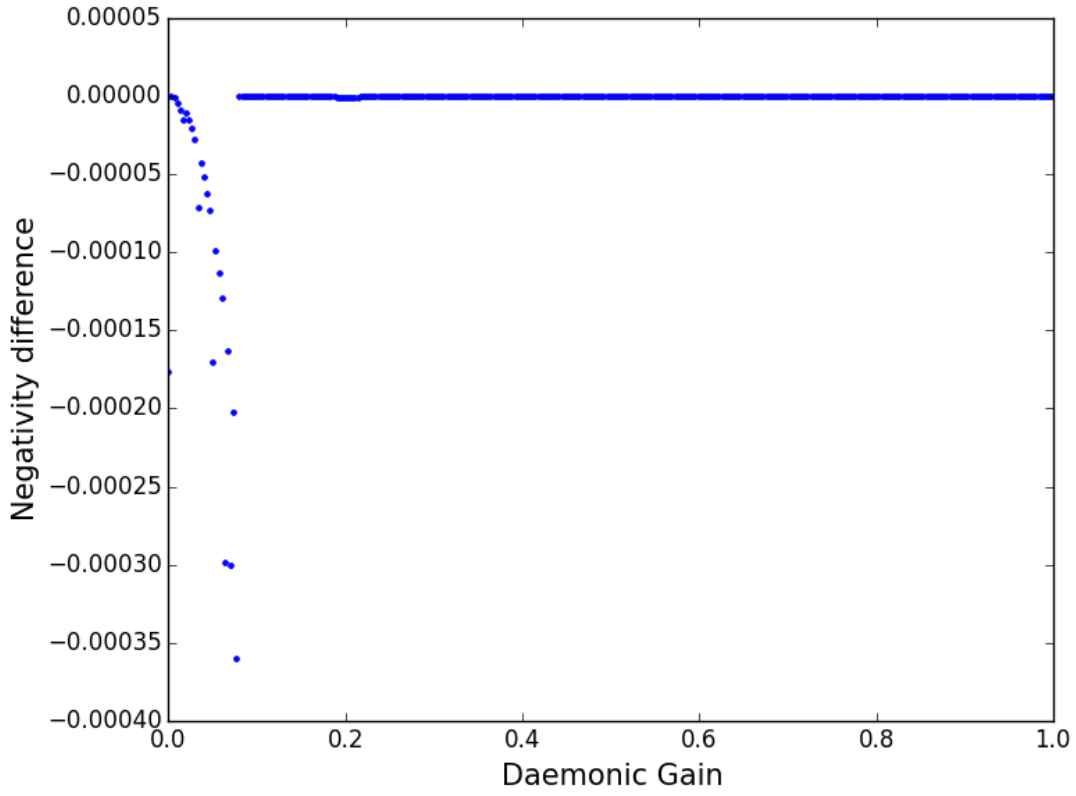

Figure 4.2: Difference between the maximal negativity at a given daemonic gain and the negativity that is exhibited by the states, which maximise the average entanglement of formation at given daemonic gain. By construction, this difference must be non-negative. The deviations from zero in the plot are due to imperfect numerics. The plot suggests, that up to corrections beyond numerical precision the same states that maximise negativity also maximise average entanglement of formation.

The Schmidt coefficients for the three qubit states in the symmetric canonical form (equation 4.3) are

$$\lambda^A_{1,2} = \frac{1}{2} \pm \sqrt{\frac{1}{4} - k_4^2(k_0^2 + k_1^2 + k_2^2) - k_3^2(k_1^2 + k_2^2)} \tag{4.14}$$

$$\lambda^B_{1,2} = \frac{1}{2} \pm \sqrt{\frac{1}{4} - k_4^2(k_0^2 + k_1^2 + k_3^2) - k_2^2(k_1^2 + k_3^2)} \tag{4.15}$$

$$\lambda^C_{1,2} = \frac{1}{2} \pm \sqrt{\frac{1}{4} - k_4^2(k_0^2 + k_2^2 + k_3^2) - k_1^2(k_2^2 + k_3^2)}. \tag{4.16}$$

From the formulas of the average entanglement of formation and the daemonic gain one can see that both are symmetric in the parameters k_2 and k_3. We now want to maximise the entanglement while satisfying two constraints: The state must be normalised and

the daemonic gain must remain constant. Instead of dealing with the complete formula for daemonic gain, it is sufficient to demand σ to be constant. Numerics show, that the boundary can still be reached even if k_0 is set to zero. This vindicates our choice of omitting the complex phase in equation (4.3). If we solve this problem with Lagrange multipliers, this yields the following set of equations

$$k_1^2 + 2k_2^2 + k_4^2 = 1 \text{ [(I) (Normalisation)]} \tag{4.17}$$

$$(k_1^2 + 2k_2^2 - k_4^2)^2 - 4k_1^2(2k_2^2 - k_4^2) = \sigma \text{ [(II) } (\delta W \text{ constant)]} \tag{4.18}$$

$$\partial_i E - \lambda_1 \partial_i(\text{I}) - \lambda_2 \partial_i(\text{II}) = 0, \tag{4.19}$$

where we have set

$$k_2 = k_3 \tag{4.20}$$

and

$$\partial_i = \frac{\partial}{\partial(k_i^2)}. \tag{4.21}$$

We can set $k_2 = k_3$ because both the normalisation constraint and the $\delta W = const.$ constraint are symmetric in both parameters and also the local maximum condition from equation (4.19) is the same for both parameters. In principle, $k_2 = k_3$ could also be a local extremum, but it turns out not to be one. Since the average entanglement of formation still seems to be too difficult to find a parametrisation for the states that maximise negativity, we replace the logarithm by its Taylor expansion up to first order at $1/2$. This, together with the condition $k_0 = 0$ and $k_2 = k_3$ yields the much simplified and approximated formula for the average entropy of formation

$$E = k_2^4 + 2k_2^2k_1^2 + 2k_2^2k_4^2 + k_1^2k_4^2. \tag{4.22}$$

Solving the equations then gives

$$k_0^2 = 0 \tag{4.23}$$

$$k_1^2 = 1 - 2k_2^2 - k_4^2 \tag{4.24}$$

$$k_2^2 = \frac{1}{4}(1 \pm \sqrt{\sigma}) \tag{4.25}$$

$$k_3^2 = k_2^2 \tag{4.26}$$

$$k_4^2 = \frac{1 + 5k_2^4 - 5k_2^2}{2 - 7k_2^2}. \tag{4.27}$$

If one plots the two solutions that arise from the fact that there is an ambiguity in the solution for k_2^2, one easily sees that only states with

$$k_2^2 = \frac{1}{4}(1 - \sqrt{\sigma}) = \frac{\delta W}{2\Delta\epsilon} \tag{4.28}$$

lie on the boundary. Plotting the daemonic gain against the negativity of these states results in the green line in figure (4.1).

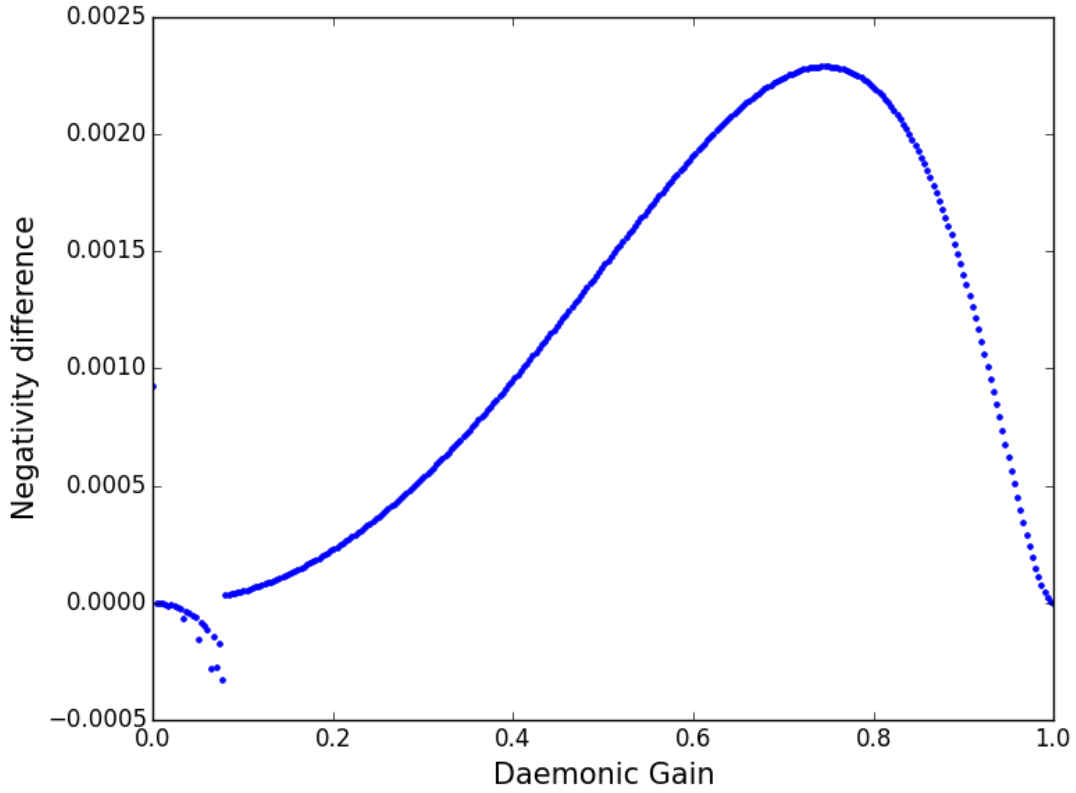

Figure 4.3: Difference between the negativities of numerical solution for the boundary and the approximated analytic solution

Figure (4.3) shows the deviation of the negativity exhibited by the states with parameters given in equations (4.23– 4.27) from the numerically found maximal negativity at given daemonic gain. This is just the distance of the green line from the orange one in figure (4.1). Again, due to numerical imperfections, the analytically found states show greater negativity for small daemonic gain. In contrast, for bigger daemonic gain the error being made by approximating the logarithm linearly causes the analytically found states to exhibit up to roughly 0.0023 less negativity than the numerically found optimal ones.

