Quantum Random Walks

ETH
Eidgenössische Technische Hochschule Zürich
Swiss Federal Institute of Technology Zurich

Quantum Random Walks

Master Thesis

Lorenzo Laneve

April 1, 2022

Advisors: Dr. Ivano Tavernelli (IBM Research)
Dr. Francesco Tacchino (IBM Research)
Supervisor: Prof. Dr. Kenny Paterson

Applied Cryptography Group
Institute of Information Security
Department of Computer Science, ETH Zürich

Abstract

Random walks are stochastic models which are extensively used in theoretical computer science: one of the best known classical algorithms for the satisfiability problem (SAT) relies on the behaviour of a random walk [21].

Research of the past decades tried to extend the idea to the quantum setting, in order to see whether it is possible to achieve further speedup using quantum effects. Notable examples are the works by Kempe [12, 13], where the *unitary walk model* is introduced: such walk consists in repeatedly applying a unitary to a pure quantum state. Despite interesting results were found [1, 2, 25], the unitary walk model cannot be seen as a generalization of classical random walks: unitary evolution, unlike the one exhibited by classical random walks, is always reversible. Moreover, randomness and decoherence, which are phenomena conjectured to provide speedup in some quantum processes [14, 15], cannot be represented in this model.

In this work we show that, using the density matrix formalism, it is possible to construct a framework that naturally extends the classical theory, and unifies the classical and unitary models, while being able to express any hybrid quantum-classical process.

Contents

Introduction

Random walks, also known as Markov chains in the classical literature, are models with a graph traversed by a walker jumping from node to node, choosing at each step which node will be visited next in a random fashion. Random walks are extensively used in theoretical computer science: for example, one of the best known classical algorithms for k-SAT relies on the behaviour of a random walk [21].

Research of the past decades tried to extend the idea of random walk to the quantum setting: notable examples are the seminal works by Kempe [12, 13], where a concept of *unitary walk* is introduced. In such model, we start from a (pure) quantum state $|\psi\rangle$, and we repeatedly apply a unitary U. Therefore, instead of random choices, here the walker visits the nodes of the graph in a *superposition*.

In the same works, Kempe proposed a formalization of *hitting times* and *mixing times*, concepts already present in the classical theory of Markov chains, which are extremely important for algorithmic purposes.

Quantum superposition often lead to unexpected behaviours: for example, a unitary walk on the line expands in space quadratically faster than its classical counterpart [12]. Moreover, a general framework proposed by Szegedy [25] shows how to construct a unitary walk with quadratically

faster mixing times from a classical Markov chain.

Unfortunately, the unitary model cannot be formally seen as a generalization of Markov chains, as not every classical Markov chain can be reproduced by a unitary walk. Moreover, phenomena such as randomness and decoherence cannot be expressed as unitary transformations. In this thesis we show a way to use the density matrix formalism to give a definition of *quantum Markov chain* (the discrete counterpart of the quantum stochastic walk [26]). This models aims to give a formalization for quantum random walks that (1) more closely resembles the classical theory of Markov chains, and (2) is able to take into account any hybrid quantum-classical process, in order to explore its potentiality.

Chapters 1 and 2 introduce to the basics of quantum theory and quantum information, explaining all the physical and mathematical concepts needed to follow the rest of the work.

Chapter 3 defines Markov processes and chains: starting from the classical definitions, we construct a quantum counterpart highlighting the analogies with the original formalization.

Chapter 4 reformulates the definitions of hitting times originally proposed by Kempe [13] in our framework and gives a comparison with the classical definition of hitting time, showing that one of these versions is the natural generalization of its classical counterpart, and give a general formula to compute these times. The rest of the chapter focuses on giving examples and applications of this formalism, as well as exploring extensions and limitations for the quantum hitting time.

Chapter 5 applies the definitions of hitting times to analyze the famous Grover's search algorithm [7] in a novel way. In particular we show that, using a particular scheme of measurement carried out during the evolution of the walk, we manage to preserve the quantum speed-up of the algorithm,

while the analysis closely follows typical approaches found in the classical theory.

Chapter 6 explores necessary and sufficient conditions for *ergodicity*, a property of random walks that implies approach to a stationary state in the long run. We collect and reprove results found in [27, 8], also giving connections with hitting times, in analogy to the classical theory.

Chapter 1

Quantum Information

This chapter, along with the next one, aims to give basic notions of quantum theory and quantum information, which are crucial in order to understand the rest of the work. The topics covered by the first two chapters are treated more thoroughly in [22].

1.1 Hilbert spaces and the braket notation

We start by giving some notions of linear algebra, which will be important for the rest of this work.

Definition 1.1 *A* inner product space $(\mathcal{X}, \cdot)$ *is a vector space equipped with a inner product, i.e. an operation taking two vectors of* $\mathcal{X}$ *and returning an element of the underlying field.*

Definition 1.2 *A inner product space is said to be a* Hilbert space *when every Cauchy sequence converges to a limit contained in the space itself.*

The definition of Hilbert space, although not straightforward to understand, is extremely important for mathematical rigorousness: it ensures that any infinite/integral sum of vectors in the space still gives us a valid element in the same space.

In quantum theory, a physical system is represented by a Hilbert space over complex field, and a (column) vector of this field represents a possible *state* of the system. We use the *braket* notation: column vectors are written as *kets* $|\phi\rangle$, while a row vector is represented by a *bra* $\langle\phi|$. In general, we can see the bra as a shorthand for:

$$\langle\psi| = (|\psi\rangle)^\dagger$$

where $\cdot^\dagger$ denotes the *conjugate transpose*. Like in standard linear algebra, one can multiply a bra and a ket, obtaining the inner product:

$$\langle\phi||\psi\rangle \equiv \langle\phi|\psi\rangle \in \mathbb{C}$$

On the other hand, outer products can be written in the form $|\phi\rangle\langle\psi|$, and they will represent matrices.

1.2 Qubits

Let us consider a simple physical system: a bit. Bits can be only in two possible states: '0' or '1'. What is the natural transposition of a bit to the quantum setting? Consider a Hilbert space $\mathcal{H}$ of dimension two! Then, we can fix a basis[1] $\{|0\rangle, |1\rangle\}$ for this space representing the two states of the bit. This basis is also called *computational basis* of the qubit, and we will use this basis as the standard basis when we write objects in vector form:

$$|0\rangle = \begin{bmatrix} 1 \\ 0 \end{bmatrix}, |1\rangle = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$$

Unlike classical bits, this bit (which we will call quantum bit, or *qubit*) can be in any state of the form

$$|\psi\rangle = \alpha|0\rangle + \beta|1\rangle$$

[1] In quantum theory, we assume every basis to be orthonormal unless specified otherwise.

where $\alpha, \beta \in \mathbb{C}$ with $|\alpha|^2 + |\beta|^2 = 1$. This last constraint is called *normalization*: every valid physical state must have unitary norm, as the entries of the vector represent some probability distribution.

1.3 Measurements

One can argue that, since a qubit can be in infinitely many states, it can carry an infinite amount of information. Unfortunately this is not true: in order to 'read' this qubit we need to *measure* it. Roughly speaking, we fix a basis of the Hilbert space, and this measurement will result in an outcome which tells us which state of the basis the system is in. In the case of the qubit, if we measure with respect to the computational basis, we can obtain a 0 or a 1, i.e. one bit of information regarding the current state of the qubit. But what happens if the qubit is in a state of the form $\alpha|0\rangle + \beta|1\rangle$, with $\alpha, \beta \neq 0$?

In this case, the qubit is said to be in a *superposition* of the states of the computational basis, and when we measure we obtain 0 with probability $|\alpha|^2$ and 1 with probability $|\beta|^2$ (and here is why we introduced the normalization constraint in the last section). Another interesting behaviour is that, once we obtain an outcome from a measurement, the state of the system will *collapse* to that state: for example, if we measure a qubit in the computational basis, and we read a '0', then the state of the qubit from that moment will be exactly $|0\rangle$.

More rigorously, a (projective) measurement is a set of orthogonal projectors $\{\Pi_x\}$ such that, if applied on a system in state $|\psi\rangle$, the process returns x with probability:

$$\Pr_\psi [x] = |\Pi_x|\psi\rangle|^2 = \langle\psi|\Pi_x|\psi\rangle$$

and, if x is measured, the system will collapse to the *post-measurement state*:

$$|\psi_x\rangle = \frac{\Pi_x|x\rangle}{|\Pi_x|x\rangle|}$$

This is also to highlight that, in order to carry out a measurement, we do not need to reveal the whole information about the state. For example, in a three-dimensional Hilbert space, one can measure with respect to the computational basis $\{|0\rangle, |1\rangle, |2\rangle\}$, but it is also possible to measure an arbitrary predicate about the state, (e.g. is the value of the state $\neq 1$?).

1.4 Unitary transformations

Transformations of the state of a quantum system can be mathematically expressed as *unitary matrices*, i.e. matrices U satisfying $UU^\dagger = U^\dagger U = \mathbb{1}$. The unitarity of the matrix is necessary because it ensures that the norm of the states are preserved and, in particular, quantum states remain mathematically valid throughout the evolution. A simple example is the Pauli X gate for qubits:

$$X = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$$

One can see that this acts like a NOT operation in the computational basis: $X|0\rangle = |1\rangle, X|1\rangle = |0\rangle$. On the other hand, consider the following *Hadamard* states:

$$|+\rangle = \frac{|0\rangle + |1\rangle}{\sqrt{2}}$$
$$|-\rangle = \frac{|0\rangle - |1\rangle}{\sqrt{2}}$$

One can see, by linearity, that $X|+\rangle = |+\rangle, X|-\rangle = -|-\rangle$, i.e. the Hadamard basis forms an *eigenbasis* for X. In general, the high-level behaviour of a unitary operator strongly depends on the basis we are considering and at the same time, by linearity, a gate is fully determined by how it acts on the elements of a given basis.

1.5 Composing systems

Considering two Hilbert spaces $\mathcal{H}_A \ni |x\rangle_A, \mathcal{H}_B \ni |y\rangle_B$, we define a *tensor product*

$$|x\rangle_A \otimes |y\rangle_B$$

which denotes the states of the system containing the two subsystems A and B, and we may explicit the subscripts on the kets indicating which state belongs to which subsystem, unless sufficiently clear from the context. The tensor product has the following properties:

- It is **distributive** over addition:

$$(|x_1\rangle + |x_2\rangle) \otimes |y\rangle = |x_1\rangle \otimes |y\rangle + |x_2\rangle \otimes |y\rangle$$
$$|x\rangle \otimes (|y_1\rangle + |y_2\rangle) = |x\rangle \otimes |y_1\rangle + |x\rangle \otimes |y_2\rangle$$

- Scalars can be taken out of the product:

$$(a|x\rangle) \otimes |y\rangle = |x\rangle \otimes (a|y\rangle) = a(|x\rangle \otimes |y\rangle)$$

- Tensor product of operators is applied independently:

$$(U_1 \otimes U_2)(|x_1\rangle \otimes |x_2\rangle) = U_1|x_1\rangle \otimes U_2|x_2\rangle$$

- Inner product acts linearly on the tensor product:

$$(\langle x_1| \otimes \langle y_1|)(|x_2\rangle \otimes |y_2\rangle) = \langle x_1|x_2\rangle\langle y_1|y_2\rangle$$

From now on, we will also write $|x\rangle|y\rangle$ or even $|xy\rangle$ to denote tensor products. Let us do again an example with qubits. If we have two qubits A and B, where the first is in state $|0\rangle_A$ and the second is in the state $|1\rangle_B$, the total system is a two-qubit string in the state:

$$|0\rangle_A \otimes |1\rangle_B = |01\rangle$$

1.6 Modeling randomness

Suppose that we have a process which sets the system in a quantum state $|\phi_i\rangle$ with probability p_i, and $\sum_i p_i = 1$. Can we find a mathematical object that can help us to conveniently describe such state? Consider the measurement probabilities: if we have a projector Π_x, then by the law of total probability

$$\begin{aligned}
\Pr[x] &= \sum_i p_i \Pr_{\phi_i}[x] \\
&= \sum_i p_i \langle\phi_i|\Pi_x|\phi_i\rangle \\
&= \sum_i p_i \operatorname{Tr}(\Pi_x|\phi_i\rangle\langle\phi_i|) \\
&= \operatorname{Tr}\left(\Pi_x \sum_i p_i |\phi_i\rangle\langle\phi_i|\right) \\
&=: \operatorname{Tr}(\Pi_x \rho)
\end{aligned}$$

Such matrix ρ is called *density matrix* or *density operator*, and it represents a general (mixed) quantum state. One can immediately see that ρ must be Hermitian, positive semi-definite and with trace 1, and this also implies that its eigenvalues form a probability distribution, where an eigenstate of ρ appears with a probability equal to its associated eigenvalue. In other words, a density matrix represents nothing more than a probability distribution of quantum states. Also, by a natural extension, post-measurement states can be computed as:

$$\rho \mapsto \frac{\Pi_x \rho \Pi_x}{\operatorname{Tr}(\Pi_x \rho)}$$

Moreover, if we apply a unitary U to the state ρ will return (whatever it will be), the density matrix of the resulting state is given by $U\rho U^\dagger$. In the rest of the work, we denote by $\bar{\mathcal{H}}$ the space of density matrices in the Hilbert space $\mathcal{H}$, and we will use the term *state* also for a density operator.

Chapter 2

Linear Algebra of Quantum Channels

Suppose now we want to introduce randomness in the evolution. For example, what if we want to apply a certain unitary U_i with probability p_i? We can model such process with following map:

$$\rho \mapsto \sum_i p_i U_i \rho U_i^\dagger$$

Also (a priori) measurements can be schematized as such mappings (ignoring terms with zero probability):

$$\rho \mapsto \sum_x \mathrm{Tr}(\Pi_x \rho) \cdot \frac{\Pi_x \rho \Pi_x}{\mathrm{Tr}(\Pi_x \rho)} = \sum_x \Pi_x \rho \Pi_x$$

Unlike unitary transformations, which are always reversible (the inverse $U^\dagger$ always exists!), these operations may not be invertible in general. Let us generalize the notion of *quantum channel* as a map $\mathcal{E} : \bar{\mathcal{H}} \to \bar{\mathcal{H}}$[1] which is:

- **linear**, because by the law of total probability we have $\sum_i p_i \mathcal{E}(\rho_i) = \mathcal{E}(\sum_i p_i \rho_i)$, and linearity is needed in order to keep states consistent with the evolution;
- **trace-preserving**, because states with trace 1 must remain with trace 1 after the evolution;

[1] In this work we only consider endomorphic maps, i.e. maps $\mathcal{E} : \bar{\mathcal{H}} \to \bar{\mathcal{H}}$. However, general theory allows maps with different input and output systems.

- **completely positive**, because density matrices after the evolution must remain positive semi-definite, even when applied on submatrices (subsystems) of larger matrices.

To sum up, these three conditions are necessary (and sufficient) to keep the states valid throughout the evolution. An important result due to Kraus (Theorems 8.1-8.3 in [18]) states that any completely positive map can be decomposed into a so-called *Kraus decomposition*:

$$\mathcal{E}(\rho) = \sum_i M_i \rho M_i^\dagger$$

where $\{M_i\}_i$ are the *Kraus operators*. Furthermore, it is straightforward to see that $\mathcal{E}$ is trace-preserving if and only if $\sum_i M_i^\dagger M_i = \mathbb{1}$.

2.1 The Hilbert-Schmidt space

A crucial observation is that, since quantum channels are linear, they can be expressed as matrices. Therefore, if we think of density operators as vectors, quantum channels can be represented as matrices. We now formalize these ideas.

Definition 2.1 *Let $\mathcal{H}$ be a Hilbert space of finite dimension d. The Hilbert-Schmidt space is the space $\bar{\mathcal{H}}$ equipped with the inner product:*

$$\langle A, B \rangle := \mathrm{Tr}(A^\dagger B)$$

Notice that $\dim \bar{\mathcal{H}} = d^2$. More about the Hilbert-Schmidt inner product can be found on [18].

Theorem 2.2 *There always exists an orthonormal basis of Hermitian operators for a Hilbert-Schmidt space.*

Proof Consider the following basis:

$$\{B_k\}_k := \{|x\rangle\langle x|\}_x \cup \{|x\rangle\langle y| + |y\rangle\langle x|\}_{x \neq y} \cup \{i|x\rangle\langle y| - i|y\rangle\langle x|\}_{x \neq y}$$

for $x, y \in [d]$. One can prove this is a basis by noticing they are d^2 pairwise orthogonal elements. □

From now on, we will fix such orthonormal basis $\{B_k\}_k$, and the matrices of the Hilbert-Schmidt space can be seen as d^2-dimensional vectors. Also notice that $\rho \in \bar{\mathcal{H}}$ is Hermitian if and only if it is a real linear combination of the elements of this basis: if there is a complex element ρ_i we would get a non-Hermitian term $\rho_i B_i$, and this term cannot cancel out with other terms by linear independence.

Let us now look at the trace of $\rho = \sum_k \rho_k B_k$:

$$\mathrm{Tr}(\rho) = \sum_k \rho_k \,\mathrm{Tr}(B_k) = \sum_{k \in [d]} \rho_k \,\mathrm{Tr}(B_k)$$

where we conveniently choose the first d elements of the basis to be the diagonal elements $\{|k\rangle\langle k|\}_k$. Here we only have the constraint that the first d entries of the vector sum up to 1. Together with the fact that the diagonal entries of ρ are always non-negative, this shows that the first d elements represent a probability distribution over the states of the computational basis, i.e. the distribution we would observe by measuring ρ in the computational basis.

