universität wien

MASTERARBEIT / MASTER'S THESIS

Titel der Masterarbeit / Title of the Master's Thesis

„Non-Markovian dynamics in quantum cooling“

verfasst von / submitted by

Philipp Alexander Schüttelkopf BSc

angestrebter akademischer Grad / in partial fulfilment of the requirements for the degree of

Master of Science (MSc)

Wien, 2020 / Vienna, 2020

Studienkennzahl lt. Studienblatt / degree programme code as it appears on the student record sheet:	UA 066 876
Studienrichtung lt. Studienblatt / degree programme as it appears on the student record sheet:	Masterstudium Physik UG2002
Betreut von / Supervisor:	Mag. Dr. Marcus Huber, Privatdoz.

Acknowledgements/Dedication

Gratefully dedicated to my family who have supported me throughout my studies.

I am thankful for my friends who have encouraged me in my scientific pursuit.

Furthermore, I want to acknowledge the advice of Mag. Dr. Marcus Huber, with whom the idea of this thesis was created and the incredibly valuable guidance of Philip Taranto, without whom this thesis would not have been completed.

Abstract

Preparing cold, or pure, states is an essential part of many quantum technologies and experiments. As such, it has been the aim of various protocols, differing in the underlying structures, assumptions and cooling mechanisms. However, due to the third law of thermodynamics, it is not possible to cool to absolute zero without infinite time or resources. In practice, one is typically limited by finite resources, and so understanding the realistic limits achievable is of utmost importance. Such a setting provided the motivation for Refs. [1, 2], which derived the ultimate cooling limits in any memoryless, or Markovian, operating paradigm, in which sequential cooling steps operate independently of those previous. Here we extend this paradigm to consider non-Markovian effects on the task of quantum cooling. By considering a physically-motivated generalised collision model endowed with memory, we show that allowing for memory carriers that accompany the target system throughout the cooling process exponentially enhances the upper bound on the asymptotic ground state population compared to Markovian case [3]. Additionally, we simulate the optimal cooling protocol for various memory structures, highlighting that memory can also lead to a significant finite-time advantage.

Zusammenfassung

Die Präparation kalter oder reiner Zustände ist ein wesentlicher Bestandteil vieler Quantentechnologien und -experimente. Als solches war es das Ziel verschiedener Protokolle, die sich in den zugrunde liegenden Strukturen, Annahmen und Kühlmechanismen unterscheiden. Aufgrund des dritten Hauptsatzes der Thermodynamik ist es jedoch nicht möglich, ohne unendliche Zeit oder Ressourcen auf den absoluten Nullpunkt abzukühlen. In der Praxis ist man in der Regel durch endliche Ressourcen begrenzt, und daher ist das Verständnis der realistischen Grenzen, die man erreichen kann, von größter Bedeutung. Eine solche Einstellung lieferte die Motivation für Refs. [1, 2], das die ultimativen Kühlgrenzen in jedem speicherlosen oder Markovschen Rahmen ableitete, in dem aufeinanderfolgende Kühlschritte unabhängig von den vorherigen ablaufen. Hier erweitern wir dieses System, um nicht-Markovsche Effekte im Hinblick auf die Quantenkühlung zu berücksichtigen. Durch die Betrachtung eines physikalisch motivierten, verallgemeinerten und mit Speicher ausgestatteten Kollisionsmodells zeigen wir, dass die Berücksichtigung von Speicherträgern, die das Zielsystem während des gesamten Kühlprozesses begleiten, die Obergrenze der asymptotischen Grundzustandspopulation im Vergleich zum Markovschen Fall exponentiell erhöht [3]. Zusätzlich simulieren wir das optimale Kühlungsprotokoll für verschiedene Speicherstrukturen, wobei wir betonen, dass Speicher auch zu einem signifikanten Vorteil für endliche Zeiten führen kann.

Contents

1 Introduction

The underlying goal of this thesis is to understand the limitations of preparing pure states with realistic resources. Such states are an important part of many applications, for instance quantum computing [4], quantum sensing [5], and quantum metrology [6]. However, physical systems typically tend to thermalise with their environment [7, 8], meaning that readily-available states are not pure, but rather mixed, and in equilibrium with respect to their environment. Therefore, we wish to develop a protocol that allows us to utilise such thermal states in order to prepare useful pure ones.

Pure states are often referred to as "cold", because they have zero entropy. A pure state of special thermodynamic interest is the energy ground state, which is synonymous with zero temperature. However, perfect preparation of such a state, and subsequently perfect cooling to absolute zero, without either infinite time, resources, or complexity is forbidden by the third law of thermodynamics [9–11]. In practice, we never have such infinite resources available. Subsequently, this poses the question of how well cooling can be performed given realistic means, such as a finite dimensional environment or limited-range interactions.

The principal framework we consider comprises a target system and a thermal environment, as well as control over their interactions, which we wish to fine-tune in order to cool the system. Subsequently, we identify the environment as a quantum *machine*, which we control via implementing appropriate operations. A single step of the cooling protocol consists of bringing the system into contact with the environment, applying a joint unitary operation and lastly, the environment is traced out. For any individual step of the total process, we consider two cases: one in which we allow for arbitrary unitary dynamics between the system and environment, and one where we restrict considerations to energy-conserving unitaries, leading to the framework of thermal operations that is often considered in resource theories of thermodynamics [12, 13].

Having said that, we are ultimately interested in the performance of multiple successive cycles. There are various strategies on how to implement multiple cycles, whose end results may markedly differ. For instance, we could consider a complete rethermalisation of the depleted machine between each time-step. This is equivalent to discarding the used machine and taking a new one that is in thermal equilibrium with the initial bath, completely erasing all correlations built up between the system and environment. As such, the sequence of operations are independent of each other and the overall process is called Markovian. In contrast, we could consider reusing some of the machines that have already come into contact with the system. As these machines contain information regarding the system's history, performing such interactions generally leads to memory effects, which have already been shown to improve some important thermodynamic tasks [14], for example quantum battery charging [15]. Therefore, this thesis aims to investigate the effects of memory for possible cooling advantages.

The fundamental cooling limitations for the memoryless (Markovian) cooling regime have already been derived by Clivaz et al. [1, 2].The authors identify a fundamental bound that limits how cool the system can be prepared in the asymptotic limit of infinitely many independent cooling cycles. Remarkably, this bound depends only on the largest energy gap of the machine, rather than any intricate microscopic details of the protocol. Here we extend the framework employed there to the non-Markovian setting, where temporal correlations play a role. We do not allow for arbitrary non-Markovianity, as this would immediately trivialise the problem (e.g., by allowing for the system and environment to be initially entangled and pure, which permits perfect cooling in a single step). Nonetheless, we propose a framework that allows for a tractable amount of non-Markovianity and permits meaningful and fair comparisons for finite times. We do this via a particular type of generalised collision model with memory [16–20].

In this thesis, we first show that memory enhances the ultimate cooling bound in the limit of infinitely many cycles, exponentially quickly in the number of memory carriers. Additionally, we focus on the finite-time analysis and compare the performance of various memory structures and control regimes. We show that protocols with memory carriers can more efficiently utilise a limited amount of machines to prepare colder states compared to memoryless protocols.

This thesis is organised as follows: In Chapter 2, the appropriate scientific context is laid out. Next, in Chapter 3, we establish the framework in which we operate. In Chapter 4 we present the main results of this work, where comparisons of various cooling regimes are discussed. Lastly, in Chapter 5, all findings are summarised. Note that many of the ideas and results of this thesis have been presented in Ref. [3].

2 Background

Highly pure states represent a valuable resource. As such, producing them is an essential goal of countless experiments. However, they typically do not occur in nature due to the tendency of systems to thermalise with their surrounding environment, which is deeply connected to the second law of thermodynamics [8, 21]. Therefore, we have to generate useful pure states from readily available resources, which has been the aim of a variety of protocols [22, 23].

One way to purify a state is to decrease its temperature, which can be uniquely assigned to states that are in thermal equilibrium. In order to cool such states, it is sufficient to drive them towards a thermal state with a lower temperature. However, for states that are out of thermal equilibrium, temperature and therefore any notion of cooling is harder to define, because temperature is not an observable, but an average over observables [24]. However, in the world of thermodynamics, one of the key observable quantities is energy, and thus many studies of purifying states from a thermodynamic perspective focus on increasing the population of the lowest energy eigenstates, which corresponds to lowering the energy, and can thus be seen as synonymous with cooling. Since the partial ordering induced by majorization implies a clear notion of cooling that agrees with any of the aforementioned notions, we rely on majorization theory to determine coolness.

Since we want to start with an open mindset, we have to ask ourselves: What are the readily available states and operations that are "easy" to implement? A lot of research has been engaged with developing an understanding of thermodynamic resources. Systems tend to thermalise and are then formally described as *Gibbs states* [8, 25]; these *thermal states* that are in equilibrium with respect to their environment have thus been posited as a reasonable set of free states. Thermal states are completely passive [7, 26], meaning that no energy can be extracted from them alone, or even from an arbitrary number of copies via unitary transformations. Because they constitute a free resource, we take these to describe the initial states of our system and environment. Moreover, thermal states are connected to thermal operations, which have been found to be free as well and which we consider later in our *incoherent control* scenario [7, 12, 13, 27].

A scheme that makes use of such reasonable resources is for example *heat bath algorithmic cooling* **(HBAC)** [28–31], where entropy from the target qubit and a number of compression qubits is transferred to reset qubits, which then rethermalise with the environment. There are other ideas, e.g.: Absorption refrigerators [32, 33], which make use of resonant energy transitions in a tripartite system to dissipate heat away from the target and another hot system and into the third one. These ideas proposes concrete mechanisms to cool quantum systems. Here, instead of beginning with such a concrete model in mind, we take a more general approach by considering physically-reasonable thermodynamic resource constraints, placed upon the generic open quantum systems framework [34]. This allows us to ask questions regarding fundamental limitations. As it turns out, HBAC is a special case of our generalised framework. HBAC considers mainly qubits (spin-$\frac{1}{2}$ particles) for the target, compression and reset system, while we do not restrict the dimension or energy spectrum of our machines. Similarly to the compression qubits in HBAC, we allow for machines to re-interact with the target system and carry the memory onward through the process, which generates an advantage over memoryless schemes.

If there is no memory involved and the operations performed at each step are independent of the previous one, the whole process becomes Markovian. This is synonymous with the non-existence of temporal correlations between the system and the environment, or the reasonable negligence thereof. In the Markovian approximation it is assumed, the environment is very much bigger than the system and any correlations get "washed out" fairly quickly [34]. Furthermore, when each step of the protocol is identical, convergence properties to the asymptotic case can be seen from the spectral properties of the dynamical maps. Memory-less, time-independent processes are typically described by a collision model for discrete times [16, 19, 35], or in the continuous time limit by a Markovian master equation in the Gorini-Kossakowski-Lindblad-Sudarshan (**GKLS**) form [36–38]. Another related collision model in the continuous time limit is a *repeated interaction system* [39, 40], where the system of interest interacts unitarily with a large amount of ancillas in rapid succession. However, this model is not suited for our intention, as we wish to compare regimes for finite times regarding the utilisation of discrete resources to achieve cooling.

On the other hand, non-Markovian processes require knowledge about the history of the process in order to calculate the further evolution, which makes them harder to consider, as arbitrary dynamics can lead to the process quickly becoming intractable [41–45]. For example, interactions and re-interactions of a system with an ever growing amount of environment-subsystems, which in turn interact among themselves, would lead to an exponentially exploding amount of parameters to track. Therefore, in

order to allow for the investigation of possible advantages in a realistic setting, these memory effects have to be restricted. Lately, non-Markovian dynamics have been found to improve the performance of quantum batteries [15] and quantum key distribution [46]. Additionally, considering non-Markovianity can have advantages in the context of quantum control [47]. For the task of cooling, we consider it meaningful to restrict to a physically-reasonable form of memory: The agent can control a (small) number of subsystems and feed these forward in time. These subsystems re-interact with the target and allow for an information backflow and therefore evidence memory effects.

This type of memory can be suitably included in a generalised collision model. Such collision models have already been used in order to investigate non-Markovian dynamics [17, 18, 48–50], because they allow for a structured, yet flexible, description of certain non-Markovian effects. Nonetheless, such models are not restricted to specific states or interactions, as the memory length only defines how many memory carriers are fed forward. As we will see later, such models can often lead to tractable non-Markovian dynamics, because they can be embedded into Markovian models by including a small number of additional systems, i.e. the memory carriers, as part of the system of interest [51].

3 Framework

In this chapter, we build a physically reasonable framework within which we analyse the role of memory on the task of cooling. The first thing we need is a concrete notion of temperature.