The Two Qubit Ancilla Case

Now that we discussed the three qubit case with a one-qubit ancilla, we discuss the case in which S only consists of one qubit and the ancilla A consists of two qubits. Again,

we demand $k_0^2 = 0$. Calculating the daemonic gain in the most general case then shows that this time we do not have a symmetry between k_3^2 and k_2^2 as in the one-qubit ancilla case but instead one between k_1^2 and k_2^2. Using this symmetry, both daemonic gain and entanglement of formation have the same formulas as they did in the one qubit case except the parameters k_1 and k_3 are swapped. We can therefore easily reuse the solution of the lower right boundary in the plot in the one-qubit ancilla case for the two-qubit case:

$$k_0^2 = 0 \tag{4.29}$$
$$k_1^2 = k_2^2 \tag{4.30}$$
$$k_2^2 = \frac{1}{4}(1 \pm \sqrt{\sigma}) \tag{4.31}$$
$$k_3^2 = 1 - 2k_2^2 - k_4^2 \tag{4.32}$$
$$k_4^2 = \frac{1 + 5k_2^4 - 5k_2^2}{2 - 7k_2^2}. \tag{4.33}$$

Note however, that this transition between one-ancilla and two-ancilla case by swapping k_1 and k_3 is only possible if $k_0 = 0$.

Daemonic gain and the violation of Mermin's and Svetlichny's inequality

In the previous section we saw in figure (4.2) that exchanging negativity as measure of entanglement for average entanglement of formation does not affect, which states maximise entanglement at a given daemonic gain. We now check, whether the same states also maximally violate Bell inequalities.

Similarly as with negativity, there will not be any violation of Svetlichny's inequality as soon as there is one qubit that is classically correlated to the others. In contrast, Mermin's inequality can already be violated if there is just one pair of qubits that share non-classical correlations. In this aspect, the behaviour of Mermin's inequality is similar to using average entanglement of formation as in equation (4.13) as entanglement criterion.

The plot in figure (4.4) shows the daemonic gain and violation of Mermin's and Svetlichny's inequality for random pure states as well as for the states we found to maximise negativity at given daemonic gain. One can easily see, that these states also maximise the violation of both Bell inequalities at given daemonic gain.

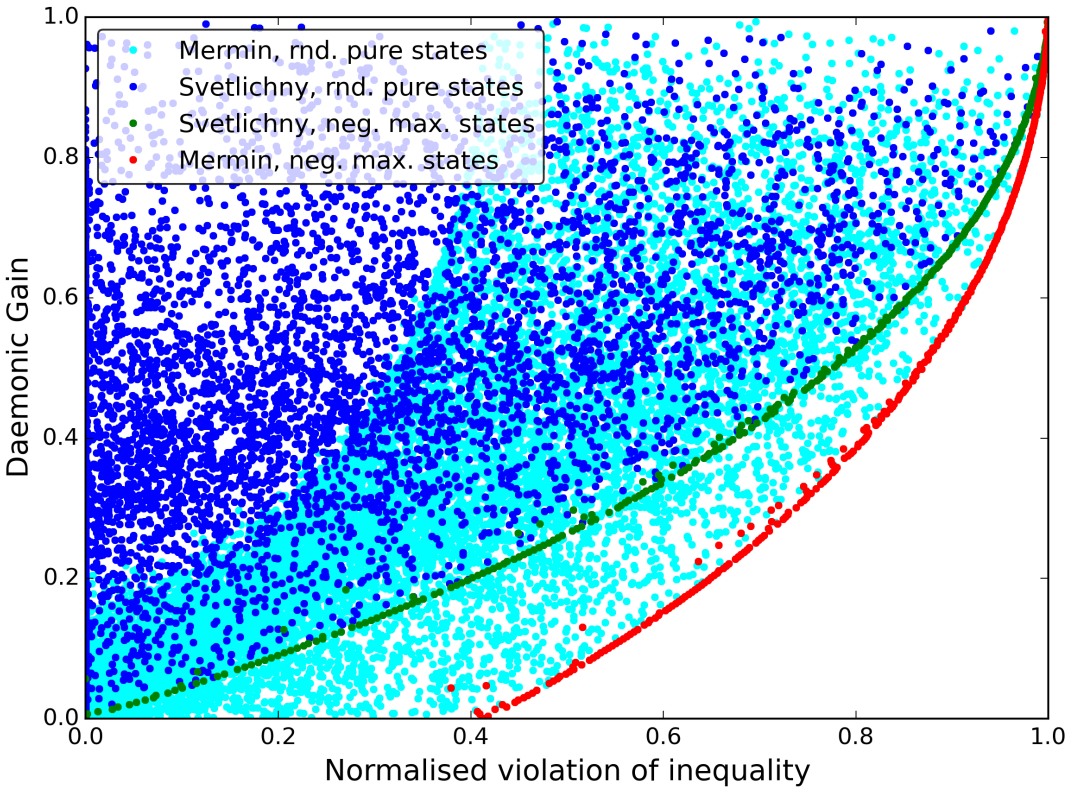

Figure 4.4: Normalised violation of Mermin's and Svetlichny's inequalities for random pure states and the states that maximise negativity at given daemonic gain. Due to faulty numerics, the true maximal violation of the Bell-inequalities is sometimes not found.

Chapter 5

Extending and Computing Daemonic Gain

In this chapter, we first extend the definition of daemonic ergotropy allowing for generalised measurements and then inquire into the problem of finding the optimal measurement that maximises daemonic ergotropy. Such a generalisation is sensible, as this can increase the amount of extractable energy, as we will see later. From now on, we will therefore always consider all kinds of POVMs instead of only projective measurements.

In the second chapter, we then suggest a genuine multipartite generalisation of daemonic gain and transfer some of our findings from the bipartite to the multipartite case.

5.1 Finding the Optimal Measurement

The optimisation over all possible measurements makes the computation of the daemonic ergotropy and daemonic gain difficult by hand and computationally costly on a computer. This is even more true, if we do not only consider projective measurements but also generalised measurements. We therefore try to reach a better understanding of this problem.

5.1.1 Two Helpful Observations

In a previous chapter, we showed that ergotropy is sublinear in both arguments (equations 3.6 and 3.7). Since one can write the daemonic ergotropy of a state ϱ and a measurement represented by a POVM P as

$$W_D(\varrho_{SA}, P) = \sum_i W\left(\mathrm{Tr}_A\left(\varrho_{SA}(\mathbb{1} \otimes P_i)\right)\right) \tag{5.1}$$

this has two immediate implications. First of all, this ensures that one can find an optimal POVM that has rank-one elements exclusively. If an element E_i of an optimal POVM E is not rank-one, then this element can split up into several non-negative rank-one effects $\tilde{E}_i^n = p_i^n \Pi_i^n$ with $E_i = \sum_n p_i^n \Pi_i^n$. Because of the sublinearity of ergotropy,

the daemonic ergotropy using the POVM $\tilde{E}$ with elements $\tilde{E}_i^n$ as defined above

$$W_D(\varrho_{SA}, \tilde{E}) = \sum_{i,n} W\left(\mathrm{Tr}_A\left(\varrho_{SA}(\mathbb{1} \otimes \tilde{E}_i^n)\right)\right) \tag{5.2}$$

$$\geq \sum_i W(\mathrm{tr}_A(\varrho_{SA}(\mathbb{1} \otimes \sum_n \tilde{E}_i^n))) \tag{5.3}$$

$$= W_D(\varrho_{SA}, E) \tag{5.4}$$

will still be optimal.