The $d(d-1)$ remaining entries all represent quantum coherences. In a classical setting where we have a classical probabilistic mixture among states, we would set all these entries to 0.

Now we can add quantum channels into the picture: notice that quantum channels are linear transformations in this space, therefore they can be represented as $d^2 \times d^2$ matrices with respect to the basis $\{B_k\}_k$. If $\mathcal{E}(\rho) = \sum_i M_i \rho M_i^\dagger$ is a Kraus decomposition for $\mathcal{E}$, then the (x, y)-entry of $P_\mathcal{E}$ is given by:

$$(P_\mathcal{E})_{x,y} = \langle B_x, \mathcal{E}(B_y)\rangle = \mathrm{Tr}(B_x^\dagger \mathcal{E}(B_y)) = \sum_i \mathrm{Tr}(B_x^\dagger M_i B_y M_i^\dagger)$$

Note that both $A = B_y^\dagger$ and $B = M_i B_x M_i^\dagger$ are Hermitian, and the trace is real. Thus, $P_\mathcal{E}$ is real. Thus, for the rest of the work we will implicitly treat quantum channels as matrices when needed.

2.2 Adjoint maps and unitality

Definition 2.3 *Let $\mathcal{E} : \bar{\mathcal{H}} \to \bar{\mathcal{H}}$ be a completely positive map. The adjoint map $\mathcal{E}^\dagger$ of $\mathcal{E}$ is uniquely defined as:*

$$\langle A, \mathcal{E}(B)\rangle = \langle \mathcal{E}^\dagger(A), B\rangle \iff \mathrm{Tr}(A\mathcal{E}(B)) = \mathrm{Tr}(\mathcal{E}^\dagger(A)B)$$

for any element $A, B \in \bar{\mathcal{H}}$.

One can see that the adjoint operation is translated to the transpose conjugate operation in the Hilbert-Schmidt space we defined earlier: $P_{\mathcal{E}^\dagger} = P_\mathcal{E}^\dagger$.

Definition 2.4 *A completely positive map $\mathcal{B} : \bar{\mathcal{H}} \to \bar{\mathcal{H}}$ is said to be* unital *if*

$$\mathcal{B}(\mathbb{1}) = \mathbb{1}$$

If $\mathcal{B}(\rho) = \sum_i M_i \rho M_i^\dagger$ is a Kraus decomposition of $\mathcal{B}$, then it must hold that

$$\sum_i M_i M_i^\dagger = \mathbb{1}$$

Theorem 2.5 *$\mathcal{E}$ is a completely positive trace-preserving map if and only if $\mathcal{E}^\dagger$ is a completely positive unital map.*

Proof Let $\mathcal{E}(\rho) = \sum_i M_i \rho M_i^\dagger$ be a Kraus decomposition for $\mathcal{E}$. Then

$$\begin{aligned}\mathrm{Tr}(\mathcal{E}(\rho)A) &= \sum_i \mathrm{Tr}(M_i \rho M_i^\dagger A)\\ &= \sum_i \mathrm{Tr}(\rho M_i^\dagger A M_i)\\ &= \mathrm{Tr}(\rho \sum_i M_i^\dagger A M_i)\end{aligned}$$

This implies $\mathcal{E}^\dagger(A) = \sum_i M_i^\dagger A M_i$ is a Kraus decomposition for $\mathcal{E}^\dagger$, which is then completely positive as well. The converse is obtained in a similar way. Now suppose $\mathcal{E}$ is trace-preserving: the Kraus operators then satisfy $\sum_i M_i^\dagger M_i = \mathbb{1}$ and we have.

$$\mathcal{E}^\dagger(\mathbb{1}) = \sum_i M_i^\dagger M_i = \mathbb{1}$$

giving the unitality of $\mathcal{E}^\dagger$. Conversely, suppose $\mathcal{E}^\dagger(\mathbb{1}) = \mathbb{1}$:

$$\begin{aligned}\mathrm{Tr}(\mathcal{E}(\rho)) &= \sum_i \mathrm{Tr}(M_i \rho M_i^\dagger) \\ &= \sum_i \mathrm{Tr}(M_i^\dagger M_i \rho) \\ &= \mathrm{Tr}(\sum_i M_i^\dagger M_i \rho) \\ &= \mathrm{Tr}(\mathcal{E}^\dagger(\mathbb{1})\rho) = \mathrm{Tr}(\rho)\end{aligned}$$

implying trace preservation. □

These definitions are treated more thoroughly in [27, 18].

2.3 Eigenvalues and eigenoperators

Since quantum channels can be represented by matrices, we can apply the spectral theory to such objects. We say that λ is an eigenvalue of the quantum channel $\mathcal{E}$ whenever

$$\mathcal{E}(X) = \lambda X$$

In this case, X is also said to be an eigenoperator associated with the eigenvalue λ. Notice that the eigenvalues of $\mathcal{E}$ are exactly the eigenvalues of the representation matrix $P_\mathcal{E}$, and this also implies that there are exactly d^2 eigenvalues (counted with multiplicity).

Observation 2.6 *For any linear map, each real eigenvalue is associated to at least one Hermitian eigenoperator.*

Proof Let λ be a real eigenvalue, with associated eigenoperator X (the geometric multiplicity is at least one for each eigenvalue).

$$\mathcal{E}(X) = \lambda X, \mathcal{E}(X^\dagger) = \lambda X^\dagger \Longrightarrow \mathcal{E}(X + X^\dagger) = \lambda(X + X^\dagger)$$

i.e. $X + X^\dagger$, which is Hermitian by construction, is also a λ-eigenoperator.□

Observation 2.7 *For Hermiticity-preserving maps (i.e. also quantum channels), eigenvalues come in complex conjugate pairs. Moreover, any Hermitian eigenoperator has to be associated to a real eigenvalue.*

Proof The first claim follows from the fact that the matrix $P_\mathcal{E}$ acting on the Hilbert-Schmidt space is real. If an Hermitian eigenoperator X is associated to a non-real eigenvalue λ, then $\mathcal{E}(X) = \lambda X$ would not be Hermitian, contradicting the assumption. □

Theorem 2.8 *$\mathcal{E}$ and $\mathcal{E}^\dagger$ have the same spectrum.*

Proof Notice that $P_\mathcal{E}$ and $P_{\mathcal{E}^\dagger} = P_\mathcal{E}^\dagger$ have the same spectrum. □

In analogy to matrices, we refer to the eigenoperators of $\mathcal{E}$ also as 'right eigenoperators', while the eigenoperators of the adjoint $\mathcal{E}^\dagger$ are called 'left eigenoperators' of $\mathcal{E}$.

Chapter 3

Classical and Quantum Markov Chains

In this chapter we are going to briefly introduce the notion of Markov process. The first part goes through the classical theory, which is covered more thoroughly in [19]. The second part defines a quantum notion of Markov chains, using some notions of quantum channels and quantum information that can be found in [27, 18].

3.1 The Markov property

In this section we introduce the concept of Markov chain as a stochastic process, which is at the heart of many classical algorithms. We first give the definitions as given in the classical theory, then we will see how we can use the elements of quantum theory introduced in Chapter 1 to define a notion of *quantum Markov chain*.

Definition 3.1 *A stochastic process is a sequence of random variables defined over some probability space:*

$$\{X(t) : t \in T\}$$

We can also see a stochastic process as a random function $X : T \to \mathbb{R}$. In usual applications, the variable t takes the meaning of time, and we split the

theory into two cases: discrete-time ($T = \mathbb{N}$) and continuous-time ($T = \mathbb{R}_0^+$). In this work, we will restrict ourselves to the discrete-time case, as a great part of what is discussed here can be naturally extended to the continuous-time case.

There are different types of stochastic processes in the literature. The one we want to focus on is the so-called *Markov process*.

Definition 3.2 *A (discrete-time) Markov process is a stochastic process $\{X_t\}_{t\in\mathbb{N}}$ which satisfies the Markov property, i.e.*

$$\Pr\left[X_{t+1} = x_{t+1} \,|\, X_t = x_t, \ldots, X_t = x_t\right] = \Pr\left[X_{t+1} = x_{t+1} \,|\, X_t = x_t\right]$$

More informally, the Markov property asserts that the state at time t, only depends on the value of the state at time $t-1$.

3.2 Classical Markov chains

Consider a Markov process $\{X_t\}_{t\in\mathbb{N}}$ where the X_t's are drawn from a finite set of states S. We want to find a relation between the distribution of the state X_t and the one of its predecessor X_{t-1}. Here we can use the law of total probability:

$$\Pr\left[X_t = y\right] = \sum_{x\in S} \Pr\left[X_t = y \,|\, X_{t-1} = x\right] \Pr\left[X_{t-1} = x\right] \tag{3.1}$$

We introduce the following notation:

Definition 3.3 *A* stochastic vector *over S is a vector $q \in [0,1]^S$ such that its entries sum up to* 1*, i.e. $\sum_{s\in S}(q_t)_s = 1$.*

You can see that such vector represents a probability distribution over S. We use q_t to denote the stochastic vector denoting the probability distribution of X_t, also called the *state probability vector* of X_t.

On the other hand, the conditional probabilities given in Eq. (3.1) can be collected into a *transition matrix* P_t:

$$(P_t)_{xy} := \Pr\left[X_t = y \,|\, X_{t-1} = x\right]$$

One can see that the rows of P_t are stochastic vectors, and this gives the definition of *stochastic matrix*.

Definition 3.4 *A* stochastic matrix *is an entry-wise non-negative matrix where each row sums up to* 1*.*

Using this new notation, Eq. (3.1) can be written as a matrix-vector multiplication:

$$q_t = q_{t-1} P_t$$

And this gives a general way to compute the distribution of X_t starting from an initial distribution:

$$q_t = q_0 P_1 P_2 \cdots P_t$$

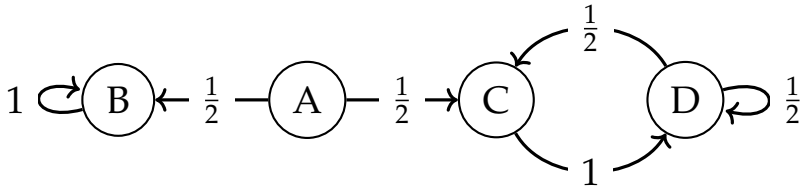
From now on, we will focus on *homogeneous* Markov processes, which means that the transition matrix $P_t \equiv P$ does not depend on time. This further simplifies Eq. (3.1):

$$q_t = q_0 P^t \tag{3.2}$$

Therefore, a Markov process can be seen as a linear, time-invariant dynamic process, and this suggests that the spectral properties of P may play a crucial role in the analysis of many interesting behaviours. This gives us everything to introduce the definition of Markov chain:

Definition 3.5 *A (discrete-time) Markov chain is a tuple* (S, P) *where*

- S *is a finite set of states;*
- $P \in [0,1]^{|S| \times |S|}$ *is a stochastic matrix called* transition matrix*.*



$$P = \begin{bmatrix} 0 & \frac{1}{2} & \frac{1}{2} & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & \frac{1}{2} & \frac{1}{2} \end{bmatrix}$$

Figure 3.1: Graphical representation of a Markov chain. The vertices of the graph are sites where the walker can be found. At each step, the walker chooses one of the outgoing edges following the weights as probabilities. Each row of the stochastic matrix gives the probability of being in each of the states in the next step, starting from the corresponding initial site.

A Markov chain is thus a model that induces a Markov process. In particular, we can see this model as a graph $G = (V, E)$ where $V = S$, and an edge from x to y is added with weight P_{xy} whenever this value is non-zero (Fig. 3.1). This is why Markov chains are also called *random walks* in the literature: they can be seen as a process involving an entity moving in a graph in a random way, following some predefined probability distributions.

We start by giving two important observations which will be useful later.

Observation 3.6 ([27]) *Any stochastic process P satisfies $P\mathbf{1} = \mathbf{1}$, where $\mathbf{1}$ is the vector of all ones.*

Proof It is sufficient to see that each entry is the inner product between a row of P and $\mathbf{1}$, which always gives 1. □

Also, we notice that 1 is the highest possible absolute value for an eigenvalue of a stochastic matrix.

Lemma 3.7 ([27]) *Every eigenvalue λ of P satisfies $|\lambda| \leq 1$.*

Proof Let us consider a (possibly complex) eigenvalue $\lambda \neq 1$ of P. Denoting with $\mathbf{x}^\top \neq \mathbf{0}^\top$ an associated (possibly complex) eigenvector we have

$$\sum_i x_i P_{ij} = \lambda x_j \Longrightarrow |\lambda||x_j| = \left|\sum_i x_i P_{ij}\right| \leq \sum_i |x_i| P_{ij}$$

implying $|\lambda|\mathbf{x}_*^\top \leq \mathbf{x}_*^\top P$, where $\mathbf{x}_*$ is computed from $\mathbf{x}$ by taking the entry-wise absolute value. Multiplying with the vector of all ones we obtain

$$|\lambda|\mathbf{x}_*^\top \mathbf{1} \leq \mathbf{x}_*^\top P\mathbf{1} = \mathbf{x}_*^\top \mathbf{1} \Longrightarrow |\lambda| \leq 1 \qquad \square$$

3.3 A quantum theory of Markov chains

Now that we have introduced classical Markov chains, we would like to extend these ideas to the quantum case. Starting from the set of states S, we define a quantum system over a Hilbert space $\mathcal{H} = \text{span}\{|s\rangle : s \in S\}$.

Now we can define a Markov process $\{|\Phi_t\rangle\}_{t\in\mathbb{N}}$ where $|\Phi_t\rangle$ can be seen as a random variable taking values in $\mathcal{H}$. From Chapter 1 we know that, in the most general case, such random variable can be represented by a distribution given by the following density operator:

$$\rho_t = \sum_i p_i |\phi_i\rangle\langle\phi_i|$$

This suggests that we can use the notion of density matrix as quantum counterpart of the state probability vectors defined earlier. Indeed, one can see that, if we use the states of S as elements of the computational basis, the diagonal elements of density operators form a stochastic vector giving the distribution for an hypothetical measurement of this state, while the off-diagonal elements can be seen as representing the quantum coherences.

Hence, assuming an homogeneous Markov process, we can see the transition from the distribution ρ_t of $|\Phi_t\rangle$ to the distribution ρ_{t+1} of $|\Phi_t\rangle$ as a

quantum channel:

$$\rho_{t+1} = \mathcal{E}(\rho_t) \Longrightarrow \rho_t = \mathcal{E}^t(\rho_0)$$

This is perfectly in line with Eq. (3.2), and starting from this analogy we give the definition of quantum Markov chain.

Definition 3.8 *A (homogeneous) quantum Markov chain is defined as a tuple* $(\mathcal{H}, \mathcal{E})$ *where:*

- $\mathcal{H}$ *is a finite-dimensional Hilbert space;*
- $\mathcal{E} : \bar{\mathcal{H}} \to \bar{\mathcal{H}}$ *is a completely positive, trace-preserving map.*

Again, we can see an analogy with the classical theory: the complete positivity of $\mathcal{E}$ in the quantum case is needed to ensure that the probabilities of the density operator are kept non-negative, in the same way the non-negativity of the entries of P ensure this for the entries of a stochastic vector.

The same thing can be seen for trace preservation: a trace preserving map preserves the fact that the probabilities sum up to 1, also ensured by the stochasticity of P for stochastic vectors.

As we did for the classical case, we give the following two observations:

Observation 3.9 *For any completely-positive trace-preserving map* $\mathcal{E}$*, the adjoint map* $\mathcal{E}^\dagger$ *satisfies* $\mathcal{E}^\dagger(\mathbb{1}) = \mathbb{1}$.

Proof We proved in Chapter 2 that the adjoint of a completely positive trace-preserving map is unital. □

This gives the exact same guarantee as Observation 3.6: in the classical theory, state probability vectors are multiplied on the left, and in other words the claim states that **1** turns into **1** again if applied to the adjoint of P.

Also Lemma 3.7 can be extended to the quantum case.

Lemma 3.10 *Every eigenvalue* λ *of* $\mathcal{E}$ *satisfies* $|\lambda| \leq 1$.

Proof The operator norm $||A||_\infty$ is defined as the highest eigenvalue of A. By the Russo-Dye inequality [4], we have that:

$$||\mathcal{E}(X)||_\infty \leq ||\mathcal{E}(\mathbb{1})||_\infty ||X||_\infty$$

Take X such that $\mathcal{E}(X) = \lambda X$. If $\mathcal{E}$ is unital, we have

$$|\lambda| ||X||_\infty = ||\mathcal{E}(X)||_\infty \leq ||\mathcal{E}(\mathbb{1})||_\infty ||X||_\infty = ||\mathbb{1}||_\infty ||X||_\infty = ||X||_\infty$$

implying $|\lambda| \leq 1$. If $\mathcal{E}$ is trace-preserving, it is sufficient to apply the argument to the dual $\mathcal{E}^\dagger$, which has the same spectrum. □

3.4 Quantum Markov chains and unitary walks

A great part of research on quantum walks, is dedicated on *unitary walks*, which is also one of the first models who tried to extend the ideas of Markov chains to the quantum setting [12, 23]. The idea is to repeatedly apply a unitary U to a quantum state $|\psi_0\rangle \in \mathcal{H}$:

$$|\psi_t\rangle = U^t |\psi_0\rangle$$

These unitary walks are shown to exhibit unexpected behaviours, mainly due to quantum interference not present in the classical theory [12, 13]. However, these walks are essentially different from classical Markov chains: not every classical Markov chain can be represented as a unitary walks, in particular due to the fact that unitary dynamics is reversible. The definition of quantum Markov chains presented in this work is therefore needed as a framework to unify the two theories, in order to better characterize the differences and to get a better understanding of where quantum speed-ups may arise. A similar program has been carried out in [26] for the continuous-time version, where the authors showed that the Lindblad equation can be also used to express a classical Chapman-Kolmogorov evolution.