3.1 Notion of Temperature

Cooling quantum systems can have many meanings, for instance decreasing its energy or entropy (purity), or for states that remain in thermal equilibrium with some bath, to decrease their temperature [2]. In general, these notions do not necessarily coincide; this is especially true for non-equilibrium dynamics or high-dimensional systems, both of which we want to include in our considerations. To see their departures, note that if a state is not in thermal equilibrium with a bath, such a standard thermal temperature cannot always be assigned. Things become even trickier when correlations are involved: for instance the overall state of some system and its environment can be entangled, potentially leading to local states that look like thermal states (at some non-zero temperature), even though the joint system is pure. Moreover, a decrease in entropy does not necessarily mean populating the lowest energy eigenstates more heavily, as the entropy is simply decreased by increasing the purity, i.e., compressing populations into any smaller number of eigenstates. Similarly, a decrease in average energy can also lead to an increase of entropy. Importantly, even though these notions differ, they have one property in common: they are all Schur concave/convex functions of the eigenvalues of the state. This fact allows us to unify the notions via a concept of temperature based on majorization theory:

Definition 1. For vectors with components sorted in decreasing value: A vector $\vec{a}$ majorizes another vector $\vec{b}$:

$$\vec{a} \succ \vec{b} \quad \text{iff} \quad \sum_{i=1}^{n} (\vec{a})_i \geq \sum_{i=1}^{n} (\vec{b})_i \quad \forall \quad n = 1, \ldots, d. \tag{1}$$

Based on this definition: we say that a state ρ_C is cooler than a state ρ_H, if the respective vector of eigenvalues $\vec{v}_C$ of ρ_C majorises that $\vec{v}_H$ of ρ_H. Based on this definition we mostly consider cooling to correspond to increasing the ground state population, i.e. satisfy the first partial sum in majorization for the vector of energy eigenvalues. Nonetheless, our results generalise. Qubits are a trivial case, i.e. majorization reduces to only the ground state population.

3.2 Open Quantum Systems

A target quantum system S of dimension d_S is considered with a corresponding, normalised ($\mathrm{Tr}[\rho] = 1$), positive semi-definite density matrix ρ_S. In general, no system is isolated, ie. there is always some form of interaction with the environment, which we have to take into account [34]. Therefore, we define an environment E with dimension d_E and density matrix ρ_E. We consider the joined arrangement of system and environment SE to be closed (for all intents and purposes) and so it evolves unitarily. If we further make the assumption that the initial system can be prepared independently of the environment (i.e. no initial correlations), then we can describe the state of the system of interest at time t via:

$$\rho_S^{(t)} = \mathrm{Tr}_E[U^{(t)} \rho_S^{(0)} \otimes \rho_E^{(0)} U^{(t)\dagger}]. \tag{2}$$

We can express this evolution in the form of dynamical maps $\Lambda^{(t,0)}$, which translate our system of interest forward in time:

$$\rho_S^{(t)} = \Lambda^{(t,0)}[\rho_S^{(0)}]. \tag{3}$$

In order to cool we wish to control the interactions between system and environment and therefore have to ask: Which kind of operations can we perform?

The most general form of interaction is a unitary operation followed by tracing out the environment. This can be described by a *CPTP-map*, a completely positive and trace preserving map. This allows to get from one valid state, meaning a positive semi-definite density matrix, to another, while keeping the normalisation, the trace of the density matrix, constant. Complete positivity assures, that this operation can be co-performed with an identity operation on a bigger quantum system and still ends up with a valid state.

This is our preliminary basis of open quantum systems. Next, we want to impose constraints on the states and interactions.

3.3 Thermal States

The framework of open quantum systems is too general for the task we wish to consider. For instance, cooling is trivial if the environment is in a pure state and the unitary simply swaps this with the system. Thus, we need some further restrictions.

The first constraint we place regards the initial state of the environment. To circumvent such scenarios as the one just described, note that typical environments found in nature are not described by pure states, but rather thermal states that indicate a long-time equilibration with its own larger environment in term. Such a state is described by a *Gibbs / thermal state* [8, 25]. This means that if we were to let a number of systems interact freely among each other, we would, after an appropriate amount of time, *typically* end up with a thermal state describing the ensemble. This physical motivation justifies our choice in restricting the initial state of the environment to be a *Gibbs/ thermal state*, $\tau(\beta)$, at some inverse temperature $\beta = \frac{1}{k_B T}$ (with k_B being the Boltzmann constant, which we set to 1 for the rest of this thesis). Thermal states, or Gibbs states, of dimension d are a one parameter family of states defined in terms of their Hamiltonian: $H = \sum_i E_i \left|i\right\rangle \left\langle i\right|$

$$\tau(\beta) = \sum_{i=0}^{d-1} \frac{e^{-\beta E_i}}{Z} \left|i\right\rangle \left\langle i\right|, \tag{4}$$

where $Z = \sum_{i=0}^{d-1} e^{-\beta E_i}$ is the partition function. For a Hamiltonian whose energy levels are ordered in a non-decreasing fashion, a Gibbs state is diagonal in the energy eigenbasis with non-increasing eigenvalues. A quantity that will become important later on is the maximum energy gap ϵ_{max}, which is the difference of the biggest and smallest energy level values of the Hamiltonian. I.e. for a zero ground state energy, ϵ_{max} is equal to the greatest energy level.

Thermal states are passive states, meaning their energy cannot be lowered by any unitary U acting solely on the state:

$$\mathrm{Tr}\left[H(\tau(\beta) - U\tau(\beta)U^\dagger)\right] \leq 0 \quad \forall \quad U. \tag{5}$$

Additionally, thermal states are completely passive. This means that even given an arbitrary number of copies, it is not possible to extract energy out of them using unitaries [26].

As such, thermal states are considered a free resource [7] and are used as a starting point for all quantum systems in this thesis. The target system and environment are both initially thermal states at the same, finite, non-zero temperature. This choice guarantees initial resource (temperature) balance and prevents trivialising the problem in the form of initialising a subsystem in a pure state and performing a swapping unitary to instantly cool to absolute zero. Therefore, the open evolution of Eq. (2) is restricted to initially thermal states:

$$\rho_S^{(t)} = \mathrm{Tr}_E[U^{(t)}\tau_S^{(0)}(\beta) \otimes \tau_E(\beta)U^{(t)\dagger}]. \tag{6}$$

The task at hand is now: We are given thermal states $\tau(\beta)$. Which interactions can we and do we want to implement in order to optimally cool the system? The restriction to initial thermal states narrows the interactions we can perform to a subset of CPTP maps. Namely those, that add a thermal environment, perform a unitary and trace out the environment. The available interactions can also be modelled via a dynamical map Λ, such that the system after one interaction is given by:

$$\rho_S^{(t)} = \Lambda^{(t,0)}[\tau_S^{(0)}(\beta)] := \mathrm{Tr}_E[U^{(t)}\tau_S^{(0)}(\beta) \otimes \tau_E(\beta)U^{(t)\dagger}]. \tag{7}$$

We consider such a dynamical map in the *coherent control* scenario, where an external coherent battery or control field supplies energy and we are not limited to energy-conserving unitaries. However, if we further restrict the operations to only energy-conserving unitaries, this would only allow us *thermal operations*, which we consider for the *incoherent control* scenario. Apart from this differentiation, we can later on consider further properties of the unitaries regarding how multi-partite they are.

3.4 Individual Cooling Cycles

After initialising all systems in thermal states at room temperature T_R (inverse room temperature β_R), a single cooling cycle is performed: We bring the target system S into contact with the environment E, implement an arbitrary cooling unitary and afterwards discard the environment.

While unitaries do not change the global set of eigenvalues of the joint system-environment SE arrangement, they do change their order. Although our notion of temperature based on majorization is necessarily blind to such re-orderings, a joint unitary can locally change the eigenvalues of S, which has implications for its temperature. In order to cool our system, we want to implement the global unitary on SE that reorders the eigenvalues in such a way that S', after the interaction, is colder than before. In other words, we want to find a unitary such that that the output state ρ'_S majorizes the input state ρ_S [1, 2].

As such, the task becomes one of reordering to achieve the optimal distribution of eigenvalues for the S subsystem. Ideally, the largest eigenvalue of SE is to be assigned to the lowest energy eigenspace of the system, and the next largest in the next lowest energy eigenstate, and so on. Such a protocol can always be performed, if the set of unitaries is not restricted, which we consider in our coherent control regime. However, any unitary can be made energy-conserving by considering an appropriate auxiliary system to compensate for the required energy. The use of thermal states and energy-conserving unitaries leads us exactly to thermal operations, which are typically considered in resource theory of quantum thermodynamics [12, 13]. The source of free energy then is a thermal bath at a higher temperature T_H. I.e. in order to be able to cool, the environment is allowed to be thermalised at the higher temperature T_H. We classify this as the *incoherent control* regime. Otherwise, one step follows the same concept as in the coherent control regime.

We have now completely defined one cooling cycle. However, like HBAC [28–30, 52] and so many other schemes [32] (e.g. quantum heat engines [7, 53], quantum batteries [54], ...) we want to consider many successive cycles. This also allows us to compare different regimes over time and resource consumption, as we will see later. When we look to compose many steps of such operations, we have a number of choices to consider, the principle of which is to allow for temporal correlations between operations (i.e. memory) or not.

3.5 Markovianity

We have outlined a single cycle of the cooling process. Now, we are interested in the performance of many successive implementations. This gives us the choice to make each cycle independent of the previous one or correlate them.

The former leads to a Markovian process, which can be described by multiple evaluations of Eq. (6). This leads to a concatenation of independent dynamical maps $\Lambda^{(i,i-1)}$ $(i = 1, ..., t)$, which take the system from step $i-1$ to step i. As such, the target system after t steps is given by:

$$\rho_S^{(t)} = \Lambda^{(t,t-1)} \circ \Lambda^{(t-1,t-2)} \circ ... \circ \Lambda^{(1,0)}[\tau_S^{(0)}(\beta)]. \tag{8}$$

The dynamical maps / interactions can be *adaptive*, meaning the unitaries are allowed to be different at each step. For instance, when complete control is allowed and one wishes to optimally cool the system, the unitary that assigns the largest eigenvalues to the system's lowest energy eigenstates *at each step* will depend upon the present joint state of the system and environment, resulting in turn in a different dynamical map between each time-step. In other cases of interest, one may wish to rather cool in a more autonomous fashion, by implementing a single unitary repeatedly, which would lead all of the dynamical maps above to be identical [52]. We will refer to this as a *non-adaptive* scheme, for which Eq. (8) reduces to a t-times concatenation of the same dynamical map. In general, this thesis focuses more on understanding the ultimate limitations of cooling with memory, so we mainly consider the more general adaptive setting.

Since the environment is reset after each cycle, it cannot carry any information forward through the process. The fundamental limitations for such a Markovian cooling process are already well understood, as analysed in two papers of *Clivaz et al.* [1, 2], which identify an upper bound for the final ground state population, for arbitrary machine dimension and both control scenarios, which is only dependent on the maximum energy gap of the machine spectrum, ϵ_{max}:

Theorem 1 *[1, 2] For any machine and control paradigm, in the limit of infinite cycles, the ground state population is upper bounded by :*

$$p_0^* = \frac{1}{\sum_{n=0}^{d_S-1}(e^{-\beta_R \epsilon_{max}})^n}. \tag{9}$$

Clivaz et al. state the optimal protocol for qubits as follows: For the incoherent control regime, a qubit target system S and a one-qubit machine M is joined with a one-qubit ancilla, A. Translated into our frame, the machine and ancilla correspond to what we call the environment. The energies are chosen to fulfil the resonance condition $E_A = E_M - E_S$, in order to create degeneracies in the joint system's energy levels. The ancilla is thermalised to T_H and together with the machine brought into contact with the system. The energy conserving cooling unitary is a swap operation, which exchanges the energy eigenvalues of energy levels $|010\rangle_{SMA}$ and $|101\rangle_{SMA}$, which is the only degenerate subspace relevant for cooling. The machine and ancilla are traced out and rethermalised using the hot bath and room temperature respectively and the cycle starts anew. Although the protocol is non-adaptive, it is possible to reach the same asymptotic limit as an adaptive protocol. However, adaptive interactions result in finite time advantages, as we will see later. With repeated Markovian cycles, and $T_H \to \infty$, the asymptotic limit of the ground state population is given by Eq. (9) [1, 2].

In the coherent control regime, the adaptive protocol simply optimally reorders the eigenvalues, as already described. The optimal non-adaptive protocol swaps the eigenvalues of the smallest population of the subspace of the ground state of the target qubit $|011\rangle_{SMA}$ with the largest population of the subspace of the excited state of the target qubit $|100\rangle_{SMA}$. The adaptive protocol performs more swaps and therefore has a higher rate of cooling, although both protocols ultimately converge to an upper bound of the target's ground state population proportional to the largest energy gap ϵ_{max} of the joined machine-ancilla arrangement [1, 2].

Here, we are interested in extending the protocol to the non-Markovian case. Allowing for arbitrary non-Markovianity would trivialise the problem by being able to almost instantly cool to absolute zero. For instance, one feature of the most general form of memory is the presence of arbitrary initial system-environment correlations: then if the joint state is almost pure, but highly entangled, a single operation would enable almost perfect cooling of a local system that initially looks extremely hot. Therefore, we keep the restriction of our systems being initially uncorrelated. Instead, we allow for temporal correlations between the time steps by retaining some parts of the environment that have previously interacted with the system, to take place in later interactions. We can analyse such a scenario using a flexible generalised *collision model* that allows for some physically reasonable memory, without trivialising the problem.