The second implication is, that daemonic ergotropy is convex in the performed measurement. This means, that given two POVMs P and Q with the same number of elements denoted by E_i and Q_j, the daemonic ergotropy for any convex combination $R = \lambda P + (1-\lambda)Q$ with elements $R_i = \lambda P_i + (1-\lambda)Q_i$ will never exceed the daemonic ergotropy for the POVMs P or Q. We directly show this by writing

$$W_D(\varrho_{SA}, \lambda P + (1-\lambda)Q) = \sum_i W(\mathrm{Tr}_A\left([\mathbb{1} \otimes (\lambda P_i + (1-\lambda)Q_i)]\varrho_{SA}\right) \tag{5.5}$$

$$\leq \lambda \sum_i W(\mathrm{Tr}_A\left([\mathbb{1} \otimes P_i]\varrho_{SA}\right) + (1-\lambda)\sum_i W(\mathrm{Tr}_A\left([\mathbb{1} \otimes Q_i]\varrho_{SA}\right) \tag{5.6}$$

$$= \lambda W_D(\varrho_{SA}, P) + (1-\lambda)W_D(\varrho_{SA}, Q). \tag{5.7}$$

This allows us to restrict our search of optimal measurements to extremal POVMs.

On a system of dimension d, any extremal POVM can at most have d^2 outcomes [8], so we can refine our search to extremal POVMs with d_A^2 outcomes, if d_A is the dimension of the ancilla. The second observation will now become useful, when we present an efficient way of finding an optimal POVM.

5.1.2 A See-Saw Algorithm to Compute Daemonic Gain

When calculating the daemonic gain

$$\delta W(\varrho_{SA}) = \min_V \mathrm{Tr}\left(\varrho_S V H V^\dagger\right) - \min_{\{E_i\}} \sum_i \min_{\{U_i\}} \mathrm{Tr}\left((U_i H U_i^\dagger \otimes E_i)\varrho_{SA}\right), \tag{5.8}$$

we already discussed that evaluating the first term comes down to diagonalising the state and the Hamiltonian. In the second term however, we are faced with two optimisations at a time. We can deal with this problem by defining

$$M := \sum_i \mathrm{Tr}\left((U_i H U_i^\dagger \otimes E_i)\varrho_{SA}\right) \tag{5.9}$$

and then alternately optimising the measurement and the unitaries in a see-saw algorithm [31]. Therefore, we first fix the number of outcomes. If we want to find an optimal POVM, we set the number of outcomes to d_A^2. Now, we initialise the unitaries U_i, for example by generating random unitaries. Then we calculate

$$\sigma_i = \mathrm{Tr}_S\left((U_i H U_i^\dagger \otimes \mathbb{1})\varrho_{SA}\right), \tag{5.10}$$

which allows us to write

$$M = \sum_i \mathrm{Tr}\left(\sigma_i E_i\right), \tag{5.11}$$

and then find the optimal POVM. This is a semidefinite programme and can thus be efficiently solved [1].

For the found POVM, one can then easily compute the conditional states

$$\varrho_{S|i} = \mathrm{Tr}_A\left((\mathbb{1} \otimes E_i)\varrho_{SA}\right) / \mathrm{Tr}\left((\mathbb{1} \otimes E_i)\varrho_{SA}\right) \tag{5.12}$$

and afterwards the optimal unitaries by diagonalising them in the diagonal basis of the Hamiltonian.

These two steps are then alternately carried out a sufficiently high number of times or until some desired precision is reached.

M can only decrease or remain constant in every iteration, since it decreases or remains constant in every step – in the minimisation over POVMs in step one and also in the minimisation over the unitaries in step two. In this way, the see-saw algorithm defines a monotonically decreasing series in M for any number of outcomes. Since the possible values for M are bounded, the algorithm must converge for any number of outcomes. However, it may converge to a non-optimal solution. For example, every time two elements E_i, E_j of the POVM are identical or linearly dependent, the algorithm will find two identical unitaries $U_i = U_j$ which will in turn cause two linearly dependent effects $E'_i \propto E'_j$. In this scenario, the number of outcomes is effectively reduced by one. We already argued that the algorithm must still converge in this case. However, the optimal solution can in general not be found anymore. In practise it happens quite often that the algorithm finds the optimal solution for projective measurements, even if a better POVM exists.

5.1.3 Optimal Measurement for Quantum-Classical States

For a quantum-classical state

$$\varrho_{SA} = \sum_j \sigma_j \otimes |j\rangle\langle j| \tag{5.13}$$

one measurement that maximises daemonic gain is the projective measurement P with elements $P_i = |i\rangle\langle i|$. We show that the daemonic gain is lower or equal for any POVM E with elements E_i and an arbitrary number of outcomes. As we already discussed above, we may without loss of generality assume that all effects E_i are rank one. This allows us to write $E_i = |\phi_i\rangle\langle\phi_i|$ and we can expand the unnormalised vectors $|\phi_i\rangle$ in the computational basis as

$$|\phi_i\rangle = \sum_k u_{ki} |k\rangle. \tag{5.14}$$

[1] In fact, M is similar to the success probability in minimum error discrimination (equation 2.52). In minimum error discrimination however, this quantity would be maximised instead of minimised, which makes this step in the algorithm equivalent to maximum error discrimination.

Since the elements of the POVM sum up to identity, we have

$$\sum_i u_{ki} u_{li}^* = \delta_{kl}. \tag{5.15}$$

Due to the basis extension theorem, one can add rows to the matrix U with elements u_{ki} to make it a unitary matrix. This means that the matrix D with elements $D_{ji} = u_{ji}u_{ji}^*$ can be interpreted as part of an extended doubly stochastic matrix. The theorem of Birkhoff and von Neumann [29] states that any doubly stochastic matrix is a convex combination of permutation matrices. With this, we can write

$$D_{ji} = u_{ji}u_{ji}^* = \sum_n p_n \pi_{ji}^n \tag{5.16}$$

with $\sum_i \pi_{ji}^n = 1$, $\sum_j \pi_{ji}^n \in \{0,1\}$ and $\pi_{ji}^n \in \{0,1\}$.

In order to calculate the daemonic gain, we first compute the unnormalised conditional states

$$\begin{aligned}
\varrho_{S|i} &= \mathrm{Tr}_A \left(\sum_j \sigma_j \otimes |j\rangle\langle j| (\mathbb{1} \otimes E_i) \right) && (5.17)\\
&= \mathrm{Tr}_A \left(\sum_j \sigma_j \otimes u_{ki} u_{li}^* |k\rangle\langle l||j\rangle\langle j| \right) && (5.18)\\
&= \sum_j \sigma_j u_{ji} u_{ji}^* && (5.19)\\
&= \sum_j \sigma_j \sum_n p_n \pi_{ji}^n. && (5.20)
\end{aligned}$$

With this, we can calculate the daemonic gain

$$\delta W(\varrho_{SA}) = \min_V \mathrm{Tr}\left(\varrho_S V H V^\dagger\right) - \min_{\{E_i\}} \sum_i \min_{\{U_i\}} \mathrm{Tr}\left(U_i^\dagger \varrho_{S|i} U_i H\right) \tag{5.21}$$

Again, we focus only on the second term, since the first term does not depend on the performed measurement. We now assume that our POVM E is the optimal one, so we can omit the optimisation over the measurement and plug in the formula for $\varrho_{S|i}$ that we derived above (equation 5.20).

$$\begin{aligned}
\sum_i \min_{\{U_i\}} \mathrm{Tr}\left(U_i^\dagger \varrho_{S|i} U_i H\right) &= \sum_i \min_{\{U_i\}} \mathrm{Tr}\left(U_i^\dagger \sum_j \sigma_j \sum_n p_n \pi_{ji}^n U_i H \right) && (5.22)\\
&\geq \sum_{i,j,n} p_n \pi_{ji}^n \min_{\{U_{ijn}\}} \mathrm{Tr}\left(U_{ijn}^\dagger \sigma_j U_{ijn} H\right) && (5.23)\\
&= \sum_n p_n \sum_i \pi_{ji}^n \sum_j \min_{\{U_j\}} \mathrm{Tr}\left(U_j^\dagger \sigma_j U_j H\right) && (5.24)\\
&= \sum_j \min_{\{U_j\}} \mathrm{Tr}\left(U_j^\dagger \sigma_j U_j H\right). && (5.25)
\end{aligned}$$

This is exactly what one obtains if one chooses the projective measurement P with elements $P_i = |i\rangle\langle i|$, so this measurement is indeed optimal. In the first step, we obtain the inequality by allowing for a different unitary for every combination of indices. One then notices in the second step, that the optimal unitary is independent of the indices i and n which yields the desired result.

If a state $\varrho_{SA} = \sum_{ij} \sigma_{ij} \otimes |i\rangle\langle j|$ is quantum-classical in some basis, we can thus find an analytic solution to the daemonic gain. To this end, one diagonalises the reduced state $\varrho_S = \mathrm{Tr}_A(\varrho_{SA})$, which yields the unitary to make the state block-diagonal. The individual blocks are then the optimal conditional states that one needs to compute the daemonic gain.