- A unitary walk of a state in $\mathcal{H}$ governed by a unitary U can be formalized as a special case of quantum Markov chain $(\mathcal{H}, \mathcal{U})$ where

$$\mathcal{U}(\rho) = U\rho U^\dagger$$

- A classical Markov chain (S, P) can be seen as a quantum Markov chain $(\mathcal{H}, \mathcal{P})$ where $\mathcal{H} = \text{span}\{|s\rangle : s \in S\}$ and

$$\mathcal{P}(\rho) = \sum_{x,y} P_{xy} |y\rangle\langle x|\rho|x\rangle\langle y|$$

  i.e. a set of Kraus operators for $\mathcal{P}$ are $\{\sqrt{P_{xy}}|y\rangle\langle x|\}_{x,y}$.

Another good reason to explore the quantum Markov chain model is that we do not exclude potential speed-ups given by unitary evolution subject to decoherence: for example, some numerical simulations by Kendon and Tregenna showed that an hybrid quantum-classical walk can achieve better performances than both a fully classical and a fully unitary counterpart [14, 15].

Moreover, a notion of quantum random walk like the one given by quantum Markov chains can capture irreversible quantum dynamics in great generality, allowing to analyze the performances of such walks under non-ideal (noisy) environments, which can be interesting for practical applications in near-term quantum computing.

Chapter 4

Hitting Times

4.1 Classical hitting times

One of the most important quantities used in algorithmic applications of Markov chains is the notion of *hitting time*. The idea is the following: we would like to reach a particular state (or any state in a particular subset). If we start from a particular state (or distribution of states), how long do we have to wait in expectation in order to *hit* such state?

We start by seeing how these times are analyzed classically. A complete approach on the classical theory of hitting times can be also found in [19].

Definition 4.1 *Define the following random variable:*

$$T_z(q) = \min\{t > 0 : X_t = y, X_0 \sim q\}$$

i.e. the time it takes to hit state z starting from state distributed according to the stochastic vector q. The hitting time *is defined as the expectation of such variable:*

$$h_z(q) := \mathbb{E}\left[T_z(q)\right]$$

Here we derive a nice well-known formula for hitting times, which can also be seen as a consequence of the recurrence relation for hitting times found in [19].

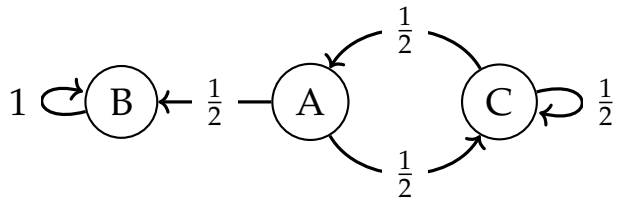


Figure 4.1: Simple example where the hitting time can be infinite. Suppose we start from A and we want to reach C: since we have a non-zero probability to end up in B (and get stuck forever there), in case this happens, the hitting time from B would be infinity. From a mathematical point of view, the hitting time from A will be infinity as well by the law of total expectation.

Theorem 4.2 *Fixed a state $z \in S$, let e_z be the z-th vector of the standard basis, and let $\Pi_{-z} = \mathbb{1} - e_z e_z^T$. The hitting time for the state z can be computed as:*

$$h_z(q) = \sum_{k=0}^{\infty} q(P\Pi_{-z})^k \mathbf{1}$$

If all the eigenvalues of $P\Pi_{-z}$ are strictly less than 1 in absolute value, than the series converges and it is equal to:

$$h_z(q) = q(\mathbb{1} - P\Pi_{-z})^{-1}\mathbf{1}$$

Before proving the claim, we stress that the above series may diverge to infinity. This is the case when it is possible to end up in a component of the chain from which we will never be able to reach z (Figure 4.1).

Proof Let E be the event occurring when we hit z after the first step. Using the law of total expectation we obtain:

$$\begin{aligned}
h_z(q) &= \mathbb{E}\left[T_z(q)\right] \\
&= \mathbb{E}\left[T_z(q) \mid E\right] \Pr\left[E\right] + \mathbb{E}\left[T_z(q) \mid \bar{E}\right] \Pr\left[\bar{E}\right] \\
&= \Pr\left[E\right] + \left(1 + h_z\left(\frac{qP\Pi_{-z}}{qP\Pi_{-z}\mathbf{1}}\right)\right) \Pr\left[\bar{E}\right] \\
&= 1 + h_z\left(\frac{qP\Pi_{-z}}{qP\Pi_{-z}\mathbf{1}}\right) \Pr\left[\bar{E}\right] \\
&= q\mathbf{1} + h_z\left(\frac{qP\Pi_{-z}}{qP\Pi_{-z}\mathbf{1}}\right) qP\Pi_{-z}\mathbf{1}
\end{aligned}$$

where one can see that the vector $\frac{qP\Pi_{-z}}{qP\Pi_{-z}\mathbf{1}}$ contains the conditional distribution $\Pr[\cdot|\bar{E}]$ of the state after one step (indeed, the quantity $qP\Pi_{-z}\mathbf{1}$ is exactly $\Pr[\bar{E}]$). By inductively applying the above relation, we obtain:

$$h_z(q) = \sum_{k=0}^{N} q(P\Pi_{-z})^k\mathbf{1} + h_z\left(\frac{q(P\Pi_{-z})^{N+1}}{q(P\Pi_{-z})^{N+1}\mathbf{1}}\right) q(P\Pi_{-z})^{N+1}\mathbf{1}$$

and, in the limit, we obtain the claimed series. The second part of the claim comes from a geometric sum. □

The results in this section can be generalized to an arbitrary subset of states $Z \subseteq S$, in order to compute the expected first time to hit any state in Z.

4.2 Quantum hitting time

From the classical theory of Markov chains we can see that the hitting time is the first time the Markov chain *hits* a state. Defining a notion of *quantum hitting time* that is satisfying enough for our purposes and easy to estimate turns out to be a real challenge: what does it mean for an element of the Markov process to *hit* a certain state? Moreover, in order to determine whether a state is hit or not in a certain step we have to measure such state, and this may compromise the behaviour of the walk. Here we present an extension of two different notions of quantum hitting time originally proposed by Kempe in [13].

Definition 4.3 *We say that $t \in \mathbb{N}$ is a p-*one-shot hitting time *for the state $|z\rangle$ if the probability of measuring $|z\rangle$ after t steps is at least p:*

$$\mathrm{Tr}\left[|z\rangle\langle z|\mathcal{E}^t(\rho)\right] \geq p$$

Definition 4.4 *Let $(\mathcal{H}, \mathcal{E})$ be a quantum Markov chain, and consider the measurement channel:*

$$\begin{aligned}\mathcal{T}_z(\rho) &= \Pi_z\rho\Pi_z + (\mathbb{1} - \Pi_z)\rho(\mathbb{1} - \Pi_z)\\ &=: \mathcal{P}_z(\rho) + \mathcal{P}_{-z}(\rho)\end{aligned}$$

i.e. $\mathcal{T}_z$ measures whether the state is z or not. The $|z\rangle$-measured quantum Markov chain is the chain $(\mathcal{H}, \mathcal{P}_{-z} \circ \mathcal{E})$.

Notice that $\mathcal{P}_{-z}$ here is not trace-preserving: it will remove any probability mass in the target subspace. This represents the fact that the chain will stop upon the hit of a target state.

Definition 4.5 *Define the following random variable:*

$$T_z(\rho) := \min \left\{ t \geq 1 \middle| |z\rangle \text{ is measured at time } t \right\}$$

The concurrent hitting time *for the state $|z\rangle$ is the expected value of this variable:*

$$h_z(\rho) := \mathbb{E}_{|\Phi_0\rangle \sim \rho} \left[T_z \right]$$

Here $|\Phi_0\rangle \sim \rho$ means that $|\Phi_0\rangle$ is a random variable distributed according to the eigenbasis of ρ.

Please note that these two notions of hitting time are essentially different, starting from the way an algorithm exploits them: while with the one-shot hitting time one has to run the Markov chain as it is, and then measure after a fixed and precise amount of steps, with the concurrent hitting time one needs to measure at each step *whether the state is hit or not*. Also we want to stress that the measurement used in the process, as specified in Definition. (4.4) is not a measurement in the computational basis: this represents an observable yielding $+1$ if the state collapses to $|z\rangle$, and -1 if it collapses within the subspace orthogonal to the one spanned by $|z\rangle$. Then, the value of z can be retrieved by doing a full measurement of the state after $+1$ is observed. We also remind that doing a full measurement at every step gives a behaviour of the walk which is efficiently reproducible by a classical Markov chain (by using the elements of the measurement basis as states).

4.3 Rewriting the quantum concurrent hitting time

We want to find a closed form expression for the concurrent hitting time of Definition 4.5. Notice that we only defined $h_z(\rho)$ for operators of trace 1: we extend this definition to any operator of trace $\neq 0$ with the following relation:

$$h_z(\rho) := \operatorname{Tr}(\rho) \cdot h_z(\rho / \operatorname{Tr}(\rho))$$

Moreover, we let $h_z(0) = 0$.

Lemma 4.6 *$h_z(\rho)$ is linear in ρ, for non-traceless operators.*

Proof Whenever $\rho = \sum_i p_i \rho_i$ is a convex combination and $\operatorname{Tr} \rho_i = 1$ for every i, linearity follows by the law of total expectation: if A_i is the event in which we have ρ_i as initial state, then

$$h_z\left(\sum_i p_i \rho_i\right) = \sum_i \mathbb{E}\left[T_z \mid A_i\right] \Pr\left[A_i\right] = \sum_i h_z(\rho_i) p_i$$

Consider an operator ρ with $\operatorname{Tr} \rho = r \neq 0$ and a scalar $a \neq 0$:

$$h_z(a\rho) = ar \cdot h_z\left(\frac{a\rho}{ar}\right) = ar \cdot h_z\left(\frac{\rho}{r}\right) = a \cdot h_z(\rho)$$

The case $a = 0$ follows from $h_z(0) = 0$.

Now consider two operators ρ, σ with $\operatorname{Tr} \rho = r, \operatorname{Tr} \sigma = s$ and $r, s, r+s \neq 0$.

$$\begin{aligned}
h_z(\rho + \sigma) &= (r+s) \cdot h_z\left(\frac{\rho + \sigma}{r+s}\right) \\
&= (r+s) \cdot h_z\left(\frac{r\rho}{r(r+s)} + \frac{s\sigma}{s(r+s)}\right) \\
&= (r+s) \cdot h_z\left(\frac{r\rho}{r(r+s)}\right) + (r+s) \cdot h_z\left(\frac{s\sigma}{s(r+s)}\right) \\
&= (r+s) \cdot h_z\left(\frac{\rho}{r+s}\right) + (r+s) \cdot h_z\left(\frac{\sigma}{r+s}\right) \\
&= h_z(\rho) + h_z(\sigma)
\end{aligned}$$

where the third equality holds because of the convex combination. □

To conclude, we can finally extend this definition to traceless operators by applying linearity:

$$h_z(\rho) = h_z(\mathbb{1} + \rho - \mathbb{1}) := h_z(\mathbb{1} + \rho) - h_z(\mathbb{1})$$

This gives us a function $h_z(\rho)$ which is linear in $\rho \in \bar{\mathcal{H}}$. We now find a solution for such function.

Theorem 4.7 *The z-concurrent hitting time can be computed as:*

$$h_z(\rho) = \sum_{k=0}^{\infty} \mathrm{Tr}(\mathcal{E}_{-z}^k(\rho))$$

where $\mathcal{E}_{-z} = \mathcal{P}_{-z} \circ \mathcal{E}$, and $\mathcal{P}_{-z}(\rho) := (\mathbb{1} - \Pi_z)\rho(\mathbb{1} - \Pi_z)$.

If all the eigenvalues λ of $\mathcal{E}_{-z}$ have $|\lambda| < 1$, then the series converges and:

$$h_z(\rho) = \mathrm{Tr}\left[(\mathcal{I} - \mathcal{E}_{-z})^{-1}(\rho)\right]$$

where $\mathcal{I}$ is the identity superoperator in $\bar{\mathcal{H}}$.

Before we prove the result, notice the condition we impose on the eigenvalues of $\mathcal{E}_{-z}$: if we have an eigenvalue $\lambda = 1$ for $\mathcal{E}_{-z}$ then $\mathcal{I} - \mathcal{E}_{-z}$ is not invertible. In terms of the Markov chain, this means there is a starting state ρ (an associated eigenstate) from which we will never measure z (in particular, since this state is always orthogonal to z, $\mathcal{P}_{-z}(\rho) = \rho$ and thus also $\mathcal{E}(\rho) = \rho$ must hold). If eigenvalues of the form $e^{i\theta} \neq 1$ are present, then the limit (and thus also the expectation) does not even exist.

Proof Consider the first step of the chain, and let E be the event in which z is measured in this step. By the law of total expectation we get:

$$h_z(\rho) = \mathbb{E}_\rho\left[T_z \mid E\right] \Pr\left[E\right] + \mathbb{E}_\rho\left[T_z \mid \bar{E}\right] \Pr\left[\bar{E}\right]$$

When E occurs, the hitting time is exactly 1 (we only took one step to measure z). Otherwise, by the Markov property the hitting time is the one given

by the chain starting at the post measurement state $\mathcal{E}_{-z}(\rho)/\operatorname{Tr}(\mathcal{E}_{-z}(\rho))$ plus one (which is the step we already took). Therefore the relation becomes:

$$\begin{aligned} h_z(\rho) &= \Pr[E] + \left(1 + h_z\left(\frac{\mathcal{E}_{-z}(\rho)}{\operatorname{Tr}(\mathcal{E}_{-z}(\rho))}\right)\right)\Pr[\bar{E}] \\ &= 1 + h_z\left(\frac{\mathcal{E}_{-z}(\rho)}{\operatorname{Tr}(\mathcal{E}_{-z}(\rho))}\right)\Pr[\bar{E}] \\ &= 1 + h_z\left(\frac{\mathcal{E}_{-z}(\rho)}{\operatorname{Tr}(\mathcal{E}_{-z}(\rho))}\right)\operatorname{Tr}(\mathcal{E}_{-z}(\rho)) \\ &= 1 + h_z(\mathcal{E}_{-z}(\rho)) \end{aligned}$$

By inductively applying this relation we obtain, for every $N \in \mathbb{N}$:

$$h_z(\rho) = \sum_{k=0}^{N-1} \operatorname{Tr}(\mathcal{E}_{-z}^k(\rho)) + h_z(\mathcal{E}_{-z}^N(\rho))$$

which, in the limit as $N \to \infty$, gives the claimed series (notice that $\mathcal{E}_z^N(\rho) \to 0$ since all the eigenvalues have absolute value < 1). The rest of the claim comes from a geometric sum. □

The above definitions and results immediately extend to an arbitrary subset of states, i.e. an arbitrary subspace of $\mathcal{H}$: it is sufficient to replace Π_z with the projector Π_S onto such subspace. One can appreciate how the expression given by Theorem 4.7 is essentially the same as the one for the classical hitting time of Theorem 4.2.

4.4 Example: single-qubit Hadamard walk

We start building intuition on this hitting time with a simple example. Consider a quantum Markov chain $(\mathcal{H}, \mathcal{E})$ where $\mathcal{H} = \operatorname{span}\{|0\rangle, |1\rangle\}$ is the Hilbert space of a single qubit and $\mathcal{E}(\rho) = H\rho H$ is the application of the Hadamard gate:

$$H = \frac{1}{\sqrt{2}}\begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}$$

We want to answer the following question: if we start from a certain state ρ and we measure in the computational basis at each step, how many steps in expectation do we need to run before we measure $|1\rangle$? In other words, what is the value of $h_1(\rho)$?

In order to compute the concurrent hitting time, we start by computing the spectral decomposition of the matrix:

$$\Pi_{-1}H = \Pi_0 H = \frac{1}{\sqrt{2}}\begin{bmatrix} 1 & 1 \\ 0 & 0 \end{bmatrix}$$

We can take the spectral decomposition of this matrix:

$$\begin{cases} \Pi_0 H|0\rangle = \frac{1}{\sqrt{2}}|0\rangle \\ \Pi_0 H|-\rangle = 0 \end{cases}$$

and this also gives the spectral decomposition of the map $\mathcal{E}_{-1}$ and, consequently, the one of $(\mathcal{I} - \mathcal{E}_{-1})^{-1}$

$$\begin{cases} \mathcal{E}_{-1}(|0\rangle\langle 0|) = \frac{1}{2}|0\rangle\langle 0| \\ \mathcal{E}_{-1}(|-\rangle\langle -|) = 0 \\ \mathcal{E}_{-1}(|-\rangle\langle 0|) = 0 \\ \mathcal{E}_{-1}(|0\rangle\langle -|) = 0 \end{cases} \implies \begin{cases} (\mathcal{I} - \mathcal{E}_{-1})^{-1}(|0\rangle\langle 0|) = 2|0\rangle\langle 0| \\ (\mathcal{I} - \mathcal{E}_{-1})^{-1}(|-\rangle\langle -|) = |-\rangle\langle -| \\ (\mathcal{I} - \mathcal{E}_{-1})^{-1}(|-\rangle\langle 0|) = |-\rangle\langle 0| \\ (\mathcal{I} - \mathcal{E}_{-1})^{-1}(|0\rangle\langle -|) = |0\rangle\langle -| \end{cases}$$

This spectral decomposition gives all we need in order to compute $h_1(\rho)$. It is sufficient to decompose ρ with respect to the eigenbasis of $\mathcal{I} - \mathcal{E}_{-1}$ and then use linearity. A notable example is given by the hitting times starting from the computational basis states:

$$h_1(|0\rangle\langle 0|) = 2$$
$$h_1(|1\rangle\langle 1|) = 2h_1(|-\rangle\langle -|) - \sqrt{2}h_1(|0\rangle\langle -|) - \sqrt{2}h_1(|-\rangle\langle 0|) + h_1(|0\rangle\langle 0|) = 2$$

where we decomposed $|1\rangle\langle 1| = 2|-\rangle\langle -| - \sqrt{2}|0\rangle\langle -| - \sqrt{2}|-\rangle\langle 0| + |0\rangle\langle 0|$ and applied linearity.