3.6 Collision Models

Now, we have already put restrictions on the system, the environment and the form of operations. We have seen the corresponding Markovian setting and results; we now wish to consider the task of cooling when memory is present. The overall Markovian dynamics can be regarded via a microscopic model in the following way: Consider breaking up the environment into an arbitrary number of identical quantum systems, here called *machines* M, of dimension d_M. In a given step, the target system is brought into contact (collides) with a chosen number of machines, the unitary interaction is performed and afterwards the systems separate. The machines that have interacted with the system are traced out, and in the next cycle a number of *fresh* machines interact with the system, and so on. The tracing out and replacing with fresh machines models a complete rethermalisation of the machines with the environment, and so no knowledge of past interactions is stored in the machines, leading to Markovian dynamics. This concept used here is the standard *collision model* picture which provides a microscopic model for Markovian dynamics (as it was originally defined for [16, 35, 55]). A schematic of such a process can be seen in Fig. 1.

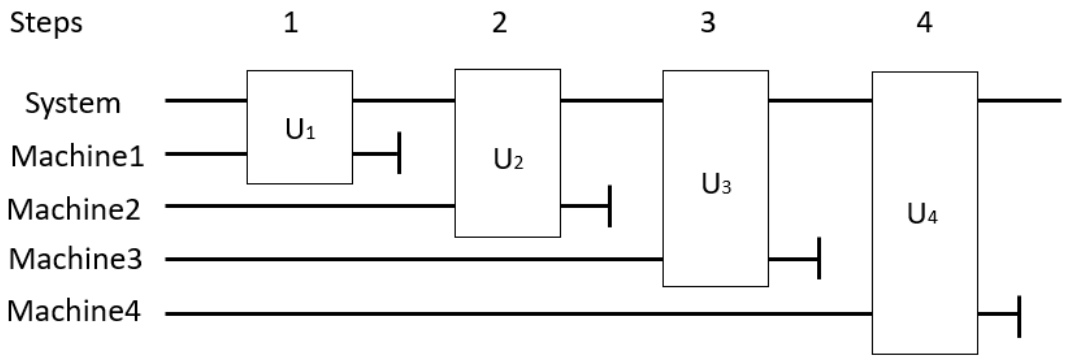

Figure 1: *Schematic of a Markovian collision model.*
A box, labelled with U_i, represents the i-th unitary operation on the incoming systems. The system is brought in contact with a machine, a unitary interaction is performed and the machine is traced out (terminated line). Then the system comes into contact with the second machine and so on.

A number of assumptions built into this standard collision model ensure that the process is Markovian, i.e. the dynamical maps are independent of each other: There must not be any i) initial correlations, ii) machine-machine interactions (like, e.g., in [17]) and iii) repeated system-machine interactions. This makes the model very practical for our endeavour, as we can introduce various types of non-Markovianity by breaking some of these assumptions, without ending up with completely general non-Markovian dynamics, allowing us to make reasonable comparisons to their Markovian counterparts.

3.7 Non-Markovianity

Now we want to extend the collision model to consider non-Markovianity. Since we want to cool our system, we would like to increase its ground state population via interactions. While repeated interactions with only the same machine is ineffective, a combination of repeated interactions and fresh machines seems reasonable. In other words, we allow some of the machines to feed forward and interact with the system later, acting as memory carriers, instead of completely rethermalising, as illustrated in Fig. 2. The idea for where a possible advantage might come from is that we wish to draw further value from the already used machines and use them as storage to compress extra population into the system's lowest energy eigenstates. Instead of being traced out, l of them take part in the next interaction. The *memory length*, or depth, l denotes how many machines are fed forward to the next cycle. For small memory lengths, the dynamics is tractable as we will see through the *Markovian embedding* that we use later. Using this extended model we can further specify: i) The effective unitary size, k, stands for how many machines are interacted with at each step; ii) The total number of machines, that have been interacted with, or are at disposal, throughout the entire protocol is denoted by m. For example for Fig. 2 the memory length is $l = 2$ and $k = 3$, because the system interacts with 3 machines at each step, of which two are fed forward to the next cycle. Since the overall protocol requires 6 machines to implement the 4 unitaries, the full circuit shown has $m = 6$. Therefore, the applied unitaries $U_{1,2,3}$ correspond to the following systems interacting: $SM_1M_2M_3$, $SM_2M_3M_4$ and $SM_3M_4M_5$, respectively. As such, the target system interacts with $k - l$, here 1, new machines at each step. Just as in the Markovian case, in the non-adaptive scenario all unitaries are the same: $U_i = U_j \quad \forall \quad i, j$, while for the adaptive protocol, these can all be different. We classify all scenarios in this thesis by these parameters. Trivially, a memory length of $l = 0$ corresponds to the Markovian case.

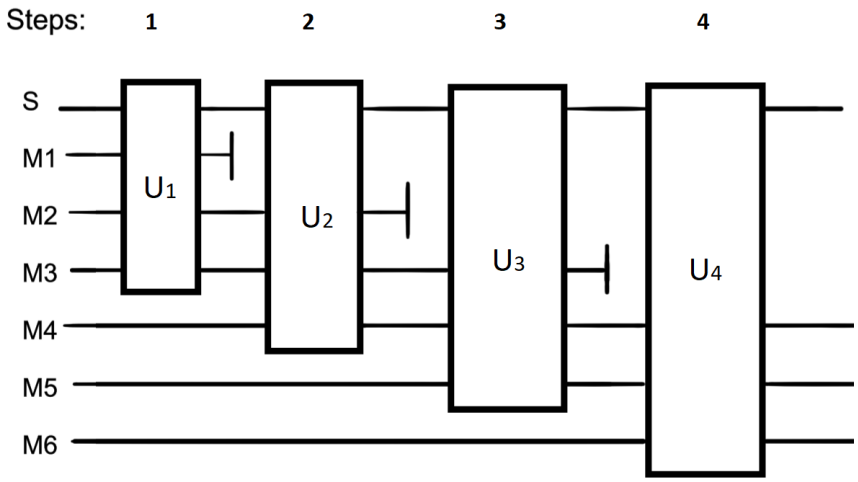

Figure 2: *Schematics of a collision model with memory.*
The shown protocol has a unitary size $k = 3$ and memory length $l = 2$. A box, labelled with U_i, represents the i-th unitary operation on the systems. A terminated line means that the machine is traced out. The system interacts with three machines at each step and two of them are being reused to cool the system, which leads to memory effects.

This framework regarding memory carriers has some intrinsic restrictions, as we now outline: For a fixed k, not all values of l are possible, as a given unitary size limits the maximally possible memory length. We cannot feed more machines forward than are being interacted with at each step. A similar reasoning follows in the opposite direction. A given memory length implies interaction with at least the same amount of machines and therefore minimum effective unitary size. Additionally, if the memory length is equal to the unitary size, the target system will interact with the same machine(s) over and over, which yields no further improvement once the machines have been depleted. Therefore, the constraint $l < k$ is imposed.

Similarly, in order to use a fixed number of machines the amount of steps n should be an integer, i.e. : $(n-1) = \frac{m-k}{k-l}$. For example with a unitary size of $k = 5$ and a memory length of $l = 2$ it is not possible to use up exactly 10 machines without breaking protocol, as m for the first few steps is: $m = k + (k-l)(n-1) = 5 + (5-2)(n-1) = 5, 8, 11, 14, \ldots$. We can look at the ultimate final ground state population regarding these two values (n and m), which generally will give us different results. However, in the limit of asymptotically infinite cycles these coincide. Otherwise, for finite times they are a good means of comparison, as we will discuss next.

As there are many parameters classifying the possible regimes, a simple comparison regarding the final ground state population does not sufficiently put the cooling protocols into relative perspective. Thus, it is necessary to set a scale which is the same for all regimes. Two ideas come to mind: Firstly, the resource theoretical view, which looks at the final ground state population after a fixed number of machines m has been interacted with. This comparison focuses on how well different regimes can utilise their limited resources, but disregards the number of steps n, or number of unitary operations performed. Secondly, a time- or step-dependent comparison focuses on comparing the final ground state population of the system after the same finite amount of interactions, steps n have taken place in order to contrast the different speeds or cooling rates, which is important if an abundant amount of machines are present.

4 Results

In order to approach the problem at hand, note that the particular type of non-Markovian cycle that we propose here can be rearranged to always reuse the same l machines, which we identify as *memory carriers*, L. Subsequently the cooling target S can be extended to include the target system as well as the l memory carriers, denoted by SL, which together interact with $k - l$ new machines at each step. This is called a *Markovian embedding*, as the non-Markovian protocol can be represented via a Markovian interaction on a finite, bigger space, while still remaining tractable for a small number of memory carriers: We only need to track the states of the target system and the l memory carriers, rather than the full system-environment dynamics (as one would for a general non-Markovian evolution). Fig. 2 can be slightly modified in order to visualise the Markovian embedding, as shown in Fig. 3.

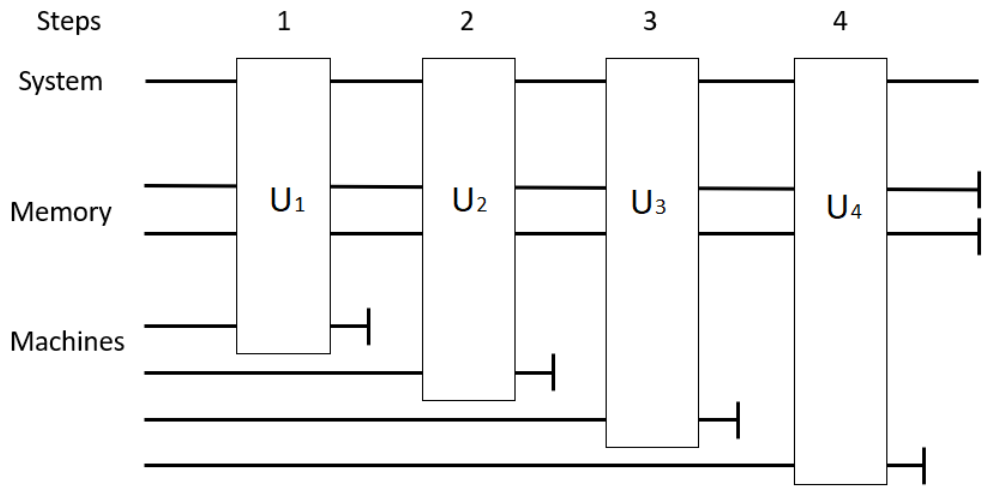

Figure 3: *Markovian embedding visualised.*
A cooling protocol for a unitary size $k = 3$ and memory length $l = 2$. Boxes labelled U_i represent unitary operations on the incoming systems. A terminated line means that the machine is traced out. Machines M1 and M2 are the memory of the process, i.e. we always feed these same machines forward and trace them out only at the very end, as such the system+memory interacts with only 1 fresh machine at each step.

4.1 Asymptotic Cooling Behaviour

An upper bound regarding the non-Markovian cooling regime has been presented by Taranto et al. [3], where it is shown that memory enhances the attainable ground state population, compared to the Markovian case, exponentially in the number of memory carriers:

Theorem 2 *[3] For any d_S-dimensional system interacting at each step with k identical d_M-dimensional machines, with l of the machines (labelled L) used at each step carrying the memory forward, in the limit of infinitely many cycles:*
i) The ground state population of S is upper bounded by

$$p^*(\epsilon_{max},\beta,k,l) = (\sum_{n=0}^{d_S-1} e^{-\beta n d_M^l (k-l)\epsilon_{max}})^{-1}. \tag{10}$$

ii) The vector of eigenvalues of the output system state is majorized by that of the following attainable state:

$$\rho_S^*(\epsilon_{max},\beta,k,l) = \sum_{n=0}^{d_S-1} \frac{e^{-\beta n d_M^l (k-l)\epsilon_{max}})}{\mathbb{Z}_S(\beta, d_M^l(k-l)\epsilon_{max})} |n\rangle\langle n|_S\,, \tag{11}$$

whenever the initial state $\tau_S(\beta)\otimes\tau_M(\beta)^{\otimes l}$ is majorized by

$$\rho_{SL}^*(\epsilon_{max},\beta,k,l) = \sum_{n=0}^{d_{SL}-1} \frac{e^{-\beta n (k-l)\epsilon_{max}})}{\mathbb{Z}_{SL}(\beta, (k-l)\epsilon_{max})} |n\rangle\langle n|_{SL}\,. \tag{12}$$

Here $\mathbb{Z}_X(\beta,\epsilon) := \sum_{n=0}^{d_X-1} e^{-\beta n\epsilon}$ is a quasi-partition function, depending only upon the maximum energy gap [3]. This exponential advantage originates in the use of memory, as these added memory systems allow for a larger space to compress population into, compared to only one target system. Although the system interacts with l-less fresh machines each step, i.e. with $k-l$ fresh machines, the repeated interactions with l memory carriers leads to the inclusion of a factor d_M^l, which scales exponentially with the amount of memory carriers. In other words: There is a vast improvement for already a small amount of memory carriers (see e.g. Fig. 7). Furthermore, the best way to improve the upper limit of a protocol is to increase the amount of memory carriers involved.