5.1.4 Daemonic Ergotropy and Quantum Correlations

The above result links the question for the optimal measurement to the question whether quantum correlations can be a resource for daemonic ergotropy. If for a state ϱ_{SA} the optimal measurement is projective, with projectors $|i\rangle\langle i|$ and we write the state as

$$\varrho_{SA} = \sum_{ij} \sigma_{ij} \otimes |i\rangle\langle j| \tag{5.26}$$

where the ancilla part is written in the basis defined by the optimal projectors, we notice that all off-block-diagonal terms $\sigma_{ij}, i \neq j$ do not contribute to the daemonic ergotropy, which is thus the same as for the quantum-classical state

$$\tilde{\varrho}_{SA} = \sum_{i} \sigma_{ii} \otimes |i\rangle\langle i|. \tag{5.27}$$

This state can be produced by performing the optimal measurement and preparing a pure state on the ancilla accordingly. This procedure destroys all the entanglement and discord $\delta_{S:A}$ while the daemonic ergotropy remains unchanged. Entanglement and discord $\delta_{S:A}$ are thus not useful, if the optimal measurement is projective.

The following example also questions the usefulness of entanglement in this context. Consider the classical-classical state with a qutrit system and a qubit ancilla

$$\varrho = \frac{1}{3}(|0\rangle\langle 0| \otimes |0\rangle\langle 0| + |1\rangle\langle 1| \otimes |1\rangle\langle 1| + |2\rangle\langle 2| \otimes |1\rangle\langle 1|). \tag{5.28}$$

For a Hamiltonian with eigenvalues $\epsilon_0 \leq \epsilon_1 \leq \epsilon_2$ one easily calculates the daemonic gain

$$\delta W(\varrho) = \frac{1}{3}(\epsilon_2 - \epsilon_0). \tag{5.29}$$

On the other hand, entanglement is maximised for pure states and for any pure state we have

$$\delta W(|\Psi\rangle) \leq \frac{1}{2}(\epsilon_1 - \epsilon_0) \tag{5.30}$$

which is smaller than the daemonic gain of ϱ for a suitably chosen Hamiltonian.

In the next section, we will see that projective measurements are not always optimal, when we calculate the daemonic gain of a classical quantum state.

5.1.5 Enhancing Daemonic Gain with Generalised Measurements

In the original definition of daemonic gain [3] the possible measurements to be performed on the ancilla are restricted to projective measurements. However, we are going to provide an example where the daemonic gain that can be achieved when using an exemplary POVM is higher than it would be with any projective measurement. The state used in this example is the 3×2 classical-quantum state

$$\varrho_{SA} = \frac{1}{3} \sum_{i=0}^{2} |i\rangle\langle i| \otimes P_i \tag{5.31}$$

with

$$P_i = \Pi\left(\frac{2\pi i}{3}, 0\right) \tag{5.32}$$

and

$$\Pi(\alpha, \beta) = \frac{1}{2} \begin{pmatrix} 1 - \cos(\alpha) & \sin(\alpha) \cdot \exp(i\beta) \\ \sin(\alpha) \cdot \exp(-i\beta) & 1 + \cos(\alpha) \end{pmatrix}. \tag{5.33}$$

The Hamiltonian is $H = \sum_i \epsilon_i |\epsilon_i\rangle\langle\epsilon_i|$ with increasing energy eigenvalues. We shall now first compute the daemonic gain, restricting ourselves to projective measurements. Since daemonic gain is convex, the effects of any POVM are either all rank one or can be split up into rank one effects without changing the daemonic gain. Therefore, we can compute the maximal daemonic gain for projective measurements by computing it for the measurement $\mathbf{\Pi} = (\Pi(\alpha, \beta), \Pi(\alpha + \pi, \beta))$ and optimize over the angles α and β

afterwards.

$$\varrho_S = \frac{1}{3}(|0\rangle\langle 0| + |1\rangle\langle 1| + |2\rangle\langle 2|) \tag{5.34}$$

$$p_\alpha \varrho_{S|\alpha} = \mathrm{Tr}\left(\varrho_{SA}(\mathbb{1} \otimes \Pi(\alpha, \beta))\right) \tag{5.35}$$

$$= \frac{1}{3}(|0\rangle\langle 0|\mathrm{Tr}\left(P_0\Pi(\alpha,\beta)\right) + |1\rangle\langle 1|\mathrm{Tr}\left(P_1\Pi(\alpha,\beta)\right) + |2\rangle\langle 2|\mathrm{Tr}\left(P_2\Pi(\alpha,\beta)\right) \tag{5.36}$$

$$\begin{aligned} = \frac{1}{3}\Bigg[& |0\rangle\langle 0|\frac{1}{2}(1+\cos(\alpha)) \\ & + |1\rangle\langle 1|\left(\frac{1}{2} - \frac{1}{4}\cos(\alpha) + \frac{\sqrt{3}}{4}\sin(\alpha)\cos(\beta)\right) \\ & + |2\rangle\langle 2|\left(\frac{1}{2} - \frac{1}{4}\cos(\alpha) - \frac{\sqrt{3}}{4}\sin(\alpha)\cos(\beta)\right)\Bigg] \end{aligned} \tag{5.37}$$

$$p_{\alpha+\pi}\varrho_{S|\alpha+\pi} = \mathrm{Tr}\left(\varrho_{SA}(\mathbb{1} \otimes \Pi(\alpha+\pi, \beta))\right) \tag{5.38}$$

$$\begin{aligned} = \frac{1}{3}(& |0\rangle\langle 0|\mathrm{Tr}\left(P_0\Pi(\alpha+\pi,\beta)\right) \\ & + |1\rangle\langle 1|\mathrm{Tr}\left(P_1\Pi(\alpha+\pi,\beta)\right) \\ & + |2\rangle\langle 2|\mathrm{Tr}\left(P_2\Pi(\alpha+\pi,\beta)\right) \end{aligned} \tag{5.39}$$

$$\begin{aligned} = \frac{1}{3}\Bigg[& |0\rangle\langle 0|\frac{1}{2}(1-\cos(\alpha)) \\ & + |1\rangle\langle 1|\left(\frac{1}{2} + \frac{1}{4}\cos(\alpha) - \frac{\sqrt{3}}{4}\sin(\alpha)\cos(\beta)\right) \\ & + |2\rangle\langle 2|\left(\frac{1}{2} + \frac{1}{4}\cos(\alpha) + \frac{\sqrt{3}}{4}\sin(\alpha)\cos(\beta)\right)\Bigg] \end{aligned} \tag{5.40}$$

From the definition of ergotropy it is already obvious that the ergotropy of the conditional states $\varrho_{S|\alpha}$ and $\varrho_{S|\alpha+\pi}$ will be maximal for $\cos(\beta) = 1$. Consider the state

$$\varrho = a|0\rangle\langle 0| + (b+c)|1\rangle\langle 1| + (b-c)|2\rangle\langle 2| \tag{5.41}$$

where $a, b, c \in \mathbb{R}$ and $c \geq 0$. Let the Hamiltonian be

$$H = e_0|e_0\rangle\langle e_0| + e_1|e_1\rangle\langle e_1| + e_2|e_2\rangle\langle e_2| \tag{5.42}$$

without demanding any particular ordering of the energy eigenvalues. Then, the ergotropy without loss of generality can be written as

$$W = \mathrm{Tr}\left(\varrho H\right) - \min_U \mathrm{Tr}\left(U\varrho U^\dagger H\right) = \mathrm{Tr}\left(\varrho H\right) - (e_0 a + e_1(b+c) + e_2(b-c)) \tag{5.43}$$

$$= \mathrm{Tr}\left(\varrho H\right) - (ae_0 + b(e_1+e_2) + c(e_1-e_2)). \tag{5.44}$$

Note, that $e_1 \leq e_2$ since $(b+c) \geq (b-c)$. Consequently, W increases with c and we can set $\beta = 0$ in the above calculation. Exploiting addition theorems, we can now write

$$p_\alpha \varrho_{S|\alpha} = \frac{1}{6}\left[|0\rangle\langle 0|(1+\cos\alpha) + |1\rangle\langle 1|\left(1+\cos\left(\alpha-\frac{2\pi}{3}\right)\right) + |2\rangle\langle 2|\left(1+\cos\left(\alpha+\frac{2\pi}{3}\right)\right)\right] \tag{5.45}$$

$$p_{\alpha+\pi} \varrho_{S|\alpha+\pi} = \frac{1}{6}\left[|0\rangle\langle 0|(1-\cos\alpha) + |1\rangle\langle 1|\left(1-\cos\left(\alpha-\frac{2\pi}{3}\right)\right) + |2\rangle\langle 2|\left(1-\cos\left(\alpha+\frac{2\pi}{3}\right)\right)\right]. \tag{5.46}$$

As one can easily see, an optimal value of α is not unique, as shifting its value by $\frac{2\pi}{3}$ can be compensated by relabeling the states, which does not affect the daemonic gain. We now aim to find the optimal α in the interval $[-\frac{\pi}{3}, \frac{\pi}{3})$. When calculating the ergotropy of the conditional states we need to know the ordering of their eigenvalues.

$$\alpha \in \left[-\frac{\pi}{3}, 0\right) \Rightarrow \cos\alpha \geq \cos\left(\alpha+\frac{2\pi}{3}\right) \geq \cos\left(\alpha-\frac{2\pi}{3}\right) \tag{5.47}$$

$$\alpha \in \left(0, \frac{\pi}{3}\right) \Rightarrow \cos\alpha \geq \cos\left(\alpha-\frac{2\pi}{3}\right) \geq \cos\left(\alpha+\frac{2\pi}{3}\right) \tag{5.48}$$