Thus, this tells us that it takes 2 steps in expectation to measure a $|1\rangle$, starting from either $|0\rangle$ or $|1\rangle$[1].

4.5 Measurement schemes

One can see, by the form of the expression and the nature of the process, that the notion of concurrent hitting time seems to be the natural quantum generalization of the classical hitting time. The main difference is that, in the classical case, measurements do not alter the state, and thus the behaviour of the walk.

As we will see (cf. Chapter 5), in some cases measuring the hit at each step may result in the loss of potential quantum speedups compared to the fully quantum coherent case. We now try to extend the notion of hitting time to explore the impact of measurement in a more general case: for example, what happens if we measure at each step only with probability $p \in (0,1)$, or we wait a certain number of steps extracted from a particular probability distribution? In order to analyze such cases we give the following definition:

Definition 4.8 *A* measurement scheme *is a discrete probability distribution for a time $T \in \mathbb{N}^+$. This represents how many steps we run before taking a measurement as given in Definition 4.4.*

Given a quantum channel $\mathcal{E}$ and a measurement scheme σ, we define the *first-measurement* channel as the expected quantum channel following σ applied before carrying out the first measurement:

$$\mathcal{E}^{\sigma}(\rho) := \mathbb{E}_{T\sim\sigma}\left[\mathcal{E}^{T}(\rho)\right] = \sum_{t=1}^{\infty} \Pr_{\sigma}\left[T = t\right] \mathcal{E}^{t}(\rho)$$

where T is a random variable we call *time to measure* (or TTM). We use this to give a generalized notion of concurrent hitting time.

[1]Please keep in mind that the hitting time from $|1\rangle$ is not 0 because we do not measure the hit on the starting state.

Definition 4.9 *Let $(\mathcal{H}, \mathcal{E})$ be a quantum Markov chain, and let σ be a measurement scheme. We define the following process: (1) sample a time $t \sim \sigma$, (2) run the chain for t steps and (3) try a measurement with $\mathcal{T}_z$ as defined in Definition 4.4. We repeat these three steps until the hit of z is measured.*

The *generalized concurrent hitting time* under σ, which we denote by $h_z^\sigma(\rho)$ is the expectation of the number $T_z^\sigma(\rho)$ of steps of the Markov chain (i.e. the number of times we apply $\mathcal{E}$) this process will take in order to stop.

Theorem 4.10 *Let $(\mathcal{H}, \mathcal{E})$ be a quantum Markov chain, and let σ be a measurement scheme. The generalized concurrent hitting time with respect to σ can be computed as:*

$$h_z^\sigma(\rho) = \mathbb{E}_{T\sim\sigma}[T] \cdot \mathrm{Tr}\left[(\mathcal{I} - \mathcal{E}_{-z}^\sigma)^{-1}(\rho)\right]$$

where $\mathcal{E}_{-z}^\sigma(\rho) = \mathcal{P}_{-z} \circ \mathcal{E}^\sigma(\rho)$.

The above formula can be intuitively interpreted as follows: the trace quantifies the expected number of *measurements* we carry out before stopping. However, the number of steps before actually *trying* a measurement is $\mathbb{E}_\sigma[T]$ in expectation.

Proof We follow a similar argument as for Theorem 4.7. By the exact same arguments, we will extend the definition of $h_z(\rho)$ such that it will be linear in $\rho \in \bar{\mathcal{H}}$. We apply the law of total expectation on the time T we wait before trying the first measurement.

$$h_z^\sigma(\rho) = \sum_{t=1}^{\infty} \left[t + h_z^\sigma(\mathcal{P}_{-z} \circ \mathcal{E}^t(\rho))\right] \cdot \Pr_\sigma[T = t]$$

The conditional expectation here is given by the fact that, whenever $T = t$, we spend t steps and then we are in the state $\mathcal{E}^t(\rho)$, just before applying the

first measurement. By splitting the sum we obtain:

$$\begin{aligned} h_z^\sigma(\rho) &= \sum_{t=1}^{\infty} t \cdot \Pr_\sigma[T=t] + \sum_{t=1}^{\infty} h_z^\sigma(\mathcal{P}_z \circ \mathcal{E}^t(\rho)) \cdot \Pr_\sigma[T=t] \\ &= \mathbb{E}_{T\sim\sigma}[T] + h_z^\sigma\left[\mathcal{P}_z \circ \left(\sum_{t=1}^{\infty} \mathcal{E}^t(\rho) \cdot \Pr_\sigma[T=t]\right)\right] \\ &= \mathbb{E}_{T\sim\sigma}[T] + h_z^\sigma\left[\mathcal{P}_z \circ \mathcal{E}^\sigma\right] \end{aligned}$$

This relation is identical to the one we found in the proof of Theorem 4.7, except that the 1 is replaced with the expectation of T. Thus, the concluding argument is the same as in Theorem 4.7. □

An interesting measurement scheme is the geometric distribution, because such measurement scheme can be easily (memorylessly) implemented by the following quantum channel:

$$\mathcal{B} = p\mathcal{T}_z \circ \mathcal{E} + (1-p)\mathcal{E}$$

i.e. at every step we apply the measurement with probability p. Under such distribution, we can also find a compact and simple way of computing $\mathcal{E}^\sigma$.

Lemma 4.11 *If $\sigma = Geom(p)$, then $\mathcal{E}^\sigma(\rho) = p\mathcal{E} \circ (\mathcal{I} - (1-p)\mathcal{E})^{-1}(\rho)$.*

Proof We use some properties of the geometric sum:

$$\begin{aligned} \mathcal{E}^\sigma(\rho) &= \sum_{t=1}^{\infty} p(1-p)^{t-1}\mathcal{E}^t(\rho) \\ &= \sum_{t=0}^{\infty} p(1-p)^t\mathcal{E}^{t+1}(\rho) \\ &= p\mathcal{E} \circ \sum_{t=0}^{\infty} \left[(1-p)\mathcal{E}\right]^t(\rho) \\ &= p\mathcal{E} \circ (\mathcal{I} - (1-p)\mathcal{E})^{-1}(\rho) \end{aligned}$$

□

4.6 Example: walk on the cycle

We now consider a simple example: suppose to have a cycle of length $2N$ (Figure 4.2), and in particular, a Hilbert space $\mathcal{H} = \mathcal{H}_C \otimes \mathcal{H}_S$ where $\mathcal{H}_C =$

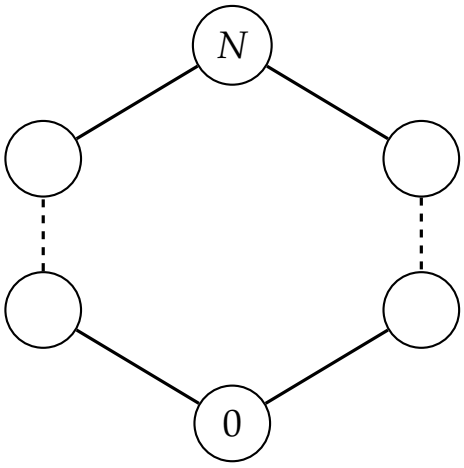


Figure 4.2: Quantum walk on a $2N$-cycle. We start from the state $|0\rangle$ (with coin register set to $|\uparrow\rangle$, and we want to reach the state $|N\rangle$. Both paths connecting states $|0\rangle$ and $|N\rangle$ are of length N.

span $\{|\uparrow\rangle, |\downarrow\rangle\}$ represents a coin register, while $\mathcal{H}_S = \text{span}\{|x\rangle : 0 \leq x < 2N\}$ represents the position of the walker on the cycle. Considering a unitary Markov chain $(\mathcal{H}, \mathcal{U})$ where $\mathcal{U}(\rho) = U\rho U^\dagger$, we start in the pure state $|\uparrow, 0\rangle$, and the unitary U is given by:

$$U = S(H_C \otimes \mathbb{1}_S)$$

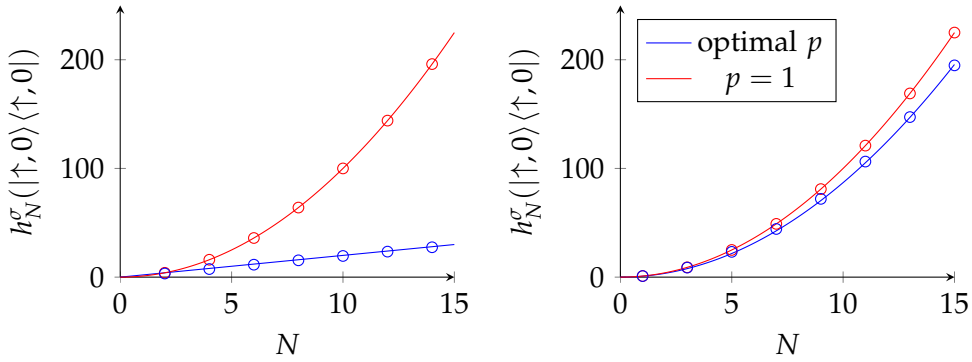
where H is the Hadamard gate and S is the shift operator defined as:

$$S|\uparrow, x\rangle = |\uparrow, x + 1 \bmod 2N\rangle$$
$$S|\downarrow, x\rangle = |\downarrow, x - 1 \bmod 2N\rangle$$

This is a so-called *coined unitary walk*, and for the case of the cycle some properties such as mixing times are already treated by Aharonov et al. [1] and Kendon et al. [14, 15]. We want to analyze the process where we apply this Markov chain and at each step we measure the hit of states $|N\rangle$ with probability p. Theorem 4.10 tells us that:

$$h_N(|\uparrow, 0\rangle\langle\uparrow, 0|) = \frac{1}{p} \operatorname{Tr}\left[(\mathcal{I} - \mathcal{P}_{-N} \circ \mathcal{U}^\sigma)^{-1}(|\uparrow, 0\rangle\langle\uparrow, 0|)\right]$$

where $\sigma = Geom(p)$, and $\Pi_{-N} = \mathbb{1} - |\uparrow, N\rangle\langle\uparrow, N| - |\downarrow, N\rangle\langle\downarrow, N|$.



(a) N even (b) N odd

Figure 4.3: Plots of the concurrent hitting time of the opposite side of a cycle of length $2N$ with respect to N. The circles are the numerical computations of the formula given by Theorem 4.10, while the continuous lines are the regressions. For $p = 1$, the regression gives exactly N^2, while for the optimal p we obtain $\sim 2N$ (even N) and $\sim N^2/1.16$ (odd N).

In the absence of an analytical solution for the above formula, we show numerical results for some values of N (a detailed explanation on how hitting times are computed with Python can be found in Appendix A.2). When $p = 1$, the hitting time seems to stick at exactly N^2, giving no speed up at all compared to the classical case (which can be seen as a Gambler's ruin [20]).

If we consider the optimal p, one can see that there is a huge difference between the case of even and odd N: when N is odd, the speed up seems to be only constant, and in particular, it settles around $\simeq N^2/1.16$. On the other hand, when N is even the hitting time consistently stays little below $2N$.

A different situation is given in the asymmetric case, where we have a cycle of length $2N + 1$ and we want to hit the state N, thus having two paths of different parity separating the initial and target states. In this case, numerical simulations give evidence that the hitting time settles around $2N$ for any choice of N, both even and odd.

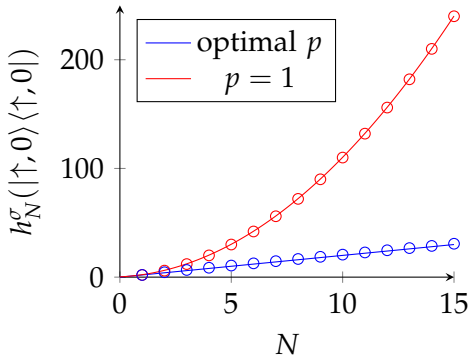


Figure 4.4: Plots of the concurrent hitting time of the opposite side of a cycle of length $2N + 1$ with respect to N. The circles are the numerical computations of the formula given by Theorem 4.10, while the continuous lines are the regressions. For $p = 1$, the regression gives $\sim (N + 1/2)^2$, while for the optimal p we obtain $\sim 2N$.

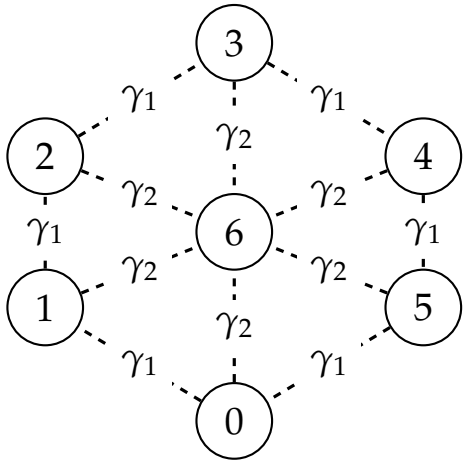


Figure 4.5: Quantum walk on a wheel on length $N = 6$. The values on the edges represent the couplings between the two states in the Hamiltonian H.

4.7 Example: quantum transport on the wheel

Here we show a slightly different example (Figure 4.5): consider N external sites, numbered from 0 to $N - 1$, plus a central site, the N-th one.

The dynamics is governed by the following Hamiltonian H:

$$H_{k,(k+1) \bmod N} = H_{k,(k-1) \bmod N} = \gamma_1$$
$$H_{k,N} = H_{N,k} = \gamma_2$$

i.e. the external sites are coupled in a cycle with strength γ_1, while the central site is coupled to each of the external sites with strength γ_2. Suppose we want to try a measurement of the hit of the central state after running the Schrödinger dynamics for some time t, i.e. we run a Markov chain $(\mathcal{H}, \mathcal{T}_N \circ \mathcal{U})$ where

$$\mathcal{U}(\rho) = e^{-iHt} \rho e^{iHt}$$

and we stop as soon as $\mathcal{T}_N$ makes the state collapse within the subspace spanned by $|N\rangle$. The expected number of steps before this happens is exactly given by $h_N(\cdot)$.

We would proceed like in Section 4.4 by computing the eigenvalues of $\Pi_{-N} e^{-iHt}$. However, notice the following:

Lemma 4.12 *For every N, the following vector is an eigenstate of H:*

$$|\psi\rangle = \frac{1}{\sqrt{N}} \sum_{x=0}^{N-1} e^{2\pi i x/N} |x\rangle$$

with associated eigenenergy $2\gamma_1 \cos(2\pi/N)$.

Proof Let $\alpha_x = e^{2\pi i x/N}$. One can see that the Hamiltonian H acts on the elements of the external sites as:

$$H|x\rangle = \gamma_1 |(x-1) \bmod N\rangle + \gamma_1 |(x+1) \bmod N\rangle + \gamma_2 |N\rangle$$

From now on we will assume that $x - 1$ and $x + 1$ are intended in $\mathbb{Z}_N$.

Applying H to $|\psi\rangle$ gives, by linearity:

$$\begin{aligned} H|\psi\rangle &= \frac{1}{\sqrt{N}} \sum_{x=0}^{N-1} \alpha_x H|x\rangle \\ &= \frac{1}{\sqrt{N}} \sum_{x=0}^{N-1} \alpha_x(\gamma_1|x-1\rangle + \gamma_1|x+1\rangle + \gamma_2|N\rangle) \\ &= \frac{\gamma_1}{\sqrt{N}} \sum_{x=0}^{N-1} (\alpha_{x-1} + \alpha_{x+1})|x\rangle + \frac{\gamma_2}{\sqrt{N}} \sum_{x=0}^{N-1} \alpha_x|N\rangle \end{aligned}$$

The second sum cancels out, as the α_x are the N-th roots of unity which sum to zero. On the other hand:

$$\begin{aligned} \alpha_{x-1} + \alpha_{x+1} &= e^{2\pi i(x+1)/N} + e^{2\pi i(x-1)/N} \\ &= e^{2\pi ix/N}(e^{+2\pi i/N} + e^{-2\pi i/N}) \\ &= \alpha_x \cdot 2\cos(2\pi/N) \end{aligned}$$

By plugging this in the above expression we obtain the claim. □

We found that $|\psi\rangle$ is also en eigenstate of e^{-iHt} for any time t of the evolution. This is reasonable, since H transports probability mass from the peripheral sites to the center in a uniform way, and since waves coming from the external sites are adding up with phases whose sum is zero, they all cancel out, giving a perfect example of destructive interference! Since $|\psi\rangle$ is orthogonal to $|N\rangle$, also $\Pi_{-N}e^{-iHt}$ has $|\psi\rangle$ associated with the same (unitary) eigenvalue.

Therefore, $|\psi\rangle\langle\psi|$ is an eigenstate of $\mathcal{U}_{-N}$ associated to 1, implying in particular that $\mathcal{I} - \mathcal{U}_{-N}$ is not invertible. This essentially yields an infinite hitting time for any initial state that has a non-zero overlap with $|\psi\rangle\langle\psi|$ (i.e. states which are a linear combination of the eigenstates of U_{-Z} with a non-zero coefficient for $|\psi\rangle\langle\psi|$).