Following Theorem 2 we wish to compare the ultimate achievable asymptotic state across various memory structures, i.e., different values of k,l, and different initial settings, i.e., different values of β,ϵ_{max}. This hierarchy of cooling regimes can be determined by [3] :

$$\rho_S^*(\epsilon_{max},\beta,k,l) \prec (\succ) \rho_S^*(\epsilon'_{max},\beta',k',l') \quad \text{if} \quad \beta(k-l)d_M^l\epsilon_{max} \leq (>)\beta'(k'-l')d_M^{l'}\epsilon'_{max}. \tag{13}$$

The first thing to note here, is that in the Markovian case ($l=0$), the unitary size k determines the ranking. However, when we introduce memory carriers, this is no longer the case, and protocols with memory, but smaller unitary size can yield better results than a Markovian regime with a greater unitary size (e.g. $k=7$, $l=0$ is worse than $k=4$, $l=3$). Although memory is very important, a greater memory length can also do worse than a lesser one (e.g. $k=4$, $l=1$ is better than $k=3$, $l=2$). As such the hierarchy has to be evaluated for every single regime in question to determine which protocol is best.

We present simulations that agree with the statements here in the following plots (Figs. 4, 5, 7). All following plots have been calculated using one qubit as a target system and single-qubit machines, $d_M=2$, unless stated otherwise. Moreover, since it has already been shown that non-adaptive protocols have a lower cooling rate compared to adaptive ones [3], the main focus of our comparison here concentrates on the latter. I.e. the plotted cooling protocols always optimally reorders all possible eigenvalues available to it such that the largest are placed in the lowest eigenstates of the system and memory carriers, thereby dissipating as much heat as possible into the machines that are discarded.

4.1.1 Coherent Control Scenario

We first compare the final ground state populations in the asymptotic limit of infinitely many cycles for various memory structures. Note that, in contrast to the finite-time comparisons we will make later on, here it does not matter whether we compare these scenarios by considering the number of total machines used m or the number of steps taken n, as they are directly proportional to each other and we are only interested in the limit. Fig. 4 compares the ground state populations of various coherent, adaptive non-Markovian cooling regimes for up to 100 steps. An external coherent battery provides free energy, hence unitaries are not restricted in any way and the adaptive scenario allows for a different unitary at each step. We expect that after 100 steps, the cooling protocols approximately reach their steady state and their ground state populations have reached their limit. The protocol operating with $k = 3$ and $l = 2$ has the worst performance of the sampled regimes and becomes constant after approximately 50 steps. Meanwhile, for values $k = 4$ and $l = 3$, the protocol has a much better performance, but takes 100 steps to reach this limit. Although most of the cooling curves have cross-overs, the final ground state populations follow Theorem 2, which depends only on a few parameters, i.e. ϵ_{max}, k, l. However the short term behaviour depends on the microscopic detail, which we will look at later in this chapter.

Another interesting comparison to make is between a memory-less protocol with a large unitary size (say $k = 7$) and protocols with smaller effective unitary size (e.g., $k = 3, 4, 5$) but that have some memory. We see that when the memory length in the latter cases is small, the Markovian scenario can still asymptotically outperform the case with memory (e.g., compare $k = 7, l = 0$ with $k = 3, l = 2$). However, for larger memory lengths, the protocols with memory typically outperform any of those without, even when the effective unitary size is smaller (e.g. compare $k = 7, l = 0$ with $k = 4, l = 3$). This demonstrates, that for the task of asymptotic cooling, memory seems to be a more important resource than the effective interaction size (although, of course, the former is fundamentally bound by the latter).

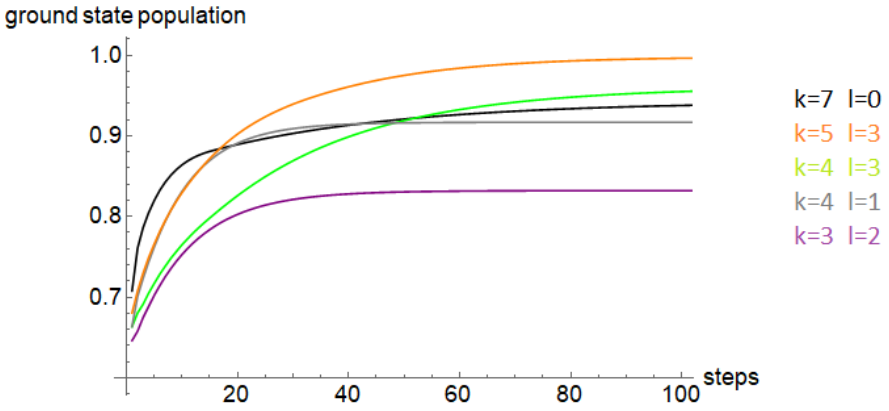

Figure 4: *Various coherent adaptive cooling curves.*
The ground state population is plotted as a function of steps for the single-qubit machine $d_M = 2$, $\epsilon_{max} = 2$ and different cooling protocols. Incorporating memory carriers can result in better cooling performances than a memoryless protocol with a bigger interaction size.

Next, we analyse the role of memory *for a fixed* unitary size, k. In Fig. 5, we consider the same coherent, adaptive Markovian cooling protocol with a unitary size of value $k = 7$ in comparison to its non-Markovian counterparts with the same unitary size, but with memory carriers. Fig. 5 shows that for larger memory lengths, here for $l \geq 4$, the calculated ultimate ground state population is very close to 1 and beyond $l = 2$ all non-Markovian regimes bunch together. Meanwhile, the Markovian regime in comparison is very much worse, i.e. reaches a ground state population of only around 0.94. This shows, that for a sufficiently large interaction size, there exists a corresponding sufficiently large memory length, for which the protocol already has an optimal cooling performance and adding any more memory carriers does not lead to significant improvements. It is also interesting to note the qualitative behaviour of the rate of cooling, i.e. how quickly a given memory structure achieves its steady state. Here we see, for example, that the greatest memory length $l = 6$ takes approximately 200 steps to reach its limit, whereas a memory length of $l = 2$ takes around 60 steps.

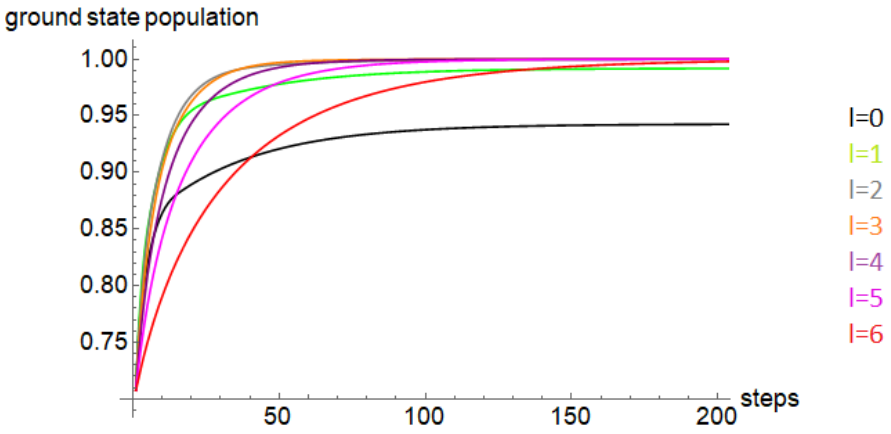

Figure 5: *Comparison of Memory lengths*
The ground state population is plotted as a function of steps for a unitary size of $k = 7$, $\epsilon_{max} = 2$ and different memory lengths. Incorporating any number of memory carriers, while keeping the interaction size the same, results in an immense cooling advantage.

4.1.2 Incoherent Control Scenario

After studying the coherent control regime, we now turn our attention to the incoherent control scenario. In the non-Markovian setting, The incoherent protocol follows that outlined for the Markovian case in Clivaz [1, 2]: To every machine qubit, an ancilla qubit A is added, such that the desired cooling transformation is rendered energy-conserving. In other words, the resonance condition $E_S = E_M - E_A$ is fulfilled, which creates degeneracies in the system's Hamiltonian that can be used for cooling in an energy conserving manner. We adapt this ancilla qubit for all of our machines and memory carriers as well. Here, the energies are chosen to be: $E_S = 1$, $E_M = 2$ and $E_A = 1$; subsequently $\epsilon_{max} = 3$. Before each cycle, the (non-memory) ancillas are thermalised at a lower inverse temperature $\beta_H = 0.000001$ than inverse room temperature $\beta_R = 0.2$, as such thermal contact is what provides the energy source to implement the required transformations in the incoherent setting. The adaptive cooling protocol is then to perform all available energy-conserving unitaries that lead to cooler system+memory carriers SL at each step. We simulate the cooling behaviour for a number of different memory structures in the incoherent scenario and plot the results in Fig. 6 for 200 steps. We also plot a comparison of different memory lengths for the same unitary size ($k = 6$) in Fig. 7.

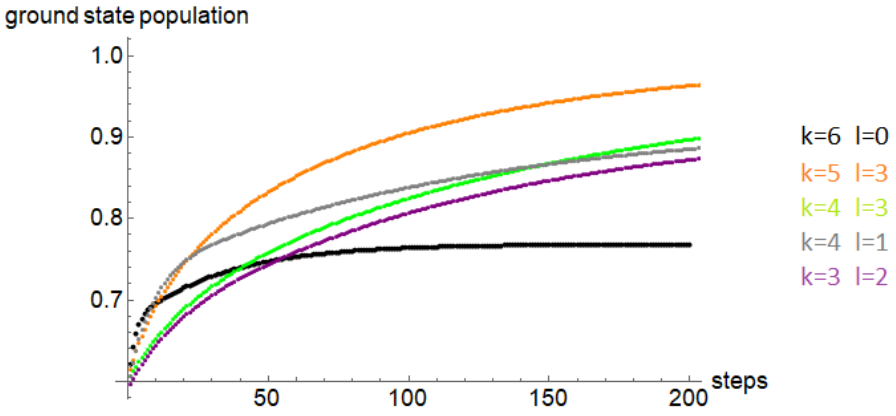

Figure 6: *Incoherent cooling protocols*
The ground state population is plotted as a function of steps for various incoherent adaptive cooling protocols with $\epsilon_{max} = 3$. Here we see that protocols with memory can have better cooling results than a Markovian protocol with a larger interaction size.

Firstly, note, that the curves are much flatter compared to the coherent control scenario. This is because we only allow for energy conserving unitaries and therefore there are less available interactions that lead to cooling at each step, resulting in much flatter cooling curves than in the coherent case,

where arbitrary reordering of the eigenvalues are permitted. subsequently, the cooling curves have not reached their steady state even after 200 steps. Nonetheless, we know that the final ground state population follows Theorem 2. Secondly, the difference in performance between Markovian and non-Markovian is even more apparent: All of the sampled non-Markovian regimes achieve better cooling than the memory-less (Markovian) protocol. interestingly, the cooling curves in purple ($k = 3$, $l = 2$) and grey ($k = 4$, $l = 1$) will lead to the same final ground state population, but have different short term behaviours, i.e. the grey curve performs better. This also goes to show that the short term behaviour in contrast to the upper limit is very much dependent on the microscopic details of the interaction.

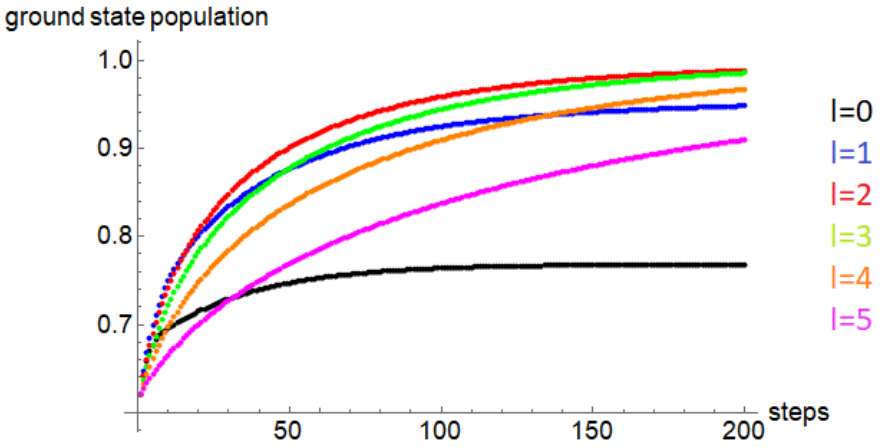

Figure 7: *Comparison of memory lengths*
In the incoherent control regime: The ground state population is plotted as a function of steps for a unitary size of $k = 6$, $\epsilon_{max} = 3$ and different memory lengths. We see very clearly that there is a performance gap between the Markovian and non-Markovian regimes. All shown protocols have the same unitary size, but the ones with memory have better cooling.

A very similar picture to the coherent scenario emerges in the incoherent control scenario when we compare protocols with the same unitary size but different memory lengths. There is a clear performance gap between the Markovian and non-Markovian regimes: any amount of memory outperforms the memoryless case. While for the domain of the plot we only show the first 200 steps, we know that (by Theorem 2), for a fixed unitary size the protocol with the greatest memory length has the highest ultimate ground state population and the rest follows in decreasing order, i.e. the the pink curve ($l = 5$) will eventually overtake all others.

We have now seen that memory carriers enhance the ultimate cooling limit in the limit of infinitely many cycles. Furthermore, this limit is only dependent on a small number of parameters. Contrary to that the finite time behaviour generally depends on the intricate microscopic details of the performed protocol. In the next section we have a closer look at the short term behaviour.