In the following calculation, the upper sign will refer to the negative and the lower sign will refer to the positive interval.

$$\delta W(\varrho_{SA}, H, \mathbf{\Pi}) = W_D(\varrho_{SA}, H, \mathbf{\Pi}) - W(\varrho_S, H) \tag{5.49}$$

$$= \mathrm{Tr}\left(\varrho_S H\right) - \min_{\mathbf{\Pi}} \sum_k \mathrm{Tr}\left(\varrho_S A(U_k^\dagger H U_k \otimes \Pi_k)\right) - \left[\mathrm{Tr}\left(\varrho_S H\right) - \min_U \mathrm{Tr}\left(\varrho_S U^\dagger H U\right)\right] \tag{5.50}$$

$$= \min_U \mathrm{Tr}\left(\varrho_S U^\dagger H U\right) - \min_{\mathbf{\Pi}} \sum_k \mathrm{Tr}\left(\varrho_S A(U_k^\dagger H U_k \otimes \Pi_k)\right) \tag{5.51}$$

$$\begin{aligned} &= \max_\alpha \{\ \frac{1}{3}(\epsilon_0+\epsilon_1+\epsilon_2) - \frac{1}{6}(\epsilon_0(1+\cos\alpha) \\ &\quad + \epsilon_1\left(1+\cos\left(\alpha \pm \frac{2\pi}{3}\right)\right) \\ &\quad + \epsilon_2\left(1+\cos\left(\alpha \mp \frac{2\pi}{3}\right)\right) \\ &\quad + \epsilon_0\left(1-\cos\left(\alpha \mp \frac{2\pi}{3}\right)\right) \\ &\quad + \epsilon_1\left(1-\cos\left(\alpha \pm \frac{2\pi}{3}\right)\right) \\ &\quad + \epsilon_2(1-\cos\alpha)\ \} \end{aligned} \tag{5.52}$$

$$= \frac{1}{6}(\epsilon_2-\epsilon_0)\max_\alpha\left(\cos\alpha - \cos\left(\alpha \mp \frac{2\pi}{3}\right)\right) = \frac{\epsilon_2-\epsilon_0}{2\sqrt{3}} \tag{5.53}$$

Now, that we computed the maximal daemonic gain for projective measurements, we compare this with the daemonic gain that can be achieved by using the POVM $\mathbf{P}$,

consisting of the effects $\frac{2}{3}P_i$, as defined in 5.32.

$$p_0\varrho_{S|P_0} = \frac{1}{3}\frac{2}{3}\left(|0\rangle\langle 0| + \frac{1}{4}|1\rangle\langle 1| + \frac{1}{4}|2\rangle\langle 2|\right) \tag{5.54}$$

$$p_1\varrho_{S|P_1} = \frac{1}{3}\frac{2}{3}\left(\frac{1}{4}|0\rangle\langle 0| + |1\rangle\langle 1| + \frac{1}{4}|2\rangle\langle 2|\right) \tag{5.55}$$

$$p_2\varrho_{S|P_2} = \frac{1}{3}\frac{2}{3}\left(\frac{1}{4}|0\rangle\langle 0| + \frac{1}{4}|1\rangle\langle 1| + |2\rangle\langle 2|\right) \tag{5.56}$$

Given the conditional states, we can now compute the daemonic gain.

$$\delta W = \epsilon_0(\frac{1}{3} - \frac{2}{3}) + \epsilon_1(\frac{1}{3} - \frac{1}{6}) + \epsilon_2(\frac{1}{3} - 16) \tag{5.57}$$

$$= -\frac{1}{3}\epsilon_0 + \frac{1}{6}(\epsilon_1 + \epsilon_2) \tag{5.58}$$

Choosing a Hamiltonian $H = |\epsilon_1\rangle\langle\epsilon_1| + |\epsilon_2\rangle\langle\epsilon_2|$ provides an example where the maximal daemonic gain can not be achieved by using projective measurements because

$$\delta W_{\text{proj}} = \frac{1}{2\sqrt{3}} < \delta W_{\mathbf{P}} = \frac{1}{3}. \tag{5.59}$$

The following plot shows the daemonic gain for the state considered here as defined in equation 5.31 and the Hamiltonian $H = |1\rangle\langle 1| + \epsilon_2|2\rangle\langle 2|$.

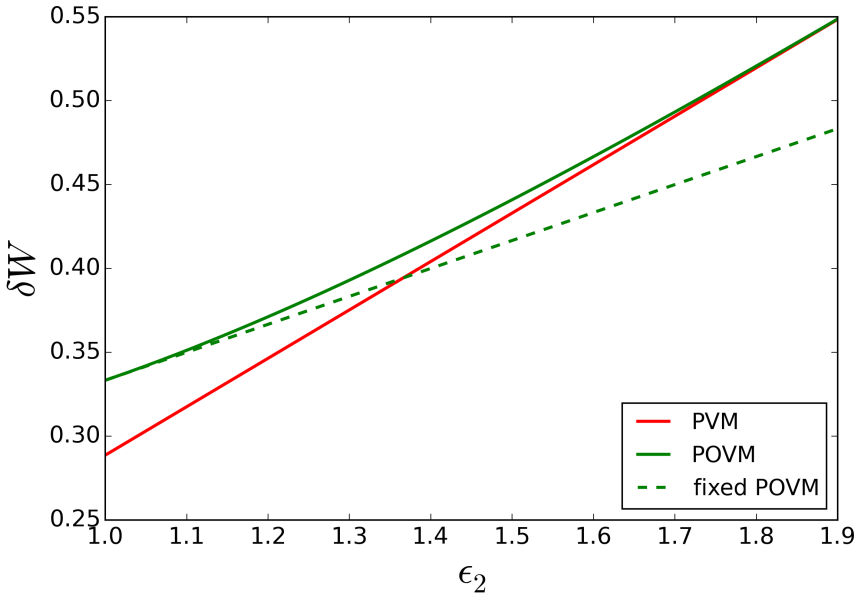

Figure 5.1: Daemonic gain for the state defined in equation 5.31 and the Hamiltonian $H = |1\rangle\langle 1| + \epsilon_2|2\rangle\langle 2|$ for the Mercedes-star POVM (green dashed line), the optimal POVM (solid green line) and the optimal projective measurement (red line).

5.2 Multipartite Daemonic Ergotropy

In chapter four, we already studied daemonic ergotropy for three qubits. Anyway, daemonic ergotropy still remained a bipartite quantity, since we grouped two qubits to form either the ancilla or the system. Now, we extend daemonic ergotropy to the multipartite case.

In bipartite daemonic ergotropy one considers a system S and an ancilla A which both play fundamentally different roles. From S one extracts energy while the ancilla can only be measured. This asymmetry should prevail in a multipartite context. Maintaining the division into systems and ancillas, there are two extremal cases of multipartite generalisations to study. In one case, there are several systems from which energy is extracted via local unitaries and a single ancilla. In the other case, there is only one system but several ancillas on which measurements may be performed locally. In the most general case, one would then be confronted with several systems and ancillas. This is illustrated in figure (5.2).

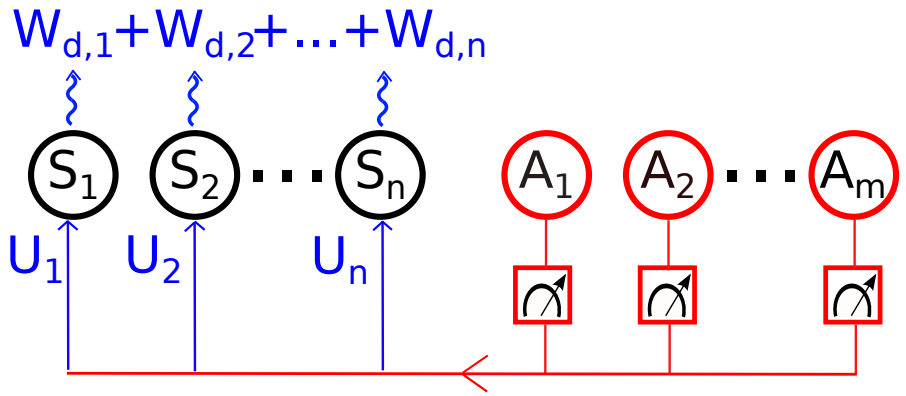

Figure 5.2: Energy extraction scheme for multipartite daemonic ergotropy

If we are concerned with several systems, we first need to specify a multipartite generalisation of ergotropy. We demand the total energy of all systems to equal the sum of the energies in the individual systems. Therefore, with the local Hamiltonians H_i acting on system i, the total Hamiltonian should be $H = \sum_i H_i$, where we write H_i short for $\mathbb{1}_1 \otimes \ldots \otimes \mathbb{1}_{i-1} \otimes H_i \otimes \mathbb{1}_{i+1} \otimes \ldots \otimes \mathbb{1}_n$ if there are n systems. With this, we can write the multipartite ergotropy as

$$W^n(\varrho_S, H) = \mathrm{Tr}\left(\varrho_S H\right) - \min_U \mathrm{Tr}\left(\varrho_S U H U^\dagger\right), \tag{5.60}$$

where the minimisation takes only product unitaries $U = U_1 \otimes \ldots \otimes U_n$ into account and ϱ_S is the joint state of all systems. With the definition of the Hamiltonian and the

product unitaries, we can rewrite the multipartite ergotropy as

$$W^n(\varrho_S, H) = \sum_{i=1}^{n} \mathrm{Tr}\left(\varrho_i H_i\right) - \sum_i \min_{U_i} \mathrm{Tr}\left(\varrho_i U_i^\dagger H_i U_i\right) \tag{5.61}$$

$$= \sum_i W(\varrho_i, H_i), \tag{5.62}$$

where ϱ_i is the reduced state of the ith system. Since the minimisation is restricted to product unitaries, multipartite ergotropy is bounded from above by the single party ergotropy $W(\varrho_S, H)$.