Although in this case one can imagine the hitting time is infinite due to the perfect destructive interference on the central site, from a mathematical point of view this is not different from the classical example of Figure 4.1:

the subspace spanned by $|\psi\rangle\langle\psi|$ is a component where we get stuck forever, and it is disjoint (orthogonal) from our target state $|N\rangle\langle N|$.

Chapter 5

Hitting Times in Grover's Search

5.1 The unstructured search setting

Consider the following search problem: given $N = 2^n$, we are given a function $f : [N] \to \{0, 1\}$ as an oracle, with the promise that exactly one element x_0 is such that $f(x_0) = 1$. We want to find and return x_0. 'Unstructured' here means that the data governing f has essentially no structure that we can exploit in an hypothetical algorithm. Therefore one can see that, in the classical case, correct deterministic algorithms have to check $N - 1$ elements. Moreover, also randomized algorithms need to check $\Omega(N)$ elements in order to achieve non-trivial success probability.

5.2 Grover's Markov chain

A different situation comes in the quantum setting, thanks to a result due to Grover [7]. The idea is simple: the oracle implementing the function can be thought as a (classical) logical circuit $\mathcal{C}_f$.

Assuming without loss of generality that $\mathcal{C}_f$ is made only of NOT and AND gates (which is always possible, by universality of these gates), one can obtain a quantum circuit $\mathcal{Q}_f$ where each NOT gate is replaced by a Pauli X gate, and each AND gate is replaced by a Toffoli gate with a newly added

(a) Pauli X gate

(b) Toffoli gate

Figure 5.1: Circuit model of Pauli X and Toffoli gates. By using the third qubit as an auxiliary (ancilla) register, and setting it to 0, one can see that the Toffoli gate flips the value of the qubit if and only if the first two qubits are in state $|1\rangle$.

qubit initialized to the state $|0\rangle$ (see Figure 5.1).

One can also see that the depth of the circuits $\mathcal{C}_f, \mathcal{Q}_f$ (in their respective models), as well as the number of gates, are equivalent. On the other hand, we obtained a circuit that acts in the computational basis as follows:

$$|x\rangle|0\cdots0\rangle \mapsto |x\rangle|\phi_x\rangle|f(x)\rangle$$

where $|\phi_x\rangle$ denotes possible intermediate values of the computation. Moreover, we can construct the inverse circuit $\mathcal{Q}_f^\dagger$ by reapplying the same gates in inverse order.

Using a call from $\mathcal{Q}_f$ and $\mathcal{Q}_f^\dagger$ we can construct the so-called *bit oracle*, i.e. a unitary $\mathcal{O}_f$ such that:

$$\mathcal{O}_f|x\rangle_A|z\rangle_B = |x\rangle_A|z \oplus f(x)\rangle_B$$

where A is the register containing input for f, and B is the qubit where the output will be stored, in a reversible way. Grover [7] showed that, using this oracle it is possible to construct the following unitary:

$$R_{x_0} = 2|x_0\rangle\langle x_0| - \mathbb{1}$$

which, in geometric terms, is a reflection of the whole Hilbert space with respect to the vector $|x_0\rangle$. Combined with another unitary, which is an in-

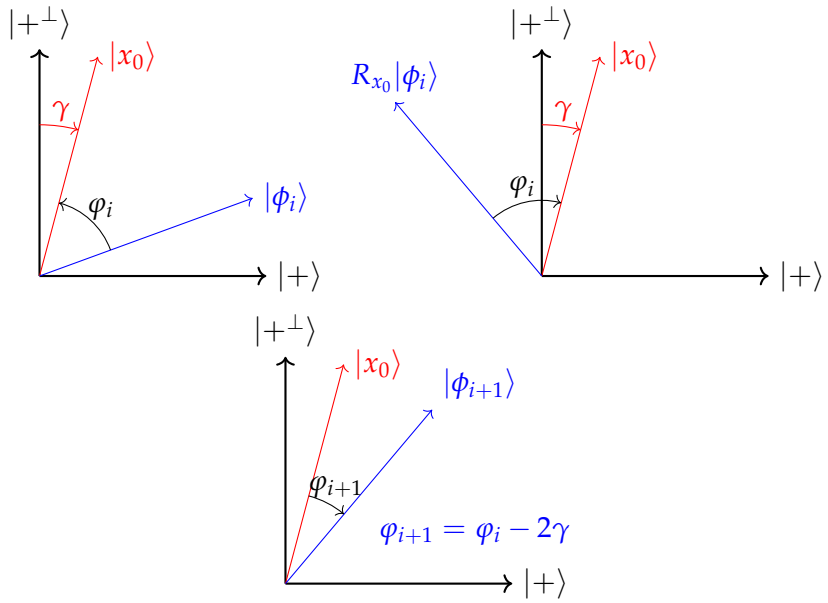


Figure 5.2: Implementation of Grover's rotation. This shows how a reflection around $|x_0\rangle$ (here represented by the transformation R_{x_0}) followed by an inversion with respect to $|+\rangle$ forms a rotation of an angle 2γ, where γ is the angle between $|x_0\rangle$ and $|+^\perp\rangle$.

version with respect to the Hadamard vector

$$|+\rangle := |+\rangle_1 \otimes \cdots \otimes |+\rangle_n = \frac{1}{\sqrt{N}} \sum_{x=0}^{N-1} |x\rangle$$

we obtain Grover's operator:

$$G = (\mathbb{1} - 2|+\rangle\langle+|)(2|x_0\rangle\langle x_0| - \mathbb{1})$$

This transformation is an inversion followed by a reflection (see Figure 5.2), and acts, in the plane spanned by $|+\rangle, |x_0\rangle$, as a rotation of angle

$$2\gamma = 2\arcsin\langle+|x_0\rangle = 2\arcsin\frac{1}{\sqrt{N}}$$

and, if we apply G for $r = \Theta(\sqrt{N})$ of times, starting from the state $|+\rangle$, we end up in a state which is very close to $|x_0\rangle$, thus a full measurement in the computational basis would give us x_0 with high probability. If we consider

a quantum Markov chain $(\mathcal{H}, \mathcal{G})$, where

$$\mathcal{G}(\rho) = G\rho G^\dagger$$

then Grover's result can be summarized in terms of hitting times as follows:

Theorem 5.1 *Grover's Markov chain $(\mathcal{H}, \mathcal{G})$ has a p-one shot hitting time for $|x_0\rangle$ at $\Theta(\sqrt{N})$, with $p = 1 - \mathcal{O}(\frac{1}{N})$, starting from the state $\rho_0 = |+\rangle\langle+|$.*

In other words, Grover's search algorithm is nothing more than a unitary walk.

5.3 Grover's chain and the concurrent hitting time

What happens if we use Grover's operator G with the concurrent hitting time paradigm, i.e. at each step we measure the hit and we stop when the state collapses to $|x_0\rangle$?

In other words, we want to compute the concurrent hitting time $h_{x_0}(|+\rangle\langle+|)$. The quantum operation representing the measurement we want to apply is:

$$\mathcal{T}_{x_0}(\rho) = \Pi_{x_0}\rho\Pi_{x_0} + (\mathbb{1} - \Pi_{x_0})\rho(\mathbb{1} - \Pi_{x_0})$$

i.e. the only coherences we destroy are only the ones between $|x_0\rangle$ and the rest of the space. Notice that such operation is efficiently computable (Figure 5.3). This adds a second call to the oracle of f at every step, but it only increases the query complexity by a constant factor.

Theorem 5.2 *Let $(\mathcal{H}, \mathcal{T} \circ \mathcal{G})$ be Grover's Markov chain under repeated measurements. The marked element x_0 is hit in expected $\Theta(N)$ steps.*

Before proving the result, we remark that this shows that the quadratic speedup is lost, even when we destroy the least amount of quantum coherences.

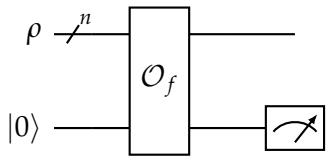


Figure 5.3: Circuit that implements $\mathcal{T}(\rho) = \Pi_{x_0}\rho\Pi_{x_0} + (\mathbb{1} - \Pi_{x_0})\rho(\mathbb{1} - \Pi_{x_0})$. We use an ancilla to store the value of $f(x)$ using the bit oracle $\mathcal{O}_f$. By measuring the ancilla, we destroy the coherence between the subspace of $|x_0\rangle$ and the rest. This qubit is then traced out.

Proof We start with two considerations: notice that we only care about the subspace $\mathcal{H}^* = \text{span}\{|x_0\rangle, |+\rangle\}$, since it is invariant with respect to G, which acts non-trivially only in this space. Moreover, G can be seen as a rotation of $2\gamma = 2\arcsin\frac{1}{\sqrt{N}}$.

Therefore, we define the chain $(\mathcal{H}^*, \mathcal{G})$ with a two-dimensional state space, where $\mathcal{G}$ applies a rotation of γ. If we take $|0^*\rangle = |+\rangle^{\otimes n}$ and $|1^*\rangle$ as the real state orthogonal to $|0^*\rangle$ in $\mathcal{H}^*$, then $|x_0\rangle$ can be written as:

$$|x_0\rangle = \frac{1}{\sqrt{N}}|0^*\rangle + \sqrt{\frac{N-1}{N}}|1^*\rangle$$

while the Grover operator G can be written as:

$$G = \begin{bmatrix} \cos 2\gamma & -\sin 2\gamma \\ \sin 2\gamma & \cos 2\gamma \end{bmatrix}$$

Thus, by Theorem 4.7, we need the trace of the matrix $(\mathcal{I} - \mathcal{G}_{-x_0})^{-1}(|0^*\rangle\langle 0^*|)$. Using a software of symbolic computation (see Appendix A.1), we obtain that such trace is:

$$h_{x_0}(\rho) = \frac{N}{4} + \frac{4}{N} - 1 = \Theta(N) \qquad \square$$

Notice that, instead of explicitly computing the matrix (which may be infeasible by hand), we could follow the same approach as in Section 4.4 i.e. (1) find the eigenvalues and eigenstates of $\Pi_{-x_0}G$, (2) use them to construct the eigenvalues and eigenvectors of $\mathcal{G}_{-x_0}$ and thus also of $(\mathcal{I} - \mathcal{G}_{-x_0})^{-1}$, then (3)

express the starting state $|0^*\rangle\langle 0^*|$ as a linear combination of these eigenstates and compute the trace by exploiting linearity. Also an interesting property of the hitting time is that, if the starting state and the destination are contained in a particular subspace of the Hilbert space, then it is sufficient to compute the eigenstates of the channel within this subspace.

This result tells us that the coherences between $|x_0\rangle$ and the rest of the Hilbert space are crucial for the speed-up of the algorithm.

5.4 Using measurement schemes in Grover's search

We found that measuring at each step does not give us considerable advantage. What happens if, at each step, we measure only with probability $p \in (0,1)$? Could we hope to recover Grover's quadratic speed-up? Let us use the results of Section 4.5 to analyze the following Grover's walk:

$$\mathcal{G}_p(\rho) = p\mathcal{T}_{x_0} \circ \mathcal{G}(\rho) + (1-p)\mathcal{G}(\rho) \tag{5.1}$$

i.e. we measure the hit of x_0 with probability p (and evolve unitarily otherwise). We stop the process if (1) the measurement occurs and (2) x_0 is measured.

Theorem 5.3 *For any constant probability $p \in (0,1)$, Grover's Markov chain $(\mathcal{H}, \mathcal{G}_p)$ with step-wise measurement probability p measures x_0 in expected $\Theta(N)$ steps. However, if $p = \Theta(\frac{1}{\sqrt{N}})$, then the marked state is hit in expected $\Theta(\sqrt{N})$ steps.*

Proof We essentially need to compute $h^{\sigma}_{x_0}(|0^*\rangle\langle 0^*|)$ in the walk defined in the proof of Theorem 5.2, where $\sigma = Geom(p)$.

Using Theorem 4.10 we obtain that:

$$h^{\sigma}_{x_0}(|0^*\rangle\langle 0^*|) = \frac{1}{p} \operatorname{Tr}\left[(\mathcal{I} - \mathcal{G}^{\sigma}_{-x_0})^{-1}(|0^*\rangle\langle 0^*|)\right]$$

where $\mathcal{G}^{\sigma}_{-x_0}(\rho)$ can be computed using Lemma 4.11. Since we are talking about a two-dimensional Hilbert space, all these maps can be represented by 4×4 matrices, which allows us to directly compute the 2×2 matrix $(\mathcal{I} - \mathcal{G}^{\sigma}_{x_0})^{-1}(|0^*\rangle\langle 0^*|)$, and its trace turns out to be (Appendix A.1):

$$\mathrm{Tr}\left[(\mathcal{I} - \mathcal{G}^{\sigma}_{x_0})^{-1}(|0^*\rangle\langle 0^*|)\right] = \frac{N^2p^2 + 16N - 20Np + 16p}{8N - 4Np}$$

Notice that, if we plug $p = 1$, we get exactly what we found in Theorem 5.2, and this extends to any constant p, where this expectation is still $\Theta(N)$. If $p = o(1)$, the expected hitting time is:

$$h^{\sigma}_{x_0}(\rho) = \frac{1}{p}\frac{N^2p^2 + 16N - 20Np + 16p}{8N - 4Np} \sim \frac{Np}{8} + \frac{2}{p} \tag{5.2}$$

If $p = \Theta(\frac{1}{\sqrt{N}})$ then one can see that the hitting time is $\Theta(\sqrt{N})$ as claimed.□

This result shows how much the speed-up given by Grover's algorithm turns out to be fragile under such imperfect unitary transformations. However, a carefully chosen probability p yields a new variant of Grover's search algorithm, recovering the original complexity.

5.5 Experiment on Qiskit

We now experimentally validate the theory by running the quantum circuit defined in the previous sections with the Qiskit[1] [3] QASM simulator. We consider n qubits, with $N = 2^n$, and we fix $x_0 = N - 1$, i.e. $|x_0\rangle = |11\cdots1\rangle$. With this choice the circuit is simpler, as Grover's reflection operator $\mathbb{1} - 2|x_0\rangle\langle x_0|$ can be implemented using a controlled Z gate. We want to implement the map $\mathcal{G}_p$ as given in Eq. (5.1), and in order to do so, we will use two ancilla qubits: the first one M is used for the measurement of the hit, as already shown in Figure 5.3, while the second one R is needed to decide, with probability p, whether we will measure or not. This

[1] https://qiskit.org

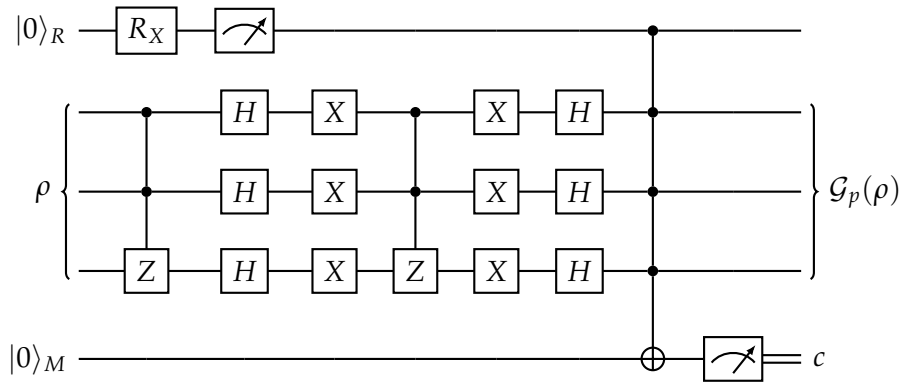


Figure 5.4: Full implementation of $\mathcal{G}_p$ for $n = 3$ qubits. One can see that, if $|x_0\rangle = |11\cdots1\rangle$, the CCZ gate implements the reflection $\mathbb{1} - 2|x_0\rangle\langle x_0|$. The second reflection $\mathbb{1} - 2|+\rangle\langle+|$ is achieved by the fact that $HX|1\rangle\langle1|XH = |+\rangle\langle+|$. The qubit R is rotated by an angle $\theta = \arcsin\sqrt{p}$ and then measured, with the effect that it is in a classical mixture of state $|1\rangle$ (with probability p) and $|0\rangle$ (with probability $1-p$), the $CNOT$ gate can then be seen as classically controlled by this random bit. At the end, c will be 1 if and only if during this step (1) the measurement occurs, and (2) the marked state x_0 is hit.