4.2 Finite-time Cooling Behaviour

We now compare several cooling regimes for finite times. While it has been shown that the asymptotic limits for infinite cycles follow a clear hierarchy, the short term behaviour of various cooling regimes may exhibit crossovers, as we have already seen in Figs. 4, 5, 7. This results from the fact that, while regimes with larger memory lengths have higher final ground state populations, they interact with and discard l less machines at each step, and as such dissipate energy to the environment slower, compared to Markovian protocols. Subsequently, protocols with less/no memory length interact with more machines and as such access more resources in a shorter amount of steps, which positively influences their short term behaviour. Because of this resource unbalance, most non-Markovian regimes (especially for large memory lengths) are initially outperformed by their Markovian counterpart, but eventually surpass them. On the other hand, if we fix the number of machines which are at our disposal, Markovian protocols take less steps than protocols with memory to complete the whole circuit, i.e. the more memory the more steps the protocol has to implement in order to use all available machines. Subsequently, in this section we compare the cooling rates for both control scenarios for both means of comparison; step-wise and machine-wise, to investigate the short term behaviour.

4.2.1 Coherent Control Scenario

We first examine the coherent control regime, an external coherent battery or control field supplies free energy to the arrangement, hence allowing for arbitrary unitaries to be implemented.

The finite time behaviour of large memory lengths can be seen in Figures 8 and 9, where the cooling curves of coherent, adaptive protocols where each machine consists of 2 qubits (i.e. $d_M = 4$) with unitary size of $k = 7$, $\epsilon_{max} = 2, 3$ respectively and memory lengths ranging from $l = 0$ (magenta) to $l = 6$ (blue) for the first 20 steps are plotted.

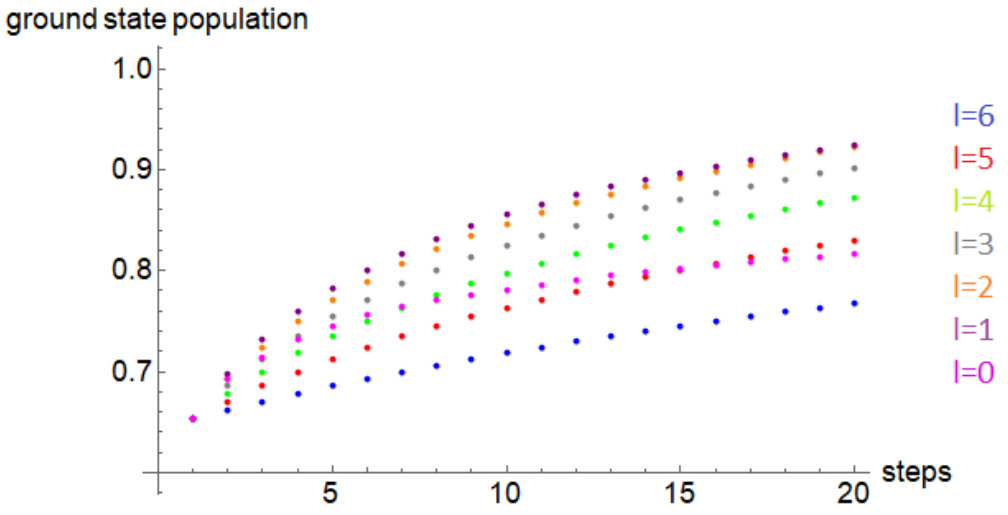

Figure 8: *Finite-time cooling behaviour*
The ground state population is plotted as a function of steps for a unitary size of $k = 7$, $d_M = 4$, $\epsilon_{max} = 2$ and different memory lengths. Protocols with more memory interact with less fresh machines per step and as such have an initial cooling disadvantage.

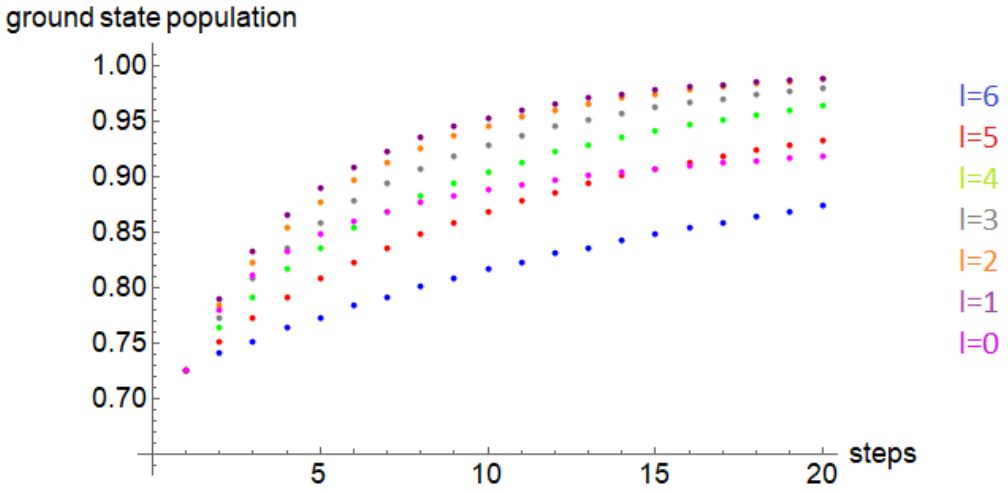

Figure 9: *Finite-time cooling behaviour 2*
The ground state population as a function of steps for a unitary size of $k = 7$, d_M=4, $\epsilon_{max} = 3$ and different memory lengths. An increase in maximum energy gap (compared to Fig. 8) increases the overall cooling performance.

For the chosen parameters it seems that: Beginning with the second step, the non-Markovian protocol's ground state populations arrange themselves in opposite magnitude of memory length, i.e. for the first 20 steps shown the protocol with the greatest memory length $l = 6$ has the lowest ground state population, whereas the regime with the smallest memory length $l = 1$ has the the highest. This

behaviour demonstrates the different rates with which protocols use their machine resources, which leads to different cooling rates. The memoryless protocol is somewhere in the middle, but very quickly gets overtaken by the non-Markovian regimes (in the domain of this plot for all but $l = 6$) in order of increasing memory length. Note, That this behaviour is the same for both Fig. 8 ($\epsilon_{max} = 2$) and Fig. 9 ($\epsilon_{max} = 3$), with the only notable difference being the overall better performance in the latter due to the higher maximum energy gap. Next we will analyse in more detail after how many steps memory holds an advantage over no memory.

The next question is about how quickly this performance gap between Markovian and non-Markovian protocols develops. To study this, in Fig. 10, we fix the machine dimension and effective unitary size, and plot the number of steps that it takes for the ground state population of the non-Markovian regimes to overtake the corresponding Markovian one, for each possible value of l. We analyse this for two choices of unitary size and three choices of maximum energy gap of the machine qubits. The amount of energy that can be dissipated at each step is proportional to the subspaces of machines to-be-discarded (= new machines interacted with). In other words, interactions with more fresh machines increase the initial cooling rate the quickest. In Fig. 10 a.), short memory lengths $l = 1, 2, 3$ take about 5 steps to outperform the memory-less cooling curve, but the greatest memory length $l = 6$ takes about 40 steps. In Fig. 10 b.), with a unitary size $k = 6$, the crossover points for all three sampled maximum energy gaps for the first 4 memory lengths coincide, before eventually diverging. This goes to show, that in some scenarios, a change in maximum energy gap can effect the performance of protocols with varying memory structures equally.

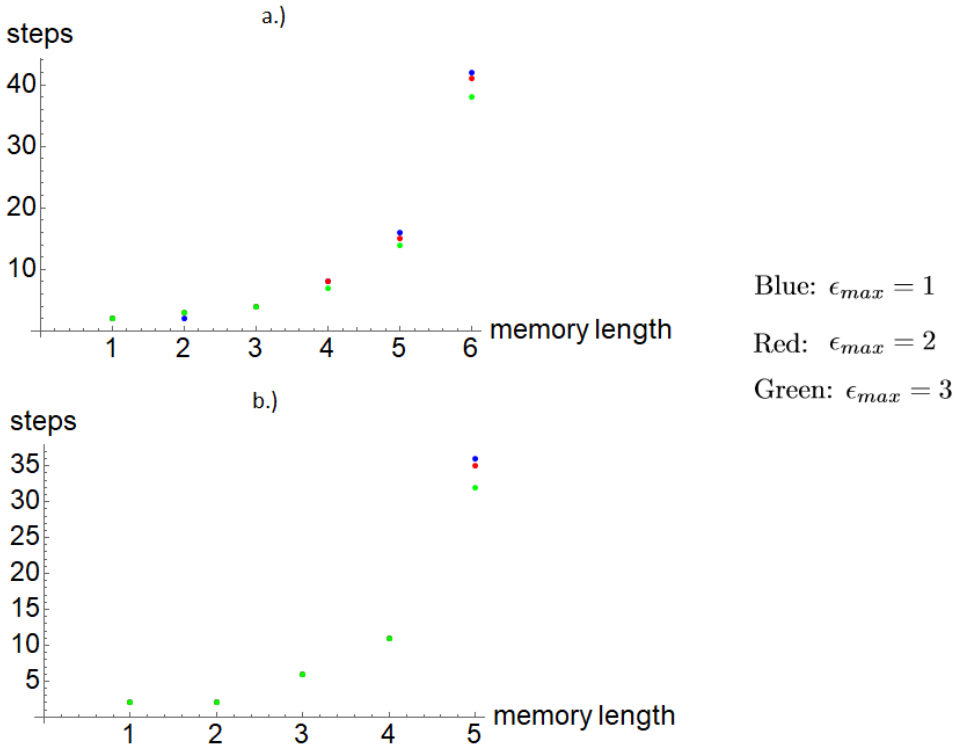

Figure 10: *The cross-over points of adaptive Markovian and non-Markovian cooling curves* The number of steps it takes for the adaptive non-Markovian regimes, with machine dimension $d_M = 2$ to have a higher ground state population than their memory-less $l = 0$ counterpart, as a function of their memory length l. For a.) $k = 7$ and b.) $k = 6$. Regimes with longer memory lengths interact with less fresh machines per step and therefore dissipate energy slower, i.e. they take more steps to outperform the Markovian cooling protocol.

Since this short term behaviour originates in the use of more machines, i.e. shorter memory lengths have access to more fresh machines per step, it is important to compare the ground state population to the total number of machines interacted with throughout the entire protocol. We show this in Fig. 11, where the same cooling regimes as in Fig. 8 are plotted, but with respect to their total machine utilisation, m, rather than the number of unitaries implemented, n.

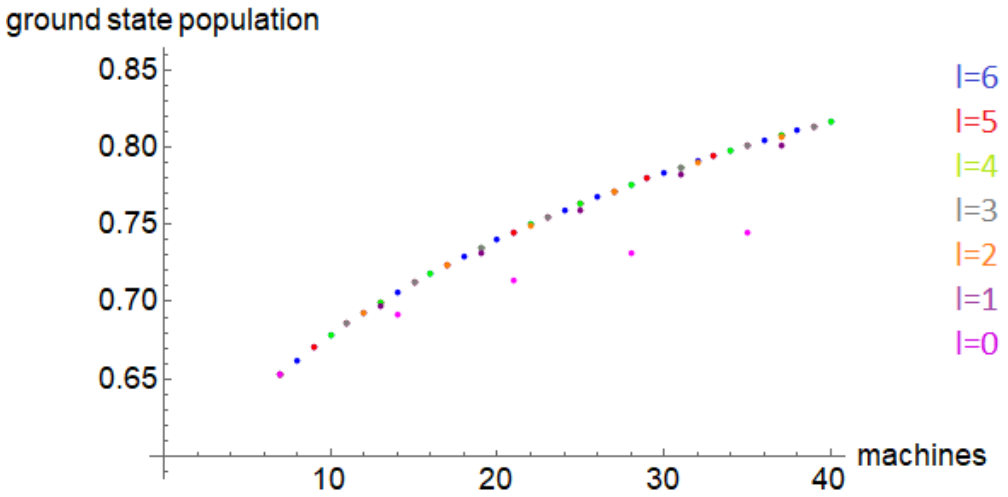

Figure 11: *Limited machine-resource comparison*
The ground state population is plotted as a function of total machines interacted with for a unitary size of $k = 7$, d_M=4, $\epsilon_{max} = 2$ and different memory lengths. Cooling protocols with memory can achieve a higher ground state population with less total machines used, compared to their memoryless counterpart.

We now see that protocols with larger memory lengths, especially $l > 1$, can better utilise a fixed set of resources to perform cooling. Although the non-Markovian regimes have different memory lengths and subsequently use up machine resources in different rates, the corresponding points in the plot follow the same path for short times, but eventually split in order of increasing memory length, leading to the asymptotic hierarchy as shown for fixed k in e.g. Fig. 5. Here, for 40 machines used all protocols with a memory length of $l > 1$ cool in the same way. In other words, they can dissipate energy by discarding machines to the same extend. Some regimes use up machines faster and as such have faster cooling in time, but their cooling-to-resource ratio is the same. Meanwhile, the Markovian protocol follows a different path already from the start and has a worse machine-wise cooling performance compared to the protocols with memory.

In order to better understand the finite times cooling behaviours, we want to change the other, unitary-independent, governing parameters. Especially, the machine dimension and the maximum energy gap, since it is still very unclear how they influence the short term cooling performance. Fig. 12 shows the ground state population for the adaptive, coherent control scenario with a machine dimension of $d_M = 2$ and maximum energy gap of $\epsilon_{max} = 2, 3$ respectively, for the total use of 40 machines.