In the bipartite setting one can express the daemonic ergotropy as

$$W_D(\varrho_{SA}, H, E) = \sum_\mu W(\varrho^\mu, H). \tag{5.63}$$

where $\varrho^\mu = \mathrm{Tr}_A(\varrho_{SA}(\mathbb{1} \otimes E_\mu))$ is the unnormalised conditional state of the system for outcome μ.

In the same fashion, we can write the daemonic ergotropy in the n-system setting as

$$W_D^n(\varrho_{SA}, H, E) = \sum_\mu W^n(\varrho_S^\mu, H) \tag{5.64}$$

$$= \sum_{\mu,i} W(\varrho_i^\mu, H_i) \tag{5.65}$$

where ϱ_i^μ is the unnormalised reduced conditional state of system i if the measurement outcome was μ.

Note, that if we are dealing with several systems, it is sufficient to consider pure states in our analysis since considering a mixed state is equivalent to appending another system with a completely degenerate local Hamiltonian and considering the purification of the state, such that tracing over the appended system will yield back the original state.

In the case of several ancillas and one system, the multipartite generalisation does not change the situation for pure states because every projective measurement with rank-one projectors is optimal, including projectors on a product basis, which the multipartite version of daemonic ergotropy allows. Thus, it is only interesting to consider mixed states, which is then equivalent to a pure state scenario with many ancillas and many systems.

Therefore, we first aim to gain a better understanding of the problem with many systems, only one ancilla and pure states.

5.2.1 Optimal Measurement for Pure States

We first note that in contrast to the bipartite case, in the multipartite case projective measurements do in general not even suffice for pure states. One directly sees this from

the example considered in section (5.1.5). First, we diagonalise the classical-quantum state from the example and write

$$\varrho_{SA} = \sum_i p_i |i\rangle\langle i|_{SA}. \tag{5.66}$$

Now, we append another system S_2 and consider the purification

$$|\Psi\rangle = \sum_i \sqrt{p_i}\, |i\rangle_{SA} \otimes |i\rangle_{S_2}. \tag{5.67}$$

For system S_2 we choose a completely degenerate Hamiltonian $H_2 = h\mathbb{1}$. Using equation (5.65) we can write

$$W_D^n(|\Psi\rangle, H, E) = \sum_\mu W(\varrho_S^\mu, H_S) + \sum_\mu W(\varrho_{S_2}^\mu, H_2). \tag{5.68}$$

Since H_2 is completely degenerate, the second term is a constant. The first term is exactly the bipartite daemonic ergotropy as considered in section (5.1.5) which was maximised for a generalised measurement. Thus, also the multipartite daemonic ergotropy for the pure state $|\Psi\rangle$ is maximised for the same generalised measurement, which proves the claim.

For GHZ states

$$|GHZ\rangle = \frac{1}{\sqrt{2}}(\,|0_{S_1}\ldots 0_{S_n} 0_{A_1}\ldots 0_{A_m}\rangle + |1_{S_1}\ldots 1_{S_n} 1_{A_1}\ldots 1_{A_m}\rangle) \tag{5.69}$$

however, the local projective measurements on $|0\rangle$ and $|1\rangle$ are optimal, since the conditional state of all systems is a pure product state independently of the outcome and its energy can thus be minimised using local unitaries.

We now investigate the multipartite daemonic ergotropy for W-like states

$$|W\rangle = \cos\phi \sin\theta\, |1_{S_1} 0_{S_2} 0_A\rangle + \sin\phi \sin\theta\, |0_{S_1} 1_{S_2} 0_A\rangle + \cos\theta\, |0_{S_1} 0_{S_2} 1_A\rangle \tag{5.70}$$

where the local Hamiltonians for the systems S_1 and S_2 are both $H = |1\rangle\langle 1|$. The results for multipartite ergotropy, multipartite daemonic ergotropy and multipartite daemonic gain are shown in the plots below. Since all quantities have a periodicity of $\pi/2$ in ϕ, the range of ϕ has been reduced accordingly.

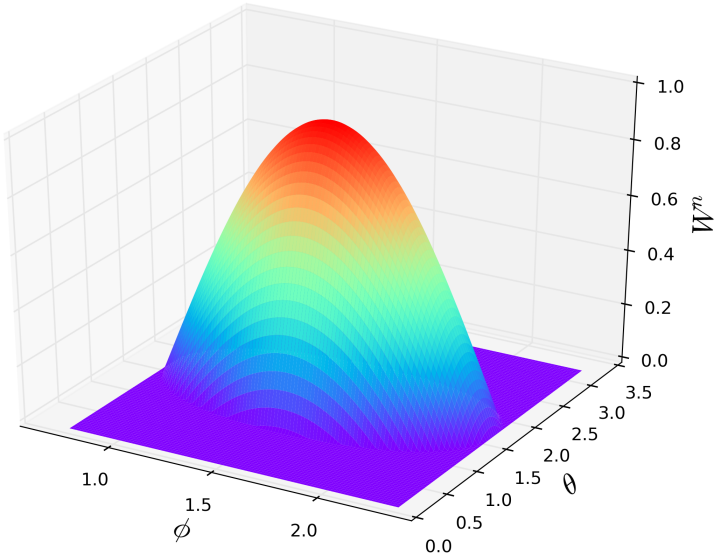

Figure 5.3: Multipartite ergotropy for $|W\rangle$ as defined in equation (5.70)

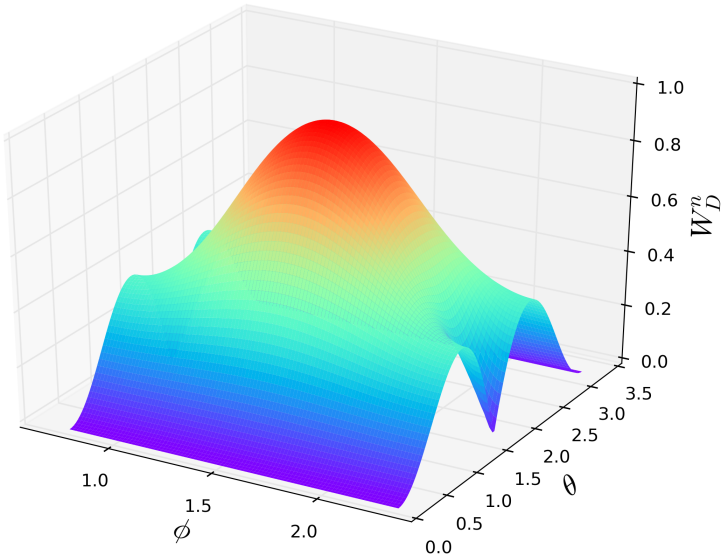

Figure 5.4: Multipartite daemonic ergotropy for $|W\rangle$ as defined in equation (5.70)

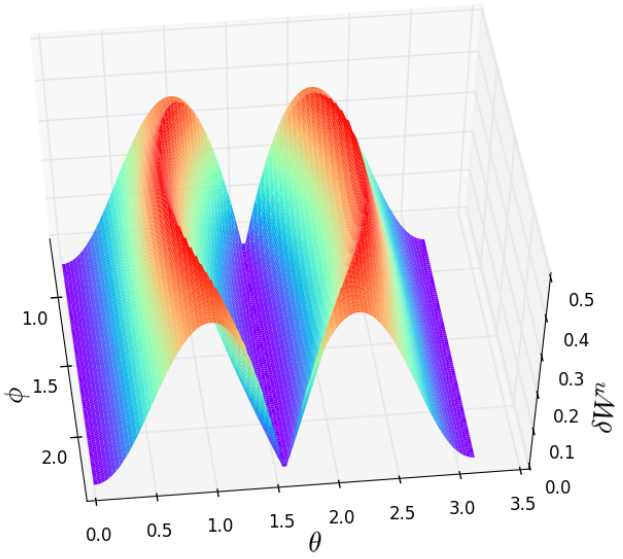

Figure 5.5: Multipartite daemonic gain for $|W\rangle$ as defined in equation (5.70)

Numerics show that a possible advantage when using generalised measurements is smaller than numerical precision in this case.