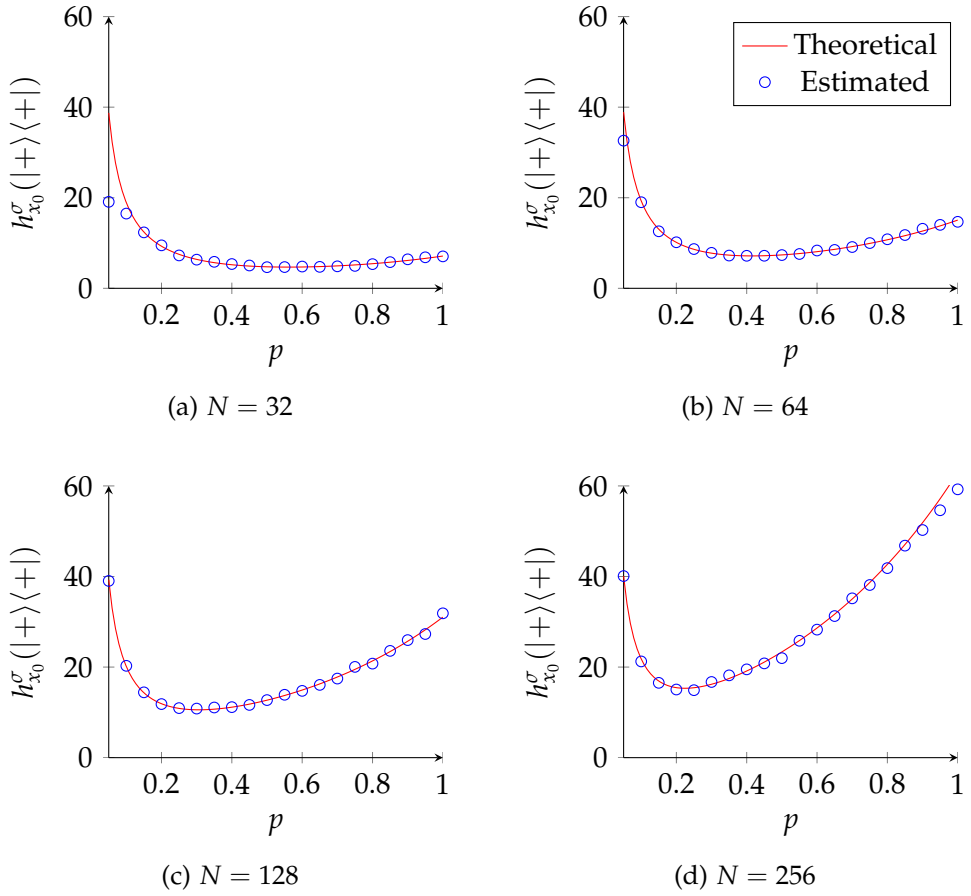
is achieved by rotating the qubit, initially in state $|0\rangle$, using $R_X(\arcsin(\sqrt{p}))$. The random bit is then obtained by measuring this bit in the computational basis.

$$|0\rangle_R \overset{R_X}{\mapsto} \sqrt{1-p}|0\rangle_R + \sqrt{p}|1\rangle_R \mapsto (1-p)|0\rangle\langle0|_R + p|1\rangle\langle1|_R$$

At this point, we can use the bit oracle as in Figure 5.3 controlled by the qubit R: if R is in state $|0\rangle$, nothing happens and the n qubits remain unentangled from M, thus any measurement on M will return 0 and will not change the state of the main register. The circuit implementing the step of the walk is summarised in Figure 5.4. One could also use Python to generate random bits which classically control whether we add the measurement circuit at each step or not. The choice to implement the random choices using a qubit turned out to give much faster computations, as the former approach requires to modify (and thus recompile) the whole quantum circuit for each iteration of the Monte Carlo estimation. Also, Qiskit currently does not

allow to stop the circuit execution as soon as we measure 1. Therefore, we carried out the experiment as follows: we run the above circuit $2N$ times, we store for each run a string of the $2N$ bits measured throughout the walk, and compute the concurrent hitting time as the position of the first 1 occurring in this string. We also estimate the probability to go over $2N$ steps without hitting x_0, by counting the strings of all zeroes. During these experiments this turned out to be very rare, giving $< 0.1\%$ probability only in extreme cases, which we can neglect in the estimation of our expectations.

As shown in Figure 5.5, experiments are consistent with the results of Theorem 5.3.



(a) $N = 32$ (b) $N = 64$

(c) $N = 128$ (d) $N = 256$

Figure 5.5: Plots of the concurrent hitting time with respect to the step-wise measurement probability p. The red line is the theoretical expectation as computed in the proof of Theorem 5.3, while the blue circles are the estimations computed by running the circuit with the given p for 1000 independent times.

Chapter 6

Quantum Theory of Ergodic Markov Chains

Aside from algorithms based on hitting times, where one looks for a particular state of the chain, Markov chains are also exploited in sampling algorithms: in the classical theory, the state probability vector of Markov chains under particular assumptions – namely *irreducibility* and *aperiodicity* – approaches a limiting (or *stationary*) distribution starting from any initial condition.

$$\lim_{t\to\infty} qP^t = \pi \quad \text{for every } q$$

Using carefully designed chains, the limiting distribution will be some distribution of interest (e.g. a uniform distribution over the elements of S), and by running the chain for a sufficiently large amount of steps, the probability distribution of the reached state is close to the stationary distribution. A notable application of this method is uniform sampling of a perfect matching from a particular class of graphs in polynomial time [6, 10].

In this chapter we explore an extension of the notions of irreducibility and aperiodicity for quantum Markov chains, and we prove that these give a

sufficient and necessary condition for the existence of the limiting state.

$$\lim_{t\to\infty} \mathcal{E}^t(\rho) = \rho^* \tag{6.1}$$

These proofs will closely follow the ones that can be found in the classical theory, and the results trivially imply their classical counterparts. In the rest of the chapter, we will use d to denote the dimension of the Hilbert space of a Markov chain $(\mathcal{H}, \mathcal{E})$.

6.1 Irreducibility

We start by discussing the first of the two aforementioned notions: *irreducibility*.

A classical Markov chain is said to be irreducible if, starting from any state $s \in S$, it is possible to eventually reach any other state $s' \in S$ with positive probability. In other words, the graph associated with the Markov chain is strongly connected.

In quantum theory, we cannot give a definition of irreducibility that 'covers' only the elements of a particular basis, because the existence of the limit as given in Eq. (6.1) does not vary under a change of basis U of the states:

$$\mathcal{U}\left[\lim_{t\to\infty} \mathcal{E}^t(\rho)\right] = \lim_{t\to\infty} (\mathcal{U}\circ\mathcal{E}\circ\mathcal{U}^{-1})^t(U\rho U^\dagger)$$

where $\mathcal{U}(\rho) = U\rho U^\dagger$. On the other hand, we have the following example:

Example 6.1 *Consider a unitary walk $(\mathcal{H},\mathcal{U})$ with $\mathcal{U}(\rho) = U\rho U^\dagger$ on a Hilbert space such that:*

$$U|x\rangle = |x+1 \bmod d\rangle$$

Starting from a state of the computational basis, we eventually reach any other state of the same basis. However, if we start from the state $|+\rangle$, we have that this is an eigenstate of U, and no other state of the Hadamard basis can be reached.

We may say that a state $|\psi\rangle$ can be 'reached' if, at some point, we are in a state ρ with a non-zero overlap with $|\psi\rangle$, i.e. a state is reachable if it is measurable from a state of the Markov process in *some* measurement basis. Hence, we are looking for definitions that do not depend on a particular measurement basis. As noted in [27, p. 102], this suggests that irreducibility does not only imply spectral properties of the map, but it is also determined by such properties.

Definition 6.2 ([8]) *A positive map $\mathcal{E} : \bar{\mathcal{H}} \to \bar{\mathcal{H}}$ is said to be* irreducible *if the following holds for any $\rho \in \bar{\mathcal{H}}$:*

$$\bigoplus_{k=0}^{\infty} \operatorname{supp}\left(\mathcal{E}^k(\rho)\right) = \mathcal{H}$$

And we say that a quantum Markov chain $(\mathcal{H}, \mathcal{E})$ is irreducible if $\mathcal{E}$ is. More informally, this notion of irreducibility states that, by evolving the chain for a sufficient amount of steps, we will see states that span the whole Hilbert space.

Theorem 6.3 ([27, 8]) *The following are equivalent for a positive map $\mathcal{E} : \bar{\mathcal{H}} \to \bar{\mathcal{H}}$*

(1) $\mathcal{E}$ is irreducible;

(2) For any projector $\Pi \in \bar{\mathcal{H}}$, $\mathcal{E}(\bar{\mathcal{H}}_\Pi) \subseteq \bar{\mathcal{H}}_\Pi$ implies $\Pi \in \{0, \mathbb{1}\}$;

(3) $(\mathcal{I} + \mathcal{E})^{d-1}(\rho)$ is positive definite for any positive semi-definite $\rho \in \bar{\mathcal{H}} \setminus \{0\}$.

where $\bar{\mathcal{H}}_\Pi$ is the subspace of $\bar{\mathcal{H}}$ onto which Π projects, and $\mathcal{I}$ is the identity super-operator in $\bar{\mathcal{H}}$.

Proof (1) $\Rightarrow$ (2). Suppose there is a projector $\Pi \neq 0, \mathbb{1}$ such that $\mathcal{E}(\bar{\mathcal{H}}_\Pi) \subseteq \bar{\mathcal{H}}_\Pi$. This implies that $\mathcal{E}(\rho) \in \bar{\mathcal{H}}_\Pi$ whenever $\rho \in \bar{\mathcal{H}}_\Pi$, and this also holds for $\mathcal{E}^k(\rho)$ for every k, implying:

$$\bigoplus_{k=0}^{\infty} \operatorname{supp}\left(\mathcal{E}^k(\rho)\right) \subseteq \mathcal{H}_\Pi \subsetneq \mathcal{H}$$

(2) $\Rightarrow$ (3). Since $\rho, \mathcal{E}(\rho)$ are always positive semi-definite it holds that

$$\ker((\mathcal{I}+\mathcal{E})(\rho)) = \ker(\rho + \mathcal{E}(\rho)) \subseteq \ker\rho$$

Now suppose that this claim holds with equality: this would imply that $\mathrm{supp}\,(\mathcal{E}(\rho)) \subseteq \mathrm{supp}\,(\rho)$ and, in general, $\mathcal{E}(\bar{\mathcal{H}}_\Pi) \subseteq \bar{\mathcal{H}}_\Pi$[1], where Π is the projector onto $\mathrm{supp}\,(\rho)$. By (2), this gives $\Pi = \mathbb{1}$ and $\mathrm{supp}\,(\rho) = \mathcal{H}$, which means ρ is already positive definite and

$$\ker((\mathcal{I}+\mathcal{E})^{d-1}(\rho)) = \ker\rho = \{0\}$$

giving that $(\mathcal{I}+\mathcal{E})^{d-1}(\rho)$ is positive definite as well. Therefore we now assume $\ker((\mathcal{I}+\mathcal{E})(\rho)) \subsetneq \ker\rho$, and this implies that the rank of the operator strictly increases at each step, reaching its maximum in at most $d-1$ steps.

(3) $\Rightarrow$ (1). If there is a proper subspace $\bar{\mathcal{H}}_\Pi$ of $\bar{\mathcal{H}}$ and a state $\rho \in \bar{\mathcal{H}}$ such that $\bigoplus_{k=0}^{\infty} \mathcal{E}^k(\rho) \subseteq \bar{\mathcal{H}}_\Pi$, then in particular we must have:

$$\mathrm{supp}\,((\mathcal{I}+\mathcal{E})(\rho)) \subseteq \mathrm{supp}\,(\rho + \mathcal{E}(\rho)) \subseteq \mathrm{supp}\,(\rho) \oplus \mathrm{supp}\,(\mathcal{E}(\rho)) \subseteq \mathcal{H}_\Pi$$

By inductively applying the argument we obtain

$$\mathrm{supp}\left((\mathcal{I}+\mathcal{E})^k(\rho)\right) \subseteq \mathcal{H}_\Pi$$

i.e. this operator is not positive definite for any $k \geq 0$, and this holds in particular for $k = d-1$. □

Here we also show that, using the result above, the notion of irreducibility can be weakened, meaning that an irreducible chain will certainly span the whole subspace within $d-1$ steps:

Corollary 6.4 *A positive map $\mathcal{E}$ is irreducible if and only if:*

$$\bigoplus_{k=0}^{d-1} \mathrm{supp}\left(\mathcal{E}^k(\rho)\right) = \mathcal{H}$$

[1] This because $\mathrm{supp}\,(\sigma) \subseteq \mathrm{supp}\,(\rho)$ implies $\mathrm{supp}\,(\mathcal{E}(\sigma)) \subseteq \mathrm{supp}\,(\mathcal{E}(\rho))$.

Proof Since $(\mathcal{I}+\mathcal{E})^{d-1}(\rho)$ is positive definite, we have:

$$\begin{aligned}\mathcal{H} &= \operatorname{supp}\left((\mathcal{I}+\mathcal{E})^{d-1}(\rho)\right) \\ &= \operatorname{supp}\left(\left(\sum_{k=0}^{d-1}\binom{d-1}{k}\mathcal{E}^k\right)(\rho)\right) \\ &= \bigoplus_{k=0}^{d-1}\operatorname{supp}\left(\mathcal{E}^k(\rho)\right)\end{aligned}$$

The converse implication is trivial. □

We now give a proof of a quantum version of the Perron-Frobenius theorem, which states that any irreducible Markov chain admits a unique stationary state.

Theorem 6.5 ([27, 8]) *Let $\mathcal{E}$ be an irreducible positive trace-preserving map. Then its eigenvalue $\lambda = 1$ is non-degenerate and admits a unique positive definite eigenstate ρ^*.*

Before proving the statement, notice that positive definiteness here means exactly that *any* state in the Hilbert space has a strictly positive probability to appear upon measurement under a suitably chosen basis.

Proof Let ρ^* be an eigenstate of $\mathcal{E}$ associated with $\lambda = 1$. By Theorem 6.3, we know that

$$0 \prec (\mathcal{I}+\mathcal{E})^{d-1}(\rho^*) = 2^{d-1}\rho^*$$

i.e. ρ^* is positive definite. Suppose now for a contradiction there is a linearly independent eigenstate σ associated to $\lambda = 1$ (we can assume without loss of generality that σ is Hermitian, otherwise we can take $\sigma + \sigma^\dagger$): this means that also $\rho^* + c\sigma$ is an eigenstate associated to $\lambda = 1$: by the same argument as above we deduce that this eigenstate is positive definite for any c, but we can choose a $c \neq 0$ such that $\rho^* + c\sigma \succeq 0$ has a non-trivial kernel. □

This state ρ^*, which satisfies $\mathcal{E}(\rho^*) = \rho^*$, is called *stationary state* of the chain, and notice that this is the only possible candidate state for the limiting state of a Markov chain, as $\lim_{t\to\infty} \mathcal{E}^t(\rho^*) = \rho^*$ trivially.

6.2 Periodicity

Notice that irreducibility alone is not sufficient to conclude that the chain dynamics will always converge to its stationary state. In order to prove this we state the following (classical) example:

Example 6.6 *Let $(\mathcal{H}, \mathcal{E})$ be a quantum Markov chain with $|\mathcal{H}| = 2$ and:*

$$\mathcal{E}(\rho) = |1\rangle\langle 0|\rho|0\rangle\langle 1| + |0\rangle\langle 1|\rho|1\rangle\langle 0|$$

One can see that this is irreducible, as any pure state (the only ones without full rank) will oscillate between $|0\rangle$ and $|1\rangle$, which span the whole space. In fact, the unique stationary distribution is

$$\rho^* = \frac{1}{2}|0\rangle\langle 0| + \frac{1}{2}|1\rangle\langle 1|$$

However, if we start from an element of the computational basis, the limit does not exist.

The problem here is that the chain has some 'periodicity', i.e. the evolution of the chain can be divided in steps where the chain is certainly in state $|0\rangle$ and steps in which the chain is certainly in state $|1\rangle$.

Definition 6.7 ([[8]) *] Let $\mathcal{E} : \bar{\mathcal{H}} \to \bar{\mathcal{H}}$ be a positive map. The* period *of ρ is defined as*

$$d(\rho) := \gcd\left\{k \geq 1 \,\middle|\, \operatorname{supp}\left(\mathcal{E}^k(\rho)\right) \supseteq \operatorname{supp}(\rho)\right\}$$

where gcd *denotes the greatest common divisor. If $d(\rho) = 1$, the state ρ is said to be* aperiodic. *The map $\mathcal{E}$ is also said to be aperiodic if every state $\rho \in \bar{\mathcal{H}}$ is.*

Unlike irreducibility, it is not straightforward to see that this notion of period extends the one defined in classical theory: the period of a state s in a classical Markov chain is given by the greatest common divisor of the lengths of the cycles in the graph containing s, and a cycle of length ℓ exists if and only if one can return to s from s with exactly ℓ steps with positive probability. On the other hand, if we start from a pure state $|\psi\rangle\langle\psi|$ it is not sufficient to have a non-zero overlap with the state to say that we 'returned' to that state according to Definition 6.7. Instead, we say that we return to a state ρ in ℓ steps if any state observable in ρ is also observable in $\mathcal{E}^\ell(\rho)$ (under a suitably chosen basis).

Theorem 6.8 ([8]) *Let $\mathcal{E}$ be an irreducible map. The following are equivalent:*

(1) $\mathcal{E}$ is aperiodic;

(2) For every state $\rho \in \bar{\mathcal{H}}$, there exists $M \equiv M(\rho)$ such that $\mathcal{E}^m(\rho)$ is positive definite for every $m \geq M$.

Proof (2) $\Rightarrow$ (1). It is sufficient to see that $1 = \gcd(M, M+1)$ divides $d(\rho)$ for every ρ.