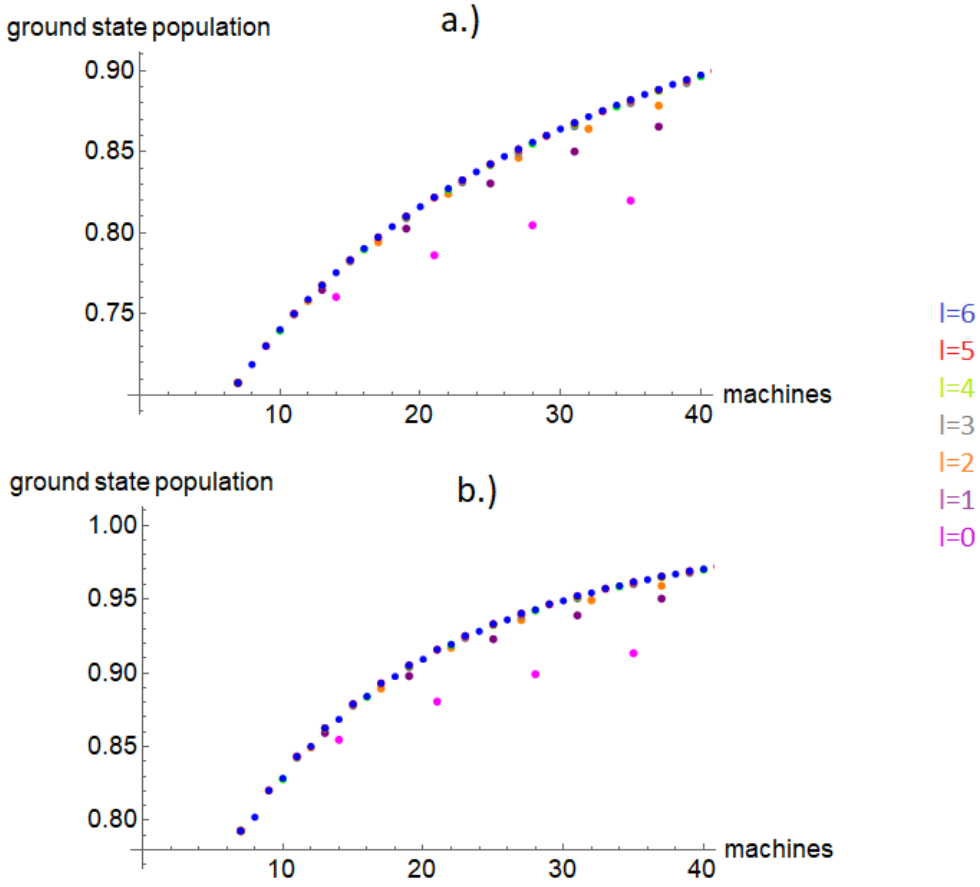

Figure 12: *Limited machine-resources comparison 2*
The ground state population is plotted as a function of total machines interacted with for a unitary size of $k = 7$, $d_M = 2$, $\epsilon_{max} = 2, 3$ for a.) and b.) respectively and different memory lengths. All non-Markovian protocols have a clear cooling advantage over the memoryless protocol, with respect to the total number of machines interacted with.

The overall behaviour in Fig. 12 is similar to that of Fig. 11, even though the machine dimension is decreased to $d_M = 2$. The non-Markovian protocol's curves split much earlier and there is a visible difference in ground state population after already 20 steps. Still, protocols with greater memory lengths, here for $l > 2$, overlap and follow the same curve. A change in maximum energy gap improves all protocol's general performance.

We have now seen the short term cooling behaviour of protocols with the same unitary size but different memory lengths and found, that if memory carriers are involved, there is an increase in machine-wise and step-wise performance, but more memory results in interactions with less fresh machines and as such slows down the cooling process. Next, we want to compare how protocols with different unitary sizes perform in the short term. Fig. 13 shows three Markovian protocols with different unitary sizes ($k = 7, 5, 3$) and their non-Markovian counterparts with a memory length of $l = 1$ each, plotted for the first 20 steps. The cooling curves follow the same outline as above (see Figs. 8, 9), i.e. A greater unitary size results in a better performance and the protocols with a memory length of $l = 1$ achieve better cooling than the corresponding memory-less ones. Furthermore, there are also crossovers of curves with different unitary sizes. For the chosen parameters it appears that though Markovian protocols with a greater unitary size can initially outperform non-Markovian protocols with smaller k, the non-Markovian curve can catch up after a relatively short amount of steps. (e.g. $k = 7$ $l = 0$ exceeds $k = 5$ $l = 1$ until step 12; not shown: $k = 7$ $l = 0$ exceeds $k = 6$ $l = 1$ until step 5)

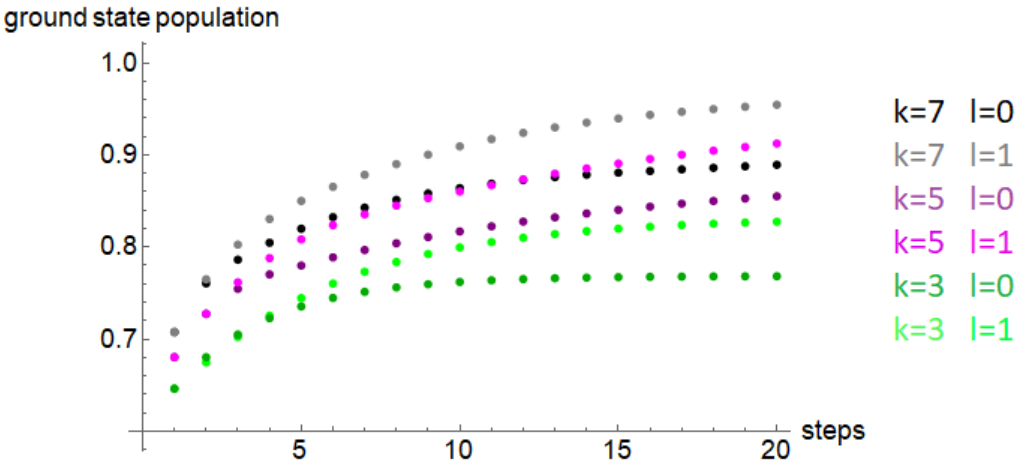

Figure 13: *Short term cooling behaviour*
The ground state population is plotted as a function of the first 20 steps for the single-qubit machine $d_M = 2$, $\epsilon_{max} = 2$ and different cooling protocols. Although a greater unitary size can initially have better cooling, a protocol with a smaller unitary size but with memory can overtake in a relatively short time (e.g.$k = 7$ $l = 0$ exceeds $k = 5$ $l = 1$ until step 12)

Next, we want to make another comparison, but regarding a limited amount of total machines interacted with. Fig. 14 demonstrates four Markovian protocols with different unitary sizes ($k = 3, 4, 5, 7$) and their non-Markovian counterparts with the greatest allowed ($k > l$) memory length respectively, plotted for 40 machines used throughout the protocol. It appears, the sampled Non-Markovian protocols with a unitary size greater than 3 (here: $k = 4, 5, 7$) have a very similar cooling performance and bunch together, with the greatest unitary size ($k = 7$) achieving the slightly better cooling, followed by the next greatest ($k = 5$). The Markovian regimes perform worse than the protocols with memory. I.e. the ground state population is only dependent on the unitary size and as such the smallest unitary size (here $k = 3$) has the worst cooling. However, Markovian protocols with large unitary sizes can achieve a very similar performance to non-Markovian protocols with smaller k (e.g. Markovian $k = 7$ and non-Markovian $k = 3$ and $l = 2$). However, for a slightly greater unitary size, this is no longer the case (e.g. $k = 4\ l = 3$ outperforms the Markovian $k = 7$). This goes to show, that non-Markovian protocols can outperform Markovian regimes with a greater unitary size in a comparison regarding finite machine resources.

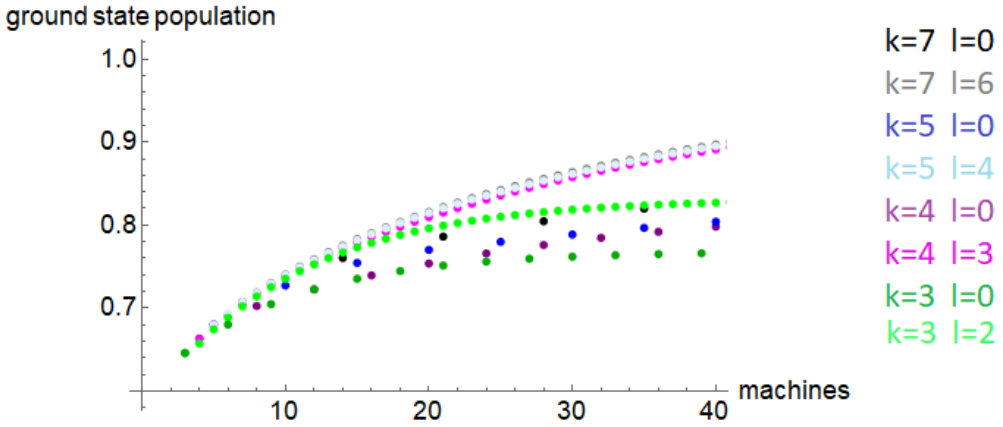

Figure 14: *Limited machine-resources comparison 3*
The ground state population is plotted as a function of total number of machines used for the single-qubit machine $d_M = 2$, $\epsilon_{max} = 2$ and different cooling protocols. We show, that Non-Markovian protocols can outperform a Markovian protocol with a bigger unitary size in a comparison regarding limited machine resources (e.g.$k = 3\ l = 2$ and $k = 5\ l = 0$).

4.2.2 Incoherent Control Scenario

Just as above for asymptotically infinite cycles, the cooling protocol follows that outlined by Clivaz [1,2]. Every machine consists of a machine qubit M and an ancilla A. The resonance condition ($E_A = E_M - E_S$) is fulfilled via the choice of: $E_S = 1$, $E_M = 2$ and $E_A = 1$, this fixes $\epsilon_{max} = 3$. The non-memory ancillas are allowed to thermalise to a hotter temperature between steps. Here we choose the inverse room and hot temperatures to be $\beta_R = 0.2$ and $\beta_H = 0.000001$ respectively. The adaptive protocol implements the optimal available energy-conserving cooling unitary at each step. Fig. 15 shows a close-up of Fig. 7 for the first 40 steps.

A picture very similar to the coherent control scenario (see Fig. 8 and Fig. 9) emerges: We know that larger memory lengths have a cooling advantage for asymptotically infinite times. However, for only a small number of steps a shorter memory length can achieve a higher ground state population. Initially the shortest memory length of $l = 1$ performs best. However, this changes after only 16 steps, when the second shortest memory length $l = 2$ takes the lead, and so on. The number of steps it takes for the non-Markovian regimes to perform better than the memory-less counterpart follows a pattern akin to the coherent control scenario (see Fig. 10).

Similarly to the coherent control scenario, we also want to see how well incoherent regimes with memory carriers make use of machine resources. Fig. 16 plots the same cooling curves as in Fig. 15, but as a function of the total machines interacted with throughout the entire protocol.

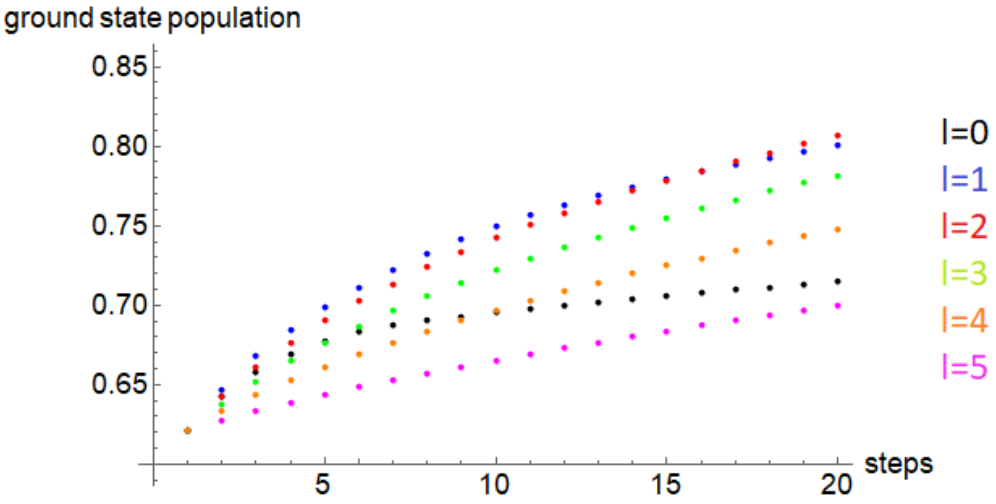

Figure 15: *Short Term cooling behaviour 2*
In the incoherent control regime with an adaptive protocol: The ground state population is plotted as a function of steps for a unitary size of $k = 6$, $\epsilon_{max} = 3$ and different memory lengths. The short term cooling behaviour can have crossovers (e.g. at step 16 memory length $l = 2$ surpasses $l = 1$). We can also see that the Markovian cooling curve very quickly starts to flatten (after about 6 steps) and gets overtaken by protocols with memory.