5.2.2 Optimal Measurement for Quantum-Classical States

If the global state of all systems $S_1, \ldots, S_n$ and the ancilla A is in a quantum-classical state

$$\varrho_{S_1 \ldots S_n A} = \sum_i \sigma^i_{S_1 \ldots S_n} \otimes |i\rangle\langle i|_A \tag{5.71}$$

we can repeat the proof presented in section (5.1.3) to show that the projective measurement with projectors $|i\rangle\langle i|$ is still optimal. This of course is still true for the special case in which the ancilla is divided into multiple subsystems and the state is

$$\varrho_{S_1 \ldots S_n A_1 \ldots A_m} = \sum_i \sigma^i_{S_1 \ldots S_n} \otimes |i\rangle\langle i|_{A_1} \otimes \ldots |i\rangle\langle i|_{A_m}, \tag{5.72}$$

for which the optimal measurement consists of the local projective measurements with elements $|i\rangle\langle i|_{A_k}$.

Conclusion

In this work, we generalised daemonic ergotropy in two ways, by extending the class of considered measurements and suggesting a genuine multipartite generalisation.

As groundwork, we examined the relation between multipartite entanglement and bipartite daemonic ergotropy for three qubit systems. We analytically found states that maximise daemonic gain at a given value of multipartite negativity and found a good analytic approximation of states that minimise daemonic gain at a given value of multipartite negativity. We showed numerically that the found states are not specific to the used entanglement measure but also turn out to be optimal for a variety of other entanglement measures. Exploiting symmetries of daemonic gain and entanglement of formation in the state parameters, we transfer our results of the two-qubit system case to the two-qubit ancilla case.

Then, we extend the definition of daemonic ergotropy to allow for generalised measurements and present a see-saw algorithm to find the optimal measurement and calculate the daemonic ergotropy. We argue that for pure states, any rank-one projective measurement is optimal in the bipartite setting. Thereafter, we prove that an optimal measurement for the quantum-classical state

$$\varrho_{SA} = \sum_i \sigma_i^S \otimes |i\rangle\langle i|^A \tag{5.73}$$

is the projective measurement with effects $|i\rangle\langle i|^A$. This result then shows that entanglement and discord $\delta_{S:A}$ (but not $\delta_{A:S}$) are not useful for energy extraction if a projective measurement is optimal.

Subsequently, we consider a classical-quantum state and show that using generalised measurements to measure the ancilla can give an advantage that depends on the Hamiltonian.

Afterwards, we suggest a multipartite generalisation of daemonic ergotropy. By considering the purification of the previous example, we show that for pure states, projective measurements are in general no longer optimal. For $|GHZ\rangle$ states however, we argue that projective measurements are always optimal and go on to compute the daemonic gain of W-like states

$$|W\rangle_{S_1S_2A} \propto a\,|100\rangle + b\,|010\rangle + c\,|001\rangle \tag{5.74}$$

for identical Hamiltonians on the first and the second system and note that in this case projective measurements are also optimal. This rises the question, how the Hamiltonian determines the optimal measurement and whether generalised measurements still perform better in multipartite systems with identical local Hamiltonians.

In the end, we remark that also in the multipartite setting, one optimal measurement for the quantum-classical state

$$\varrho_{SA} = \sum_i \sigma^i_{S_1\ldots S_n} \otimes |i\rangle\langle i|_{A_1} \otimes \ldots \otimes |i\rangle\langle i|_{A_m} \tag{5.75}$$

is given by the projectors on $|i_{A_1} j_{A_2} \ldots k_{A_m}\rangle$.

Appendix A

Semidefinite Programming

Semidefinite programmes (SDPs) form a subclass of convex optimisation problems. A convex optimisation problem is a problem of the form

$$\begin{aligned} \min \quad & f(x) \\ \text{subject to} \quad & g_i(x) \leq 0 \quad i \in \{1, \ldots, m\} \end{aligned} \tag{A.1}$$

where both the objective function f and the constraint functions g_i are convex functions [32]. A semidefinite programm is a convex optimisation problem on the cone of positive semidefinite matrices with a linear objective function and affine constraints.

One way to state a semidefinite programme is thus

$$\begin{aligned} & \min_{X \in \mathbb{S}^n} \langle C, X \rangle_{\mathbb{S}^n} \\ \text{subject to} \quad & \langle A_k, X \rangle_{\mathbb{S}^n} = a_k, \quad k = 1, \ldots, m \\ & X \succeq 0, \end{aligned} \tag{A.2}$$

which is called the conic form of the SDP [21]. By $\mathbb{S}^n$, we denote the class of (real) symmetric $n \times n$ matrices, $\succeq$ is the Loewner order and $a_k \in \mathbb{R}$ and A_k are $n \times n$-matrices. Note, that SDPs also cover inequality constraints such as

$$\langle B_l, X \rangle_{\mathbb{S}^n} \leq b_l, \quad l = 1, \ldots, s. \tag{A.3}$$

Those can, however, be transformed into equality constraints by introducing unknown but positive slack variables c_k in the following way

$$\langle B_l, X \rangle_{\mathbb{S}^n} = b_l - c_l, \quad c_l \geq 0. \tag{A.4}$$

Conversely, equality constraints are just a special case of inequality constraints in which the slack variables are fixed. This is why one can slim down the definition of semidefinite programmes by only choosing one way of representing constraints.

The Dual Problem

To every SDP there is a *dual* SDP which for problem A.2 is defined as

$$\begin{aligned} &\max_{y\in\mathbb{R}^m} \langle a, y\rangle_{\mathbb{R}^m} \\ \text{subject to } &\sum_{i=1}^{m} y_i A_i \succeq C. \end{aligned} \tag{A.5}$$

The dual SDP is another SDP, that is related to the original, also called *primal* SDP through the optimal values of their objective functions. Specifically, the achieved minimum in the primal SDP is always greater or equal than the maximum that can be achieved in the dual problem. This is called *weak duality* and the difference between the two optima is called *duality gap*. Oftentimes, the duality gap closes, in which case the duality between the primal and the dual SDP is called *strong duality*. We will now first show that the dual SDP as written above is indeed a proper SDP by showing that it can be converted into the conic form and discuss duality afterwards.

In order to do this, we first introduce a slack variable S

$$\begin{aligned} &\min_{y\in\mathbb{R}^m} \langle -a, y\rangle_{\mathbb{R}^m} \\ \text{subject to } &\sum_{i=1}^{m} y_i A_i + S = C \\ &S \succeq 0. \end{aligned} \tag{A.6}$$

Now that the third line matches the one in equation A.2, we rewrite the first line by defining a matrix G_0 that satisfies

$$\langle G_0, A_i\rangle = a_i, \quad i = 1, \ldots, m. \tag{A.7}$$

This allows us to write

$$-\sum_i y_i a_i = -\sum_i y_i \langle G_0, A_i\rangle \tag{A.8}$$

$$= -\langle G_0, \sum_i y_i A_i\rangle \tag{A.9}$$

$$= -\langle G_0, C - S\rangle \tag{A.10}$$

$$= \langle G_0, S\rangle - \langle G_0, C\rangle. \tag{A.11}$$

Since G_0 and C are fixed, minimising $\langle -a, y\rangle_{\mathbb{R}^m}$ is equivalent to minimising $\langle G_0, S\rangle$ over all possible S. For any given SDP, the set of possible values of S is determined by the m parameters y_m as $S = C - \sum_i y_i A_i$. This allows us to characterise the set of possible values of S as the set of symmetric matrices that obey $n(n+1)/2 - m$ linear constraints. This is, because symmetric matrices have $n(n+1)/2$ free parameters, which leaves m

free parameters after imposing $n(n+1)/2 - m$ constraints. We can write these constraints in matrix form by introducing matrices G_i that fulfil

$$\langle G_i, S\rangle = d_i, i = 1, \ldots, \frac{n(n+1)}{2} - m \tag{A.12}$$

for some real numbers d_i [21]. We can now write the original problem A.6 in the conic form as

$$\begin{aligned} &\min_{S\in\mathbb{S}^n} \langle G_0, S\rangle_{\mathbb{S}^n} \\ \text{subject to} \quad &\langle G_i, S\rangle_{\mathbb{S}^n} = d_i, \quad i = 1, \ldots, \frac{n(n+1)}{2} - m \\ &S \succeq 0. \end{aligned} \tag{A.13}$$

The optimal value of the primal SDP is always greater or equal than the optimal value of its dual. This is called weak duality and it can be easily shown by simply computing the duality gap [32]

$$\langle C, X\rangle - \langle a, y\rangle = \langle C, X\rangle - \sum_{i=1}^{m} y_i a_i \tag{A.14}$$

$$= \langle C, X\rangle - \sum_i y_i \langle A_i, X\rangle \tag{A.15}$$

$$= \langle C - \sum_i y_i A_i, X\rangle \geq 0. \tag{A.16}$$

The final inequality holds, because both $S = C - \sum_i y_i A_i$ and X are positive semidefinite matrices.