(1) $\Rightarrow$ (2). Consider the set

$$C = \left\{ k \geq 1 \,\middle|\, \operatorname{supp}\left(\mathcal{E}^k(\rho)\right) \supseteq \operatorname{supp}(\rho) \right\}$$

Notice that, whenever $x, y \in C$ we have $\mathcal{E}^{x+y}(\rho) = \mathcal{E}^x(\mathcal{E}^y(\rho))$ and

$$\operatorname{supp}\left(\mathcal{E}^{x+y}(\rho)\right) = \mathcal{E}^x(\operatorname{supp}(\mathcal{E}^y(\rho))) \supseteq \mathcal{E}^x(\operatorname{supp}(\rho)) \supseteq \operatorname{supp}(\rho)$$

i.e. $x + y \in C$. From number theory we know that C contains all but a finite number of elements. Thus let us take M' as the minimum integer after which every number is included in C. If $M = M' + d - 1$, for any $m \geq M$ we have $\operatorname{supp}\left(\mathcal{E}^{m-i}(\rho)\right) \supseteq \operatorname{supp}(\rho)$ for every $0 < i < d$. This implies

$$\operatorname{supp}(\mathcal{E}^m(\rho)) = \mathcal{E}^i\left(\operatorname{supp}\left(\mathcal{E}^{m-i}(\rho)\right)\right) \supseteq \mathcal{E}^i(\operatorname{supp}(\rho)) = \operatorname{supp}\left(\mathcal{E}^i(\rho)\right)$$

In particular, supp $(\mathcal{E}^m(\rho))$ contains the sum of these subspaces, which equals $\mathcal{H}$ by Corollary 6.4. □

6.3 Limiting behaviour

In the previous sections we gave the definitions of irreducibility and aperiodicity, proving that these together give a necessary and sufficient condition for the following definition:

Definition 6.9 *A positive map $\mathcal{E}$ is called* primitive *if there exists some M such that $\mathcal{E}^M(\rho)$ is positive definite for every $\rho \in \bar{\mathcal{H}}$.*

Positivity of the map ensures that positive definiteness is preserved by $\mathcal{E}$, thus the condition given by Definition 6.9 holds also for every $m \geq M$. A Markov chain with a primitive map is also said to be *ergodic*.

Theorem 6.10 (Ergodic theorem of quantum Markov chains [27]) *Let $\mathcal{E}$ be a positive trace-preserving map. The following are equivalent.*

(1) $\mathcal{E}$ is irreducible and aperiodic;

(2) There exists $n \in \mathbb{N}$ such that $\mathcal{E}^n(\rho)$ is positive definite for every $\rho \in \bar{\mathcal{H}}$, i.e. $\mathcal{E}$ is primitive;

(3) The only eigenvalue λ of $\mathcal{E}$ with $|\lambda| = 1$ is $\lambda = 1$, which is regular and has a unique positive definite corresponding eigenoperator ρ^*;

(4) For any initial state ρ we have

$$\lim_{t\to\infty} \mathcal{E}^t(\rho) = \rho^*$$

Proof We already know the equivalence of (1) and (2) from Theorem 6.8.

(4) ⇒ (2). We know from Theorem 6.5 that the limiting distribution ρ^* (which must necessarily be an eigenstate associated to the eigenvalue 1) is positive definite. Since the limit is positive definite, there must exist some

integer T such that $\mathcal{E}^t(\rho)$ is positive definite for every $t \geq T$. Taking the maximum of T over any choice of ρ gives an integer satisfying (2).

(3) $\Rightarrow$ (4). Let $\mathcal{E}_\phi$ be the map obtained from $\mathcal{E}$ by zeroing out every eigenvalue other than $\lambda = 1$. This means that:

$$\mathcal{E}_\phi(\rho) = \rho^* \operatorname{Tr} \rho$$

If $\lambda = 1$ is the only eigenvalue with $|\lambda| = 1$, we have that:

$$\rho^* = \mathcal{E}_\phi(\rho) = \lim_{t \to \infty} \mathcal{E}^t(\rho)$$

as every other eigenvalue tends to zero in the limit.

(2) $\Rightarrow$ (3). If there is another $e^{i\theta} \neq 1$ with magnitude one in $\mathcal{E}$, then we can select a sequence of integers $\{t_i\}_i$ such that $e^{i\pi t_i}$ becomes arbitrarily close to 1. This means that there are two independent eigenstates ρ and σ with eigenvalue 1 for the limiting state. By choosing $c \neq 0$ such that $\rho + c\sigma$ has a kernel, this disproves condition (2) of Theorem 6.8, as the fact that $\mathcal{E}^m(\rho)$ is positive definite for every $m \geq M$ means that the same must hold for the limit as $m \to \infty$. □

6.4 Application: mixing the cycle

Now we try to apply what we developed in this chapter, also giving some observations that can be used in practice to design ergodic chains. Consider a graph with a cycle of N elements, which we conveniently label with elements of $\mathbb{Z}_N$. For simplicity we consider an odd N (the reason for this choice will be explained later). As a first result, we notice that using purely unitary dynamics never leads to ergodicity.

Theorem 6.11 *Unitary walks on Hilbert spaces of dimensions at least* 2 *are never ergodic.*

Proof Considering a unitary U, it always has d eigenstates $|\phi_1\rangle, \ldots, |\phi_d\rangle$ with associated unitary eigenvalues. This means that the map $\mathcal{U}(\rho) = U\rho U^\dagger$ has d eigenstates $|\phi_1\rangle\langle\phi_1|, \ldots, |\phi_d\rangle\langle\phi_d|$ associated to the eigenvalue $\lambda = 1$, contradicting the necessary condition for irreducibility as given in Theorem 6.5. □

More intuitively, if we start from any pure eigenstate of the unitary, the dynamical evolution will never be able to access other ones, and we have at least two such eigenstates in any case. Thus, the limiting distribution actually *depends* on initial distribution, whereas ergodicity also implies that the chain 'forgets' its initial conditions on the long run.

Another interesting question is whether any amount of coherent dynamics is sufficient to prevent ergodicity: for example, what happens if we mix a unitary walk with an ergodic chain, i.e. we run a step of the unitary walk with some probability and a step of an ergodic chain otherwise? The following result confirms that ergodicity is actually always preserved.

Theorem 6.12 *Let $(\mathcal{H}, \mathcal{E})$ be an ergodic Markov chain and $(\mathcal{H}, \mathcal{G})$ be a (non necessarily ergodic) chain. Then the chain $(\mathcal{H}, (1-p)\mathcal{G} + p\mathcal{E})$ is ergodic for any $p > 0$.*

Proof Let $\mathcal{E}' = (1-p)\mathcal{G} + p\mathcal{E}$. We show that the new chain is irreducible and aperiodic. Irreducibility is implied by the fact that:

$$\operatorname{supp}\left(\mathcal{E}'(\rho)\right) \supseteq \operatorname{supp}\left(\mathcal{E}(\rho)\right)$$

which implies the same for $\mathcal{E}^k$ for every k, by an inductive argument. Thus we have

$$\bigoplus_{k=0}^{d-1} \operatorname{supp}\left(\mathcal{E}'^k(\rho)\right) \supseteq \bigoplus_{k=0}^{d-1} \operatorname{supp}\left(\mathcal{E}^k(\rho)\right) = \mathcal{H}$$

Now we address aperiodicity: fixed a starting state ρ, define the following set

$$C = \left\{ k \geq 1 : \text{supp}\left(\mathcal{E}^k(\rho)\right) \supseteq \text{supp}\left(\rho\right) \right\}$$

and C' analogously for $\mathcal{E}'$. From the fact that

$$\text{supp}\left(\mathcal{E}'^k(\rho)\right) \supseteq \text{supp}\left(\mathcal{E}^k(\rho)\right)$$

as proven in the first part of the argument we deduce that $C' \supseteq C$, and thus the greatest common divisor over C' cannot be greater than the one over C, which is 1. □

This theorem in particular implies that we can 'decorate' any chain we want with unitary dynamics, and ergodicity will be preserved. It may be possible to use this result to construct ergodic chains that approach their limiting state *more efficiently*, i.e. the distance (with respect to some fixed metric) between our state and the limit becomes negligible in less steps. A conjecture is that coherent dynamics can be useful for a speed-up in this sense [25].

We now go back to our example. We consider a coined quantum walk on the cycle, i.e. a Hilbert space $\mathcal{H}_C \otimes \mathcal{H}$ where $\mathcal{H}_C = \text{span}\{|\uparrow\rangle, |\downarrow\rangle\}$ denotes a qubit register containing a 'coin', which indicates the direction we want to take in the walk. A well-known approach to achieve a unitary walk in this case is the following [12, 1]: we first flip the coin register using an operator (e.g. the Hadamard gate H), and then we apply the following unitary (shift operator):

$$S|\uparrow\rangle|x\rangle = |x + 1 \bmod N\rangle$$
$$S|\downarrow\rangle|x\rangle = |x - 1 \bmod N\rangle$$

Also, although obvious, it is interesting to remark that the quantum construction as shown in Section 3.4 of a classical Markov chain that satisfies the

classical definitions of irreducibility and aperiodicity yields an irreducible and aperiodic quantum Markov chain. This is also useful because classical irreducibility and aperiodicity are easier to prove: one only needs to see that the graph associated to the Markov chain is connected for irreducibility, while aperiodicity can be proven by exhibiting, for each state, a pair of closed walks of co-prime length within the graph, containing the state.

Speaking about the cycle, let $U = S \cdot (H_C \otimes \mathbb{1})$. If $\mathcal{U}(\rho) = U\rho U^\dagger$ and $\mathcal{T}(\rho)$ is a channel implementing a full measurement with respect to the computational basis (for both coin and main registers), one can see that the chain $(\mathcal{H}_C \otimes \mathcal{H}, \mathcal{T} \circ \mathcal{U})$ is a classical Markov chain, where any state of the form $|\uparrow\rangle|x\rangle$ or $|\downarrow\rangle|x\rangle$ becomes one of $|\uparrow\rangle|x+1\rangle$ or $|\downarrow\rangle|x-1\rangle$ with equal probability, regardless of the initial value in the coin register. One may argue that we may start from a particular initial state, which is highly non-classical, i.e. it has superpositions in such a way that the application of $\mathcal{U}$ in the first step will give non-classical behaviour that diverges from the behaviour of the classical Markov chain. This, however, can only happen in the first step, before the first application of $\mathcal{T}$ makes the state collapse to a completely classical state. Therefore we can assume without loss of generality that we already start from a classical state.

Claim 6.13 *The chain $(\mathcal{H}_C \otimes \mathcal{H}, \mathcal{T} \circ \mathcal{U})$ is irreducible and aperiodic.*

Proof The chain is irreducible since it is always possible to reach a state $|\downarrow\rangle|y\rangle$ from any other state: it is sufficient to reach $|y\rangle$ from $|y+1\rangle$ in order to obtain the $|\downarrow\rangle$ on the coin register, and the same holds for the states with $|\uparrow\rangle$.

For aperiodicity we need to use the fact that the length N of the cycle is odd: in this way, starting from $|\uparrow\rangle|x\rangle$, among all the possible walks we have two

paths: the first one is simply doing

$$|\uparrow\rangle|x\rangle \to |\downarrow\rangle|x-1\rangle \to |\uparrow\rangle|x\rangle$$

which has length 2. The second alternative is to visit the whole cycle

$$|\uparrow\rangle|x\rangle \to |\uparrow\rangle|x+1\rangle \to \ldots \to |\uparrow\rangle|x-1\rangle \to |\uparrow\rangle|x\rangle$$

which has length N. Therefore, the period of the state $|\uparrow\rangle|x\rangle$, given by the greatest common divisor among the possible closed walks, has to divide $\gcd(2, N) = 1$, i.e. $|\uparrow\rangle|x\rangle$ is an aperiodic state for every $|x\rangle$. The analogous reasoning applies to states of the form $|\downarrow\rangle|x\rangle$ □

This argument, along with the results proved in the rest of the chapter, tells us that this Markov chain approaches a limiting distribution. If we measure only with a certain probability, i.e. we consider the chain $(\mathcal{H}_C \otimes \mathcal{H}, \mathcal{E})$ with:

$$\mathcal{E} = (1-p)\mathcal{U} + p\mathcal{T} \circ \mathcal{U}$$

then Theorem 6.12 tells us that ergodicity is preserved, and a stationary state is reached in the limit also in this case. Moreover, for this particular case one can see that $\mathcal{E}(\mathbb{1}) = \mathbb{1}$, i.e. the fully mixed state $\frac{\mathbb{1}}{N}$ is an eigenstate associated to the eigenvalue 1, which must be unique by irreducibility, and so it must also be the limiting state, i.e. for any starting state ρ:

$$\lim_{t\to\infty} \mathcal{E}^t(\rho) = \frac{\mathbb{1}}{N}$$

6.5 Irreducibility and hitting times

We close this section by showing another important property for irreducible Markov chains. Since irreducibility implies that basically any two states can be reached from one another, this intuitively implies that hitting times are always finite under this assumption, i.e. we will always hit any given state sooner or later, as the chain goes on. This turns out to be the case.

Theorem 6.14 *Let $(\mathcal{H}, \mathcal{E})$ be an irreducible quantum Markov chain. Then, for any subspace $\mathcal{S} \subseteq \mathcal{H}$ and any starting state $\rho \in \bar{\mathcal{H}}$, the hitting time $h_{\mathcal{S}}(\rho)$ exists and is finite.*

Proof We know from Theorem 4.7 that $h_{\mathcal{S}}$ is finite if the map

$$\rho \mapsto \Pi_{-\mathcal{S}} \mathcal{E}(\rho) \Pi_{-\mathcal{S}} =: \mathcal{E}_{-\mathcal{S}}(\rho)$$

has all eigenvalues λ with $|\lambda| < 1$, where $\Pi_{-\mathcal{S}} = \mathbb{1} - \Pi_{\mathcal{S}}$ is the projector onto the complement space of $\mathcal{S}$. We already know the absolute value of the eigenvalues cannot exceed 1, therefore suppose for a contradiction we have some eigenstate ρ such that:

$$\mathcal{E}_{-\mathcal{S}}(\rho) = e^{i\theta} \rho$$

for some θ. By Theorem 6.5, we know that the only eigenvalue with absolute value 1 can be 1 itself, i.e. $e^{i\theta} = 1$ necessarily. This implies that:

$$\mathrm{Tr}[\Pi_{-\mathcal{S}} \mathcal{E}(\rho)] = \mathrm{Tr}\left[\mathcal{E}_{-\mathcal{S}}(\rho)\right] = \mathrm{Tr}\, \rho$$

and, in particular, that $\Pi_{-\mathcal{S}} \mathcal{E}(\rho) \Pi_{-\mathcal{S}} = \mathcal{E}(\rho)$. This gives that $\mathcal{E}(\rho) = \rho$ as well. Moreover, ρ must be positive definite, by irreducibility. This contradicts the fact that $\mathrm{Tr}[\Pi_{\mathcal{S}} \rho] = \mathrm{Tr}[\Pi_{\mathcal{S}} \mathcal{E}(\rho)] = 0$. □

This result marks a clear connection between the notion of irreducibility and the hitting times defined in Chapter 4: in particular, this result implies that walks such as the one of Section 4.7 (along with measurements), are not irreducible. In general, a quantum Markov chain partitions the Hilbert space into subspaces called *bottom strongly connected components*, which are basically the equivalence classes of the relation of reachability that we informally gave at the beginning of this chapter. An irreducible chain gives only the trivial partition consisting of one component containing the whole space (since any two states are reachable from one another). Readers that are interested in this induced partitioning can find more information in [8].

Related Work

In the past 30 years, a lot of research was made on quantum walks, starting from the seminal works by Kempe [12, 23]. Kempe defined these walks based on unitary evolutions, along with the definitions of hitting times we further analyzed in this work: Kempe proved that, on the hypercube, one can use a particular coined quantum walk to reach a vertex from its opposite end in linear time, which was argued to be exponentially faster than any classical walk on the same structure [13].

Kempe's quantum walk formalism was used to solve a bunch of problems: a notable example is the *k-collision* problem, where one has to find a subset of k equal items out of N values. Ambainis [2] showed a $\mathcal{O}(N^{k/(k+1)})$ quantum algorithm for this problem, exploiting a unitary walk. Childs and Eisenberg [5] showed that Ambainis' algorithm could be used to solve the more general problem of *k-subset finding*, where one has to find a subset of size k satisfying some given, arbitrary property.

Other interesting formalizations of quantum walk is Szegedy's construction [25], which transforms an arbitrary classical Markov chain into a (bipartite) unitary walk, and the TOM/TEM formalism by Gudder [9], which tries to generalize the concept of Markov chain using a stochastic matrix of quantum channels. For the latter model, Lardizabal [16] gave and applied a

definition of hitting time. Neither of the mentioned formalizations, however, is as expressive as the model we introduced in this work.

Always on hitting times, a notable result is the one by Magniez et al. [17]: using an approach similar to ours for estimating hitting times, they found out that, for a restricted class of classical Markov chains, it is possible to construct a unitary walk using Szegedy's method achieving hitting times with quadratic speed-up, and also showed a general construction for searching algorithms, based on Grover-like rotations.

While all the aforementioned works only treat coherent dynamics, some research was made also towards a model that takes into account randomness and decoherence: the Quantum Stochastic Walk model [26], which can be seen as the continuous-time counterpart of our formalization, unifies both classical Markov chains and unitary walks under one single model. Numerical experiments from Kendon and Tregenna [14, 15] also suggested that decoherence could even be beneficial for the speed-up of some walks.

A more comprehensive overview showing the latest results on quantum walks can be found in [11].

Conclusions and Future Work

In this thesis we presented a different and more general way to formalize quantum random walks, including both coherent dynamics and classical randomness. Within this framework, the notions of one-shot and concurrent hitting time given by Kempe [12, 13] for unitary walks can be naturally extended. Moreover, we found that the concept of concurrent hitting time is the natural generalization of the notion well-understood in the classical theory [19]. Other aspects of Markov chains, such as the ergodic theory, can find their quantum counterpart with this framework.

One possible future direction is to see, using this framework, how much quantum coherences of the states throughout the evolution of the walk influence the hitting times: what is the difference between a measuring the hit and a full step-wise measurement? An answer to this question also gives clarifications on how much speed-up can be achieved using the notion of hitting time, since the evolution of a quantum walk with full step-wise measurement could be reproduced classically.