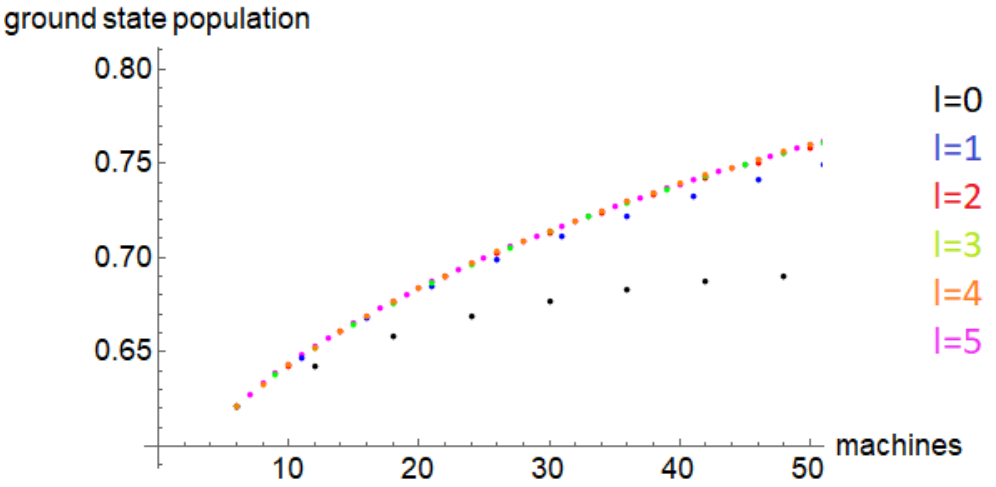

Figure 16: *Limited machine-resources comparison 4*
In the incoherent control regime with an adaptive protocol: The ground state population is plotted as a function of the total machines interacted with for a unitary size of $k = 6$, $\epsilon_{max} = 3$ and different memory lengths. Similarly to the coherent control regime, we see that regimes with memory can better utilise machine resources.

The Markovian dots (in black) are quickly outperformed by the non-Markovian ones. However, there is negligible difference between the trajectories of memory lengths $l > 1$, regarding using up to 50 machines, which means that all of these memory regimes have similar cooling improvements machine-wise. In other words, the energy that every machine individually extracts follows the same curve and it does not matter how much memory or machines are used at each step. The graph very much resembles the coherent control scenario: The non-Markovian regimes compress population into their memory carries until they become as cold as possible, which eventually results in diminishing cooling performances. The smallest memory length $l = 1$ splits first, followed by the second smallest and so on.

5 Conclusion

The aim of this thesis has been to understand how one can produce valuable pure states out of readily-available resources, and if memory can offer any advantage. We thus began by imposing some restrictions on the allowed transformations: in particular, we considered cooling arbitrary-dimensional quantum systems via unitary interactions with a thermal environment, from which the system is initially uncorrelated. The further choice of allowing either general or only energy-conserving unitaries led us to distinguish between the coherent and incoherent control paradigms, respectively. The latter coincides with the already established framework of thermal operations [12, 13], permitting the contextualisation of our results into a resource theory. Regarding the temporal control of the operation cycles, we proposed a framework that allows for comparison between various memory structures by means of a generalised collision model, in which a finite portion of the environment is controlled to act as memory carriers. Stipulating the size of the unitary interactions and the number of memory carriers provides a particular memory structure of the protocol, amongst which various performance comparisons can be fairly made. Moreover, for any given memory structure, we characterised our overall protocols according to two strategies: adaptive or non-adaptive.

In this setting, we have shown that memory greatly enhances the cooling performance for both control scenarios (in/coherent) compared to memoryless protocols [3] . We first considered the asymptotic cooling behaviour. The introduction of a finite memory length exponentially increases the asymptotic ground state population of the system via the availability of a larger low-energy subspace to compress populations into. This exponential improvement makes it possible for a protocol with memory to ultimately reach far colder states than a memoryless protocol, even when the latter has access to more fresh machines at each step (e.g., see Fig. 4). Ultimately, although the asymptotic ground state population increases as the protocol parameters (access to fresh machines and memory length) increase, the resulting hierarchy is not determined by either of the two protocol parameters *alone*, but rather the combined dynamics.

We then went on to analyse the short-term behaviour of various cooling protocols. Here we found that although any Markovian protocol eventually gets outperformed by its non-Markovian counterpart with the same unitary size in order of increasing memory length, in line with intuition, having more memory leads to a disadvantage for the initial rate of cooling (e.g., see Figs. 8, 9). A large memory length is unfavourable for short-term cooling, because the system interacts with less fresh machines at each step and as such less energy can be dissipated at each interaction. Nonetheless, a protocol with memory can fairly quickly outperform its memoryless counterpart, even if the latter has a larger unitary size (e.g., see Fig. 13).

Following this, we considered another means of comparison: We compared the performance of protocols with respect to how many machines the system has been in contact with throughout the whole process. We have shown that protocols with memory can better utilise machine-resources, i.e., achieve higher ground state populations given a smaller number of machines, compared to their Markovian counterparts (e.g., see Figs.12, 16). This advantage is also present when comparing protocols with memory to Markovian ones with a larger interaction size (e.g., see Fig. 14).

The presented results hold for arbitrary, finite dimensional systems as well as for both control scenarios (in-/coherent) and both interaction strategies (adaptive/non-adaptive). They can be viewed as a benchmark on how well cooling could ultimately be performed given our restrictions and resources, given perfect control over not only the unitary interactions—implying perfect timing abilities [56]—but also the implementation of the generalised collision model itself, which neglects any uncontrolled coupling between the target system and the broader environment beyond the memory carriers, and between the memory carriers and their broader environment. We now discuss some of these practical limitations and pose a number of open questions.

6 Outlook

In this thesis, we have considered repeated machine-system interactions to implement memory into our generalised collision model. However, this is only one possible way to endow collision models with memory. One could consider, for instance, machine-machine interactions, as already proposed in Refs. [17, 51, 57–59]. It would be interesting to understand how such different types of memory affects the cooling performance. For instance one could envisage a scenario in which the already used machines are again used in order to pre-cool the fresh machines that will in turn interact with the system, to extract further value from them. There, we would also have the choice to only use them only once, or to collect them and use them all together to additionally prepare the fresh machines, which could be advantageous.

Another possible cooling advantage could reasonably arise from weak initial correlations. It is clear that initially pure and correlated (i.e., entangled) states can lead to perfect cooling, which is why we restricted our system and environment to be initially uncorrelated. Nonetheless, one could ease this restriction to allow for weaker forms of initial correlations and analyse their affect on the cooling limits. Some progress to this end has already been made regarding the impact of initial correlations on collision models [18, 60], but not specifically in the context of cooling. We expect initial correlations to bring advantages if controlled appropriately, because it constitutes a further memory-type resource; as we have shown, similar temporal correlations between the system and the memory carriers are very beneficial to the cooling limit.

Regarding correlations, another point to note is that the system and memory carriers build up correlations throughout the protocol, which can be finally used to further cool the system . It is both this, and the fact that the memory carriers provide a larger subspace to compress population into over the course of the protocol, which leads to the cooling advantages seen. An interesting avenue for pursuit would be to study the effects of partial rethermalisation of the memory carriers throughout the protocol: in the coherent control paradigm, destroying correlations in this way could only ever be detrimental, but in the incoherent control scenario, it may prove beneficial (at least for finite times) as the rethermalisation could help to "free up" more energy-conserving subspaces that are relevant to cooling.

Further to this, we have examined various behaviours of different cooling protocols throughout this thesis. Additionally, we have seen that the asymptotic ground state population is only dependent on a few parameters, whereas the short-term behaviour is susceptible to the intricate details of the arrangement. This can result in different protocols reaching the same limit but in distinct ways, which poses the question of optimisation. What is the best (fastest, most resource economical, etc.) way to reach a certain degree of coldness in a given time with the available, limited resources, control and other parameters (e.g., machine dimension, maximum energy gap)? This is expected to be a difficult question, as the short-term behaviour of different cooling regimes can exhibit crossovers and re-crossovers (e.g., see Figs. 13, 15), and the parameter space is large. Another example of the delicate short term behaviour is the splitting of curves regarding the memory-aided machine-wise cooling comparisons (see Figs. 11, 14. When we studied the coherent control scenario, we saw that for some choices of parameters, the machine-wise cooling curves followed (approximately) the same path for a significant number of total machines used before splitting off from each other; however, comparing this to a scenario with the same memory structure but decreased machine dimension, we saw these cooling curves split much earlier (see Fig. 12). Determining the point of departure in terms of a few relevant parameters would provide an immense step towards understanding optimal cooling strategies. It is partly this intricate short term behaviour that can make the realisation of such cooling protocols susceptible to errors and imperfections. We have not taken into account the practical restrictions, regarding for example the ability to exactly implement unitaries. In order to perfectly implement every operation, a perfect time-keeping device would be needed [61], which is also impossible without infinite resources [56]. It could be interesting to pursue further research into the energy cost required to implement such unitaries with respect to the resources required to run the protocol and compare them to the cooling success. Additionally, we considered an adaptive strategy in order to achieve the best possible short-term performance, which can be difficult to manage and we often implement multi-partite unitaries, which also can be difficult to realise in an experimental setup. For asymptotically infinite times this would not make a difference, since we can always: i) use the non-adaptive strategy ii) split the optimal reordering-unitaries into a sequence of (smaller) swap operations and still could reach the same final state. However for finite times (and operations), as we have shown, the cooling performance is highly dependent on the petition of the implemented unitaries. Additionally, we have, just as in HBAC, neglected the coupling between

the system and the rest of the environment (i.e., the fresh non memory-carrying machines); in realistic scenarios, such considerations should be taken into account. Overall, the framework proposed in this thesis allows for the systematic treatment of memory effects in various cooling strategies, and it is flexible enough to include many practical considerations such as those suggested above to lead to more practical limitations of quantum cooling.

References

[1] F. Clivaz, R. Silva, G. Haack, J. B. Brask, N. Brunner, and M. Huber, "Unifying paradigms of quantum refrigeration: Fundamental limits of cooling and associated work costs," *Phys. Rev. E*, vol. 100, p. 042130, 2019. https://doi.org/10.1103/PhysRevE.100.042130.

[2] F. Clivaz, R. Silva, G. Haack, J. B. Brask, N. Brunner, and M. Huber, "Unifying Paradigms of Quantum Refrigeration: A Universal and Attainable Bound on Cooling," *Phys. Rev. Lett.*, vol. 123, p. 170605, 2019. https://doi.org/10.1103/PhysRevLett.123.170605.

[3] P. Taranto, F. Bakhshinezhad, P. Schüttelkopf, and M. Huber, "Exponential improvement for quantum cooling through finite memory effects," 2020. https://arxiv.org/abs/2004.00323.

[4] M. A. Nielsen and I. L. Chuang, *Quantum Computation and Quantum Information: 10th Anniversary Edition*. USA: Cambridge University Press, 10th ed., 2011. https://doi.org/10.1017/CBO9780511976667.

[5] C. Degen, F. Reinhard, and P. Cappellaro, "Quantum sensing," *Rev. Mod. Phys.*, vol. 89, no. 3, p. 035002, 2017. http://doi.org/10.1103/RevModPhys.89.035002.

[6] V. Giovannetti, S. Lloyd, and L. Maccone, "Advances in quantum metrology," *Nat. Photonics*, vol. 5, no. 4, p. 222–229, 2011. http://doi.org/10.1038/nphoton.2011.35.

[7] J. Goold, M. Huber, A. Riera, L. d. Rio, and P. Skrzypczyk, "The role of quantum information in thermodynamics—a topical review," *J. Phys. A*, vol. 49, no. 14, p. 143001, 2016. https://doi.org/10.1088/1751-8113/49/14/143001.

[8] A. Riera, C. Gogolin, and J. Eisert, "Thermalization in Nature and on a Quantum Computer," *Phys. Rev. Lett.*, vol. 108, p. 080402, 2012. https://doi.org/10.1103/PhysRevLett.108.080402.

[9] N. Freitas, R. Gallego, L. Masanes, and J. P. Paz, "Cooling to absolute zero: The unattainability principle," in *Thermodynamics in the Quantum Regime: Fundamental Aspects and New Directions* (F. Binder, L. A. Correa, C. Gogolin, J. Anders, and G. Adesso, eds.), vol. 195, pp. 597–622, Springer International Publishing, Cham, 2018. https://doi.org/10.1007/978-3-319-99046-0_25.

[10] R. Silva, G. Manzano, P. Skrzypczyk, and N. Brunner, "Performance of autonomous quantum thermal machines: Hilbert space dimension as a thermodynamical resource," *Phys. Rev. E*, vol. 94, p. 032120, 2016. https://doi.org/10.1103/PhysRevE.94.032120.

[11] L. Masanes and J. Oppenheim, "A general derivation and quantification of the third law of thermodynamics," *Nat. Commun.*, vol. 8, no. 1, 2017. http://doi.org/10.1038/ncomms14538.

[12] F. G. S. L. Brandão, M. Horodecki, J. Oppenheim, J. M. Renes, and R. W. Spekkens, "Resource Theory of Quantum States Out of Thermal Equilibrium," *Phys. Rev. Lett.*, vol. 111, p. 250404, 2013. https://doi.org/10.1103/PhysRevLett.111.250404.