If the primal objective is bounded below and the problem is strictly feasible, which in our formulation of the primal problem means that there is a positive definite operator $\bar{X}$ that satisfies $\langle A_k, X\rangle = a_k$, then there is an optimal solution to the dual problem and the duality gap closes. In complete analogy, if the dual problem is bounded above and strictly feasible, this is there exists a vector y that satisfies $\sum_i y_i A_i \succ C$, then there is an optimal solution to the primal problem and the duality gap closes. These conditions are called *Slater conditions*[32].

One can also directly a semidefinite programme in complex variables. In this case, the primal problem is

$$\max \quad \langle A, X\rangle \tag{A.17}$$

$$\text{subject to} \quad \Phi(X) = B \tag{A.18}$$

$$X \succeq 0 \tag{A.19}$$

and its dual problem is

$$\min \quad \langle B, Y\rangle \tag{A.20}$$

$$\text{subject to} \quad \Phi^*(Y) \succeq A \tag{A.21}$$

$$Y \text{ hermitian} \tag{A.22}$$

where Φ is a hermiticity preserving map [33]. Many software packages do however not accept to pose the problem in this form. In this case one needs to convert the complex problem into a real one. This can for example be done [34] by converting a hermitian matrix ϱ into a real, symmetric matrix

$$\sigma = \begin{pmatrix} \mathrm{Re}(\varrho) & \mathrm{Im}(\varrho) \\ -\mathrm{Im}(\varrho) & \mathrm{Re}(\varrho) \end{pmatrix}. \tag{A.23}$$

Karush-Kuhn-Tucker Conditions

The Karush-Kuhn-Tucker Conditions (KKT conditions) are a set of sufficient conditions for an optimal solution $\bar{x}$ of a convex optimisation problem. For a convex optimisation problem as stated in the very beginning in equation (A.1), the KKT conditions state that if $\bar{x}$ is optimal, then there must exist multpliers λ_i, such that [32]

$$\nabla f(\bar{x}) + \sum_{i=1}^{m} \lambda_i \nabla g_i(\bar{x}) = 0 \tag{A.24}$$

$$g_i(\bar{x}) \leq 0 \tag{A.25}$$

$$\lambda_i g_i(\bar{x}) = 0 \tag{A.26}$$

$$\lambda_i \geq 0 \tag{A.27}$$

hold for all $i \in \{1, \ldots, m\}$. One can easily proove that if $\bar{x}$ satisfies the above conditions, it must be optimal. Since f and g_i are convex functions, we have

$$f(x) - f(\bar{x}) \geq (x - \bar{x})^T \nabla f(\bar{x}) \tag{A.28}$$

$$g_i(x) - g_i(\bar{x}) \geq (x - \bar{x})^T \nabla g_i(\bar{x}), \tag{A.29}$$

where the second equation implies

$$\sum_i \lambda_i g_i(x) - \sum_i \lambda_i g_i(\bar{x}) \geq (x - \bar{x})^T \sum_i \lambda_i \nabla g_i(\bar{x}). \tag{A.30}$$

Due to the third KKT condition, the second term on the left hand side vanishes and the fourth KKT condition together with the feasibility condition implies that the first term is smaller or equal zero. Now, we can sum up this inequality and the inequality (A.28) to yield

$$f(x) - f(\bar{x}) + \sum_i \lambda_i g_i(x) \geq (x - \bar{x})^T [\nabla f(\bar{x}) + \sum_i \lambda_i \nabla g_i(\bar{x})]. \tag{A.31}$$

From the first KKT condition, we know that the term in square brackets must be zero and since the sum on the left hand side must be non-positive, $\bar{x}$ must indeed be a global minimum [32].

Bibliography

[1] John Goold, Marcus Huber, Arnau Riera, Lídia del Rio, and Paul Skrzypczyk. The role of quantum information in thermodynamics—a topical review. *Journal of Physics A: Mathematical and Theoretical*, 49(14):143001, 2016.

[2] A. E. Allahverdyan, R. Balian, and Th. M. Nieuwenhuizen. Maximal work extraction from finite quantum systems. *EPL (Europhysics Letters)*, 67(4):565, 2004.

[3] Gianluca Francica, John Goold, Francesco Plastina, and Mauro Paternostro. Daemonic ergotropy: enhanced work extraction from quantum correlations. *Nature*, 2017.

[4] Teiko Heinosaari and Mário Ziman. *The Mathematical Language of Quantum Theory: From Uncertainty to Entanglement.* Cambridge University Press, 2011.

[5] Man-Duen Choi. Completely positive linear maps on complex matrices. 10:285–290, 06 1975.

[6] Vern Paulsen. *Completely Bounded Maps and Operator Algebras.* Cambridge Studies in Advanced Mathematics. Cambridge University Press, 2003.

[7] M.A. Nielsen and I.L. Chuang. *Quantum Computation and Quantum Information.* Cambridge Series on Information and the Natural Sciences. Cambridge University Press, 2000.

[8] Giulio Chiribella, Giacomo Mauro D'Ariano, and Dirk Schlingemann. How continuous quantum measurements in finite dimensions are actually discrete. *Phys. Rev. Lett.*, 98:190403, May 2007.

[9] Asher Peres. *Quantum Theory: Concepts and Methods.* 1995.

[10] Otfried Gühne and Géza Tóth. Entanglement detection. 2009.

[11] Michał Horodecki, Paweł Horodeckib, and Ryszard Horodecki. *Physics Letters A*, 223:1–8, 1996.

[12] Private communication with Otfried Guehne and Chau Nguyen.

[13] Thomas Scheidl, Rupert Ursin, Johannes Kofler, Sven Ramelow, Xiao-Song Ma, Thomas Herbst, Lothar Ratschbacher, Alessandro Fedrizzi, Nathan K. Langford, Thomas Jennewein, and Anton Zeilinger. Violation of local realism with freedom of choice. *Proceedings of the National Academy of Sciences*, 2010.

[14] J Lavoie, R Kaltenbaek, and K J Resch. Experimental violation of svetlichny's inequality. *New Journal of Physics*, 11(7):073051, 2009.

[15] W Dür, G Vidal, and J I. Cirac. Three qubits can be entangled in two inequivalent ways. 62, 05 2000.

[16] A Acín, A Andrianov, E Jané, and R Tarrach. Three-qubit pure-state canonical forms. *Journal of Physics A: Mathematical and General*, 34(35):6725, 2001.

[17] Stephen M. Barnett and Sarah Croke. Quantum state discrimination. *Adv. Opt. Photon.*, 1(2):238–278, Apr 2009.

[18] Ulrike Herzog and János A. Bergou. Distinguishing mixed quantum states: Minimum-error discrimination versus optimum unambiguous discrimination. *Phys. Rev. A*, 70:022302, Aug 2004.

[19] I.D. Ivanovic. How to differentiate between non-orthogonal states. *Physics Letters A*, 123(6):257 – 259, 1987.

[20] Asher Peres. How to differentiate between non-orthogonal states. *Physics Letters A*, 128(1):19, 1988.

[21] Henry Wolkowicz, Romesh Saigal, and Lieven Vandenberghe. *Handbook of Semidefinite Programming: Theory, Algorithms, and Applications*. International Series in Operations Research and Management Science. Springer, 2000.

[22] Maximilian Schlosshauer. Decoherence, the measurement problem, and interpretations of quantum mechanics. *Rev. Mod. Phys.*, 76:1267–1305, Feb 2005.

[23] Quantum coherence. 453(7198):1003–1049.

[24] L Henderson and V Vedral. Classical, quantum and total correlations. *Journal of Physics A: Mathematical and General*, 34(35):6899, 2001.

[25] Harold Ollivier and Wojciech H. Zurek. Quantum discord: A measure of the quantumness of correlations. *Phys. Rev. Lett.*, 88:017901, Dec 2001.

[26] F. Galve, G. L. Giorgi, and R. Zambrini. Orthogonal measurements are almost sufficient for quantum discord of two qubits. *EPL (Europhysics Letters)*, 96(4):40005, 2011.

[27] Asma Al Qasimi and Daniel F. V. James. Comparison of the attempts of quantum discord and quantum entanglement to capture quantum correlations. *Phys. Rev. A*, 83:032101, Mar 2011.

[28] L. Mirsky. A trace inequality of John von Neumann. *Monatshefte für Mathematik*, 79(4):303–306, Dec 1975.

[29] A. Lenard. Thermodynamical proof of the Gibbs formula for elementary quantum systems. *Journal of Statistical Physics*, 19(6):575–586, Dec 1978.

[30] Private communication with Mauro Paternostro.

[31] Private communication with Roope Uola.

[32] S. Boyd and L. Vandenberghe. *Convex Optimization*. Cambridge University Press, 2004.

[33] John Watrous. *Theory of Quantum Information*. 2011.

[34] Private communication with Matthias Kleinmann.

Erklärung

Hiermit erkläre ich, dass ich die vorliegende Masterarbeit selbständig verfasst und keine anderen als die angegebenen Quellen und Hilfsmittel benutzt, sowie Zitate und Ergebnisse Anderer kenntlich gemacht habe.

....

(Ort) (Datum) (Unterschrift)