Since we talked about necessary and sufficient conditions for the convergence to a stationary distribution, another very interesting question is about quantifying *how fast* a Markov chain converges. This gives the notion of *mixing time*. Aharonov et al. [1] already tried to formalize the concept of

mixing for unitary walks, by slightly changing the definition to overcome the fact that unitary walks are never ergodic (Theorem 6.11). However, with our framework one could possibly extend the original definition of mixing time, and a conjecture is that the mixing time is tightly related to the second highest eigenvalue of the quantum channel, much like in the classical case [24].

Appendix A

Symbolic and Numerical Computations

A.1 Hitting times in Grover's algorithm with Wolfram Mathematica

In this section we briefly show how to use Wolfram Mathematica in order to carry out a symbolic computation of the formula of Theorem 4.10. Here we show how to do it for the case of Grover's algorithm (proof of Theorem 5.3).

We start the notebook by defining symbols for the two elements of the computational basis, and the identity matrix:

```
In[1]:= s0 = {{1},{0}}

In[2]:= s1 = {{0},{1}}

In[3]:= I2 = IdentityMatrix[{2, 2}]

In[4]:= I4 = IdentityMatrix[{4, 4}]
```

Recall that, in the proofs of Theorems 5.2 and 5.3, the element $|0^*\rangle$ corresponds to the starting state $|+\rangle$, while $|1^*\rangle$ is its orthogonal state within span $\{|x_0\rangle, |+\rangle\}$.

We can now define a symbol for the vector $|x_0\rangle$, our target state:

```
In[5]:= x0 = 1/Sqrt[N] s0 + Sqrt[(N-1)/N] s1
```

Moving on to the matrices, we need to define $\Pi_{-x_0} = \mathbb{1} - |x_0\rangle\langle x_0|$, and Grover's operator G which, within this two-dimensional space, is simply a rotation of the angle 2γ:

```
In[6]:= Pnx0 = Simplify[I2 - x0.Transpose[x0],
        Element[N, PositiveIntegers]]
```

$$\text{Out[6]=} \quad \begin{pmatrix} \frac{-1+N}{N} & -\frac{\sqrt{-1+N}}{N} \\ -\frac{\sqrt{-1+N}}{N} & \frac{1}{N} \end{pmatrix}$$

```
In[7]:= γ = ArcSin[1/Sqrt[N]]
```

$$\text{Out[7]=} \quad \text{ArcSin}\left[\frac{1}{\sqrt{N}}\right]$$

```
In[8]:= G = {{Cos[2γ], -Sin[2γ]}, {Sin[2γ], Cos[2γ]}}
```

$$\text{Out[8]=} \quad \begin{pmatrix} \text{Cos}[2\,\text{ArcSin}[\frac{1}{\sqrt{N}}]] & -\text{Sin}[2\,\text{ArcSin}[\frac{1}{\sqrt{N}}]] \\ \text{Sin}[2\,\text{ArcSin}[\frac{1}{\sqrt{N}}]] & \text{Cos}[2\,\text{ArcSin}[\frac{1}{\sqrt{N}}]] \end{pmatrix}$$

The crucial part comes now: we need to define the positive maps in such a way that Mathematica can invert them. Here the arguments of Section 2.1 come into play: we will construct the maps as 4×4 matrices, and the density operators as 4-dimensional vectors. For this purpose, we will represent everything with the basis

$$\{|0\rangle\langle 0|, |0\rangle\langle 1|, |1\rangle\langle 0|, |1\rangle\langle 1|\}$$

This basis will be particularly helpful for the trace computation as the elements of the diagonal of a density matrix are exactly the first and the last entries of the corresponding vector representation. We define the CP maps $\mathcal{G}$ and $\mathcal{P}_{-x_0}$ as Mathematica functions:

```
In[9]:= Gmap = Function[ρ, G.ρ.ConjugateTranspose[G]]
```

```
In[10]:= Pmap = Function[ρ, Pnx0.ρ.ConjugateTranspose[Pnx0]]
```

In order to derive the matrices, we need to apply the following formulas:

$$G_{ij} = \mathrm{Tr}(B_i^\dagger \mathcal{G}(B_j)) \tag{A.1}$$

$$P_{ij} = \mathrm{Tr}(B_i^\dagger \mathcal{P}_{-x_0}(B_j)) \tag{A.2}$$

where B_i is the i-th element of the basis we chose above. We define the array of basis elements in Mathematica, and initialize two empty 4×4 matrices:

```
In[11]:= Bs = { s0.Transpose[s0], s0.Transpose[s1],
         s1.Transpose[s0], s1.Transpose[s1] }
In[12]:= Gmx = ConstantArray[0, {4, 4}]
In[13]:= Pmx = ConstantArray[0, {4, 4}]
```

We obtain the matrix representation of our maps by applying (A.1):

```
In[14]:= For[i = 1, i <= 4, i++, For[j = 1, j <= 4, j++, Gmx[[i,j]] =
         FullSimplify[Tr[ConjugateTranspose[Bs[[i]]].Gmap[Bs[[j]]]],
         Element[N, PositiveIntegers]]]]
In[15]:= For[i = 1, i <= 4, i++, For[j = 1, j <= 4, j++, Pmx[[i,j]] =
         FullSimplify[Tr[ConjugateTranspose[Bs[[i]]].Pmap[Bs[[j]]]],
         Element[N, PositiveIntegers]]]]
```

Now we have all we need to compute the hitting time formula. First, we apply Lemma 4.11 in order to compute the matrix representation for $\mathcal{G}^\sigma_{-x_0}$:

```
In[16]:= Bmx = FullSimplify[(I4 - (1 - p)Gmx)/p,
         Element[N, PositiveIntegers]]
In[17]:= Gsmx = FullSimplify[Pmx.Gmx.Inverse[Bmx],
         Element[N, PositiveIntegers]]
```

We can compute the vector representation of the matrix $(\mathcal{I} - \mathcal{G}^\sigma_{-x_0})(|0^*\rangle\langle 0^*|)$ as follows:

```
In[18]:= ρ0=Transpose[{Flatten[s0.Transpose[s0]]}]
In[19]:= HittingState = FullSimplify[Inverse[I4 - Gsmx].ρ0,
         Element[N, PositiveIntegers]]
```

And the trace is taken as anticipated, by summing the first and last entry of the vector representation:

```
In[20]:= FullSimplify[HittingState[[1]] + HittingState[[4]],
         Element[N, PositiveIntegers]]
```

$$\text{Out[20]=} \quad \left\{\frac{16N + 16p - 20Np + N^2p^2}{8N - 4Np}\right\}$$

Notice that this is not the hitting time, but only the simplified expression for $\mathrm{Tr}\,(\mathcal{I} - \mathcal{G}^{\sigma}_{-x_0})(|0^*\rangle\langle 0^*|)$. The generalized hitting time for Grover's algorithm in this setting is given by multiplying this quantity by $\frac{1}{p}$, which is the expected number of steps before the first measurement occurs.

A.2 Hitting times on the cycle with Numpy

In this section we show how to compute the hitting time formula numerically using numpy, and we take the case of the quantum walk on the (even) cycle as example (Section 4.2).

```
import numpy as np
import json
import matplotlib.pyplot as plt


N = 10 # cycle length is 2N
UP = 0
DOWN = 1
```

As in the previous section, the most challenging part is to represent quantum channels in matrix form since we want to invert them. For this purpose, we define some helper functions to handle basis elements.

```
# constructs the x-th standard basis element for a
# size-dimensional Hilbert space
def bsi(x, size):
```

```
    return (np.arange(size) == x).astype(int)

# constructs the element |x, d>, where x is the site of
# the cycle and d is the value of the coin register (0 or 1)
def bs(d, x, size):
    return bsi(2*x + d, size)

# constructs the matrix |x, d><x, d|
def proj(d, x, size):
    v = bs(d, x, size)
    return np.outer(v, v)
```

Using the above functions to construct the basis states, we can use them to construct the basis states for the Hilbert-Schmidt space.

```
# returns the x-th element of the Hilbert-Schmidt basis
def hsch_bs(x, size):
    x1 = x % (size)
    x2 = x // (size)
    if x2 > size: raise ValueError
    return np.asmatrix(np.outer(bsi(x1, size), bsi(x2, size)))
```

The following functions compute the matrix/vector representations of given CP maps/density matrices, so that the Hilbert-Schmidt product $\langle A, B\rangle := \mathrm{Tr}(A^\dagger B)$ then becomes a dot product. Here, the CP maps are passed as python functions, taking a $n \times n$ density matrix as input and returning a density matrix in output.

```
# returns the matrix representation of the function cpmap.
# This assumes the function to be linear.
def to_matrix(cpmap, size):
    mxsize = size ** 2
```

```
    mx = np.zeros((mxsize, mxsize))
    for i in range(mxsize):
        for j in range(mxsize):
            mx[i, j] = np.trace(hsch_bs(i, size).H @
                                cpmap(hsch_bs(j, size)))
    return mx

# returns the vector representation of
# the size by size matrix given on input.
def to_vector(rho, size):
    vecsize = size ** 2
    vec = np.zeros((vecsize))
    for i in range(vecsize):
        vec[i] = np.trace(hsch_bs(i, size).H @ rho)
    return vec
```

Lastly, we need a way to compute the trace of a matrix given its vector representation. This is nicely achieved by doing a dot product with the vector representation of the identity matrix. This because $\langle \mathbb{1}, A \rangle = \mathrm{Tr}(\mathbb{1}^\dagger A) = \mathrm{Tr}\, A$.

```
# returns the vector representation of the identity.
# A Hilbert-Schmidt product with this vector
# yields the trace of the original matrix.
def trace_vector(size):
    return to_vector(np.eye(size), size)

# returns the trace of the matrix, given its
# vector representation.
def vx_trace(vrho, size):
    return np.dot(trace_vector(size), vrho)
```

Now we have all the tools to compute the hitting time formula. We construct the unitary U for the quantum walk on a cycle:

$$|x,\uparrow\rangle \mapsto \frac{|x+1,\uparrow\rangle + |x-1,\downarrow\rangle}{\sqrt{2}}$$

```
def up_shift(x):
    return (bs(UP,   (x+1)%(2*N), size=4*N) +
            bs(DOWN, (x-1)%(2*N), size=4*N))/np.sqrt(2)
```

$$|x,\uparrow\rangle \mapsto \frac{|x+1,\uparrow\rangle - |x-1,\downarrow\rangle}{\sqrt{2}}$$

```
def down_shift(x):
    return (bs(UP,   (x+1)%(2*N), size=4*N) -
            bs(DOWN, (x-1)%(2*N), size=4*N))/np.sqrt(2)
```

Here, we defined the indices of the columns of the matrix as follows: if the state is $|x,d\rangle$, then the corresponding column is $2x+d$, where $d=0$ means $\uparrow$, and $d=1$ means $\downarrow$. This can also be seen in the implementation of the function `bs(d, x, size)`.

```
U = np.zeros((4*N, 4*N))


for x in range(2*N):
    U[:, 2*x + UP]   = up_shift(x)
    U[:, 2*x + DOWN] = down_shift(x)
```

Moreover, we need the projector $\Pi_{-N} = \mathbb{1} - \Pi_{N,\uparrow} - \Pi_{N,\downarrow}$:

```
P = np.eye(4*N) - proj(UP, N, size=4*N) - proj(DOWN, N, size=4*N)
```

We define the maps $\mathcal{U}(\rho) = U\rho U^\dagger$ and $\mathcal{U}_{-N}(\rho) = \Pi_{-N}U\rho U^\dagger\Pi_{-N}$, and we compute their matrix representations using the functions we defined above.

```
U = np.asmatrix(U)
P = np.asmatrix(P)
```

```
PU = P @ U

def Umap(rho):
    return U @ rho @ U.H

def PUmap(rho):
    return PU @ rho @ PU.H

Umx = to_matrix(Umap, size=4*N)
PUmx = to_matrix(PUmap, size=4*N)
Imx = np.eye((4*N) ** 2)
```

Now we can compute the formula: setting the starting state to $|0,\uparrow\rangle\langle 0,\downarrow|$, we first apply Lemma 4.11 in order to compute the matrix representation of $\mathcal{U}^{\sigma}_{-N}$.

```
p = 1
cht = {}

starting_state = to_vector(proj(UP, 0, size=4*N), size=4*N)
while p <= 100:
    PUsigmamx = p/100*PUmx @ np.linalg.inv(Imx - (1-p/100)*Umx)
    cht[p] = vx_trace(np.linalg.inv(Imx - PUsigmamx)
                      @ starting_state, size=4*N)

    p += 1
```

The computed hitting time with measurement probability p is stored in `cht[p]`.

Acknowledgements

IBM, the IBM logo, and ibm.com are trademarks of International Business Machines Corp., registered in many jurisdictions worldwide. Other product and service names might be trademarks of IBM or other companies. The current list of IBM trademarks is available at https://www.ibm.com/legal/copytrade.

Bibliography

[1] Dorit Aharonov, Andris Ambainis, Julia Kempe, and Umesh Vazirani. Quantum walks on graphs. *Conference Proceedings of the Annual ACM Symposium on Theory of Computing*, 01 2001.

[2] A. Ambainis. Quantum walk algorithm for element distinctness. In *45th Annual IEEE Symposium on Foundations of Computer Science*, pages 22–31, 2004.

[3] MD SAJID ANIS, Héctor Abraham, et al. Qiskit: An open-source framework for quantum computing, 2021.

[4] Jean-Christophe Bourin and Eun-Young Lee. On the russo-dye theorem for positive linear maps. *Linear Algebra and its Applications*, 571, 02 2019.

[5] Andrew M. Childs and Jason M. Eisenberg. Quantum algorithms for subset finding. *Quantum Information and Computation*, 5(7), Nov 2005.

[6] Martin Dyer and Haiko Müller. Counting perfect matchings and the switch chain. *SIAM Journal on Discrete Mathematics*, 33, 05 2017.

[7] Lov K. Grover. A fast quantum mechanical algorithm for database search. In *Proceedings of the Twenty-Eighth Annual ACM Symposium on*

Theory of Computing, STOC '96, page 212–219, New York, NY, USA, 1996. Association for Computing Machinery.

[8] Ji Guan, Yuan Feng, and Mingsheng Ying. Decomposition of quantum markov chains and its applications. *Journal of Computer and System Sciences*, 95, 03 2018.

[9] Stan Gudder. Quantum markov chains. *Journal of Mathematical Physics*, 49, 07 2008.

[10] Mark Jerrum. *Counting, Sampling and Integrating: Algorithm and Complexity*. Birkhäuser Basel, Basel, 2003.

[11] Karuna Kadian, Sunita Garhwal, and Ajay Kumar. Quantum walk and its application domains: A systematic review. *Computer Science Review*, 41:100419, 2021.

[12] J Kempe. Quantum random walks: An introductory overview. *Contemporary Physics*, 44(4):307–327, Jul 2003.

[13] Julia Kempe. Discrete quantum walks hit exponentially faster. In Sanjeev Arora, Klaus Jansen, José D. P. Rolim, and Amit Sahai, editors, *Approximation, Randomization, and Combinatorial Optimization. Algorithms and Techniques*, pages 354–369, Berlin, Heidelberg, 2003. Springer Berlin Heidelberg.

[14] Viv Kendon. Decoherence in quantum walks – a review. *Mathematical Structures in Computer Science*, 17(06), Nov 2007.

[15] Viv Kendon and Ben Tregenna. Decoherence can be useful in quantum walks. *Physical Review A*, 67(4), Apr 2003.

[16] Carlos F. Lardizabal. Mean hitting times of quantum markov chains in terms of generalized inverses. *Quantum Information Processing*, 18(8), Jul 2019.

[17] Frédéric Magniez, Ashwin Nayak, Peter C Richter, and Miklos Santha. On the hitting times of quantum versus random walks. *Algorithmica*, 63(1):91–116, 2012.

[18] Michael A Nielsen and Isaac L Chuang. Quantum computation and quantum information. *Phys. Today*, 54(2):60, 2001.

[19] J. R. Norris. *Markov Chains*. Cambridge Series in Statistical and Probabilistic Mathematics. Cambridge University Press, 1997.

[20] Greg Orosi. Duration of play in the gambler's ruin problem: A novel derivation. *Applied Mathematics E-Notes*, 18:268–274, 2018.

[21] T. Schoning. A probabilistic algorithm for k-sat and constraint satisfaction problems. In *40th Annual Symposium on Foundations of Computer Science (Cat. No.99CB37039)*, pages 410–414, 1999.

[22] Benjamin Schumacher and Michael Westmoreland. *Quantum Processes Systems, and Information*. Cambridge University Press, 2010.

[23] Neil Shenvi, Julia Kempe, and K. Birgitta Whaley. Quantum random-walk search algorithm. *Physical Review A*, 67(5), May 2003.

[24] Alistair Sinclair. *Markov chains and rapid mixing*, pages 42–62. Birkhäuser Boston, Boston, MA, 1993.

[25] M. Szegedy. Quantum speed-up of markov chain based algorithms. In *45th Annual IEEE Symposium on Foundations of Computer Science*, pages 32–41, 2004.

[26] James D. Whitfield, César A. Rodríguez-Rosario, and Alán Aspuru-Guzik. Quantum stochastic walks: A generalization of classical random walks and quantum walks. *Physical Review A*, 81(2), Feb 2010.

[27] Michael M. Wolf. Quantum channels & operations: Guided tour. Lecture notes available at https://www-m5.ma.tum.de/foswiki/pub/M5/Allgemeines/MichaelWolf/QChannelLecture.pdf.