[13] D. Janzing, P. Wocjan, R. Zeier, R. Geiß, and T. Beth, "Thermodynamic Cost of Reliability and Low Temperatures: Tightening Landauer's Principle and the Second Law," *Int. J. Theor. Phys.*, vol. 39, pp. 2717–2753, 2000. https://doi.org/10.1023/A:1026422630734.

[14] B. Bylicka, M. Tukiainen, J. Piilo, C. Dariusz, and S. Maniscalco, "Thermodynamic meaning and power of non-Markovianity," *Sci. Rep.*, vol. 6, 2015. https://doi.org/10.1038/srep27989.

[15] F. H. Kamin, F. T. Tabesh, S. Salimi, F. Kheirandish, and A. C. Santos, "Non-Markovian effects on charging and self-discharging processes of quantum batteries," 2019. https://arxiv.org/abs/1910.07751.

[16] M. Ziman, P. Štelmachovič, and V. Bužek, "Description of Quantum Dynamics of Open Systems Based on Collision-Like Models," *Open Syst. Inf. Dyn.*, vol. 12, no. 01, pp. 81–91, 2005. https://doi.org/10.1007/s11080-005-0488-0.

[17] F. Ciccarello, G. M. Palma, and V. Giovannetti, "Collision-model-based approach to non-Markovian quantum dynamics," *Phys. Rev. A*, vol. 87, p. 040103, 2013. https://doi.org/10.1103/PhysRevA.87.040103.

[18] N. K. Bernardes, A. R. R. Carvalho, C. H. Monken, and M. F. Santos, "Environmental correlations and Markovian to non-Markovian transitions in collisional models," *Phys. Rev. A*, vol. 90, p. 032111, 2014. https://doi.org/10.1103/PhysRevA.90.032111.

[19] B. Vacchini, "General structure of quantum collisional models," *Int. J. Quantum Inf.*, vol. 12, no. 02, p. 1461011, 2014. https://doi.org/10.1142/S0219749914610115.

[20] Z.-X. Man, Y.-J. Xia, and R. Lo Franco, "Validity of the Landauer principle and quantum memory effects via collisional models," *Phys. Rev. A*, vol. 99, no. 4, 2019. http://doi.org/10.1103/PhysRevA.99.042106.

[21] S. Popescu, A. Short, and A. Winter, "Entanglement and the foundations of statistical mechanics," *Nat. Phys.*, vol. 2, pp. 754–758, 2006. https://doi.org/10.1038/nphys444.

[22] F. Ticozzi and L. Viola, "Quantum resources for purification and cooling: fundamental limits and opportunities," *Sci. Rep.*, vol. 4, no. 1, 2014. http://doi.org/10.1038/srep05192.

[23] J. Fischer, D. Basilewitsch, C. P. Koch, and D. Sugny, "Time-optimal control of the purification of a qubit in contact with a structured environment," *Phys. Rev. A*, vol. 99, no. 3, 2019. http://doi.org/10.1103/PhysRevA.99.033410.

[24] J. Gemmer, M. Michel, and G. Mahler, "18 Equilibrium Properties of Model Systems," in *Quantum Thermodynamics: Emergence of Thermodynamic Behavior Within Composite Quantum Systems*, pp. 185–219, Springer Berlin Heidelberg, 2004. https://doi.org/10.1007/978-3-540-44513-5_18.

[25] N. Linden, S. Popescu, A. J. Short, and A. Winter, "Quantum mechanical evolution towards thermal equilibrium," *Phys. Rev. E*, vol. 79, p. 061103, 2009. https://doi.org/10.1103/PhysRevE.79.061103.

[26] W. Pusz and S. Woronowicz, "Passive states and KMS states for general quantum systems," *Commun. Math. Phys.*, vol. 58, pp. 273–290, 1978. http://doi.org/10.1007/BF01614224.

[27] N. H. Y. Ng and M. P. Woods, "Resource theory of quantum thermodynamics: Thermal operations and second laws," in *Thermodynamics in the Quantum Regime: Fundamental Aspects and New Directions* (F. Binder, L. A. Correa, C. Gogolin, J. Anders, and G. Adesso, eds.), vol. 195, pp. 625–650, Springer International Publishing, Cham, 2018. https://doi.org/10.1007/978-3-319-99046-0_26.

[28] P. O. Boykin, T. Mor, V. Roychowdhury, F. Vatan, and R. Vrijen, "Algorithmic cooling and scalable NMR quantum computers," *Proc. Natl. Acad. Sci. U.S.A.*, vol. 99, no. 6, pp. 3388–3393, 2002. https://doi.org/10.1073/pnas.241641898.

[29] L. J. Schulman and U. Vazirani, "Scalable NMR Quantum Computation," 1998. https://arxiv.org/abs/quant-ph/9804060.

[30] S. Raeisi and M. Mosca, "Asymptotic Bound for Heat-Bath Algorithmic Cooling," *Phys. Rev. Lett.*, vol. 114, p. 100404, 2015. https://doi.org/10.1103/PhysRevLett.114.100404.

[31] L. J. Schulman, T. Mor, and Y. Weinstein, "Physical Limits of Heat-Bath Algorithmic Cooling," *Phys. Rev. Lett.*, vol. 94, p. 120501, 2005. https://doi.org/10.1103/PhysRevLett.94.120501.

[32] M. T. Mitchison and P. P. Potts, "Physical implementations of quantum absorption refrigerators," in *Thermodynamics in the Quantum Regime: Fundamental Aspects and New Directions* (F. Binder, L. A. Correa, C. Gogolin, J. Anders, and G. Adesso, eds.), pp. 149–174, Springer International Publishing, Cham, 2018. http://doi.org/10.1007/978-3-319-99046-0_6.

[33] M. T. Mitchison, M. P. Woods, J. Prior, and M. Huber, "Coherence-assisted single-shot cooling by quantum absorption refrigerators," *New J. Phys.*, vol. 17, no. 11, p. 115013, 2015. https://doi.org/10.1088/1367-2630/17/11/115013.

[34] H.-P. Breuer and F. Petruccione, *The Theory of Open Quantum Systems*. Oxford University Press, 2002. https://doi.org/10.1093/acprof:oso/9780199213900.001.0001.

[35] J. Rau, "Relaxation phenomena in spin and harmonic oscillator systems," *Phys. Rev.*, vol. 129, pp. 1880–1888, 1963. https://doi.org/10.1103/PhysRev.129.1880.

[36] G. Lindblad, "On the generators of quantum dynamical semigroups," *Commun. Math. Phys.*, vol. 48, no. 2, pp. 119–130, 1976. https://doi.org/10.1007/BF01608499.

[37] V. Gorini, A. Kossakowski, and E. C. G. Sudarshan, "Completely positive dynamical semigroups of N-level systems," *J. Math. Phys.*, vol. 17, no. 5, pp. 821–825, 1976. https://doi.org/10.1063/1.522979.

[38] V. Franke, "On the general form of the dynamical transformation of density matrices," *Theor. Math. Phys.*, vol. 27, pp. 406–413, 1976. https://doi.org/10.1007/BF01051230.

[39] D. Grimmer, D. Layden, R. B. Mann, and E. Martín-Martínez, "Open dynamics under rapid repeated interaction," *Phys. Rev. A*, vol. 94, p. 032126, 2016. https://doi.org/10.1103/PhysRevA.94.032126.

[40] S. Attal and Y. Pautrat, "From repeated to continuous quantum interactions," 2003. https://arxiv.org/abs/math-ph/0311002.

[41] H.-P. Breuer, E.-M. Laine, J. Piilo, and B. Vacchini, "Colloquium: Non-Markovian dynamics in open quantum systems," *Rev. Mod. Phys.*, vol. 88, p. 021002, 2016. https://doi.org/10.1103/RevModPhys.88.021002.

[42] I. de Vega and D. Alonso, "Dynamics of non-Markovian open quantum systems," *Rev. Mod. Phys.*, vol. 89, p. 015001, 2017. https://doi.org/10.1103/RevModPhys.89.015001.

[43] L. Li, M. J. Hall, and H. M. Wiseman, "Concepts of quantum non-Markovianity: A hierarchy," *Phys. Rep.*, vol. 759, pp. 1 – 51, 2018. https://doi.org/10.1016/j.physrep.2018.07.001.

[44] Á. Rivas, S. F. Huelga, and M. B. Plenio, "Quantum non-Markovianity: characterization, quantification and detection," *Rep. Prog. Phys.*, vol. 77, no. 9, p. 094001, 2014. https://doi.org/10.1088/0034-4885/77/9/094001.

[45] F. A. Pollock, C. Rodríguez-Rosario, T. Frauenheim, M. Paternostro, and K. Modi, "Non-Markovian quantum processes: Complete framework and efficient characterization," *Phys. Rev. A*, vol. 97, p. 012127, 2018. https://doi.org/10.1103/PhysRevA.97.012127.

[46] U. Shrikant, R. Srikanth, and S. Banerjee, "Ping-pong quantum key distribution with trusted noise: non-Markovian advantage," 2020. https://arxiv.org/abs/2004.05689.

[47] V. Mukherjee, V. Giovannetti, R. Fazio, S. F. Huelga, T. Calarco, and S. Montangero, "Efficiency of quantum controlled non-Markovian thermalization," *New J. Phys.*, vol. 17, no. 6, p. 063031, 2015. http://dx.doi.org/10.1088/1367-2630/17/6/063031.

[48] S. Lorenzo, F. Ciccarello, G. M. Palma, and B. Vacchini, "Quantum Non-Markovian Piecewise Dynamics from Collision Models," *Open Syst. Inf. Dyn.*, vol. 24, no. 04, p. 1740011, 2017. http://doi.org/10.1142/S123016121740011X.

[49] L. Mazzola, C. A. Rodríguez-Rosario, K. Modi, and M. Paternostro, "Dynamical role of system-environment correlations in non-Markovian dynamics," *Phys. Rev. A*, vol. 86, no. 1, 2012. http://doi.org/10.1103/PhysRevA.86.010102.

[50] P. Taranto, "Memory effects in quantum processes," *Int. J. Quantum Inf.*, vol. 18, no. 02, p. 1941002, 2020. http://doi.org/10.1142/S0219749919410028.

[51] S. Campbell, F. Ciccarello, G. M. Palma, and B. Vacchini, "System-environment correlations and Markovian embedding of quantum non-Markovian dynamics," *Phys. Rev. A*, vol. 98, no. 1, 2018. http://doi.org/10.1103/PhysRevA.98.012142.

[52] S. Raeisi, M. Kieferová, and M. Mosca, "Novel Technique for Robust Optimal Algorithmic Cooling," *Phys. Rev. Lett.*, vol. 122, p. 220501, 2019. https://doi.org/10.1103/PhysRevLett.122.220501.

[53] T. Feldmann and J. P. Palao, "Performance of Quantum Thermodynamic Cycles," in *Thermodynamics in the Quantum Regime: Fundamental Aspects and New Directions* (F. Binder, L. A. Correa, C. Gogolin, J. Anders, and G. Adesso, eds.), pp. 67–85, Cham: Springer International Publishing, 2018. https://doi.org/10.1007/978-3-319-99046-0_3.

[54] F. Campaioli, F. A. Pollock, and S. Vinjanampathy, "Quantum Batteries," in *Thermodynamics in the Quantum Regime: Fundamental Aspects and New Directions* (F. Binder, L. A. Correa, C. Gogolin, J. Anders, and G. Adesso, eds.), pp. 207–225, Cham: Springer International Publishing, 2018. https://doi.org/10.1007/978-3-319-99046-0_8.

[55] R. Alicki and K. Lendi, *Quantum Dynamical Semigroups and Applications*. Lecture Notes in Physics, Springer Berlin Heidelberg, 2007. https://doi.org/10.1007/3-540-70861-8.

[56] P. Erker, M. T. Mitchison, R. Silva, M. P. Woods, N. Brunner, and M. Huber, "Autonomous Quantum Clocks: Does Thermodynamics Limit Our Ability to Measure Time?," *Phys. Rev. X*, vol. 7, p. 031022, 2017. https://doi.org/10.1103/PhysRevX.7.031022.

[57] R. McCloskey and M. Paternostro, "Non-Markovianity and system-environment correlations in a microscopic collision model," *Phys. Rev. A*, vol. 89, p. 052120, 2014. https://doi.org/10.1103/PhysRevA.89.052120.

[58] S. Lorenzo, F. Ciccarello, and G. M. Palma, "Class of exact memory-kernel master equations," *Phys. Rev. A*, vol. 93, p. 052111, 2016. https://doi.org/10.1103/PhysRevA.93.052111.

[59] B. Çakmak, M. Pezzutto, M. Paternostro, and O. E. Müstecaplıoğlu, "Non-Markovianity, coherence, and system-environment correlations in a long-range collision model," *Phys. Rev. A*, vol. 96, p. 022109, 2017. https://doi.org/10.1103/PhysRevA.96.022109.

[60] T. Rybár, S. N. Filippov, M. Ziman, and V. Bužek, "Simulation of indivisible qubit channels in collision models," *J. Phys. B*, vol. 45, no. 15, p. 154006, 2012. https://doi.org/10.1088/0953-4075/45/15/154006.

[61] A. S. L. Malabarba, A. J. Short, and P. Kammerlander, "Clock-driven quantum thermal engines," *New J. Phys.*, vol. 17, no. 4, p. 045027, 2015. https://doi.org/10.1088/1367-2630/17/4/045027.