Università degli studi di Padova

Scuola di Ingegneria

Corso di laurea magistrale in Ingegneria dell'Automazione

Iterated projection methods with classical and quantum applications

Relatore: Dott. Francesco Ticozzi

Laureando: Luca Zuccato

Anno accademico 2015 - 2016

Contents

1 Introduction **1**
1.1 History . . . 1
1.2 Mathematical background . . . 1

2 Basic Theory **4**
2.1 Iterated Projection Theorem . . . 4
2.1.1 Rate of Convergence . . . 5
2.2 Extensions . . . 7
2.2.1 Row-Action Methods . . . 7
2.2.2 Dykstra's Algorithm . . . 8
2.3 Typical applications . . . 9
2.3.1 Solving Constrained L-S Matrix Problems . . . 9
2.3.2 MMUP: Matrix Model Updating Problem . . . 11
2.3.3 Projection Methods on Quantum Information Science . . . 12
2.3.4 Fixed-Order Control Design for LMI Control problems using MAP . . 13

3 Iterated Quantum Maps **17**
3.1 Statistical Description of Finite-Dimensional Quantum Systems . . . 17
3.2 Open System Dynamics . . . 19
3.3 CPTP Projections . . . 21
3.4 Iterated CPTP Map Theorem . . . 24
3.5 Quantum Application . . . 25
3.5.1 Example . . . 27

4 Bregman's Theory **30**
4.1 Bregman's Divergences and their Properties . . . 30
4.2 Generalized Pythagorean Theorem . . . 32
4.3 Iterated Convergence Results . . . 33
4.4 Quantum Bregman's Divergences . . . 34

A Tensor Product and Partial Trace **36**

B MATLAB Code **38**

Chapter 1

Introduction

1.1 History

In various scientific fields, it is often required to solve the following problem: find a point in the intersection of closed subspaces (or sets) that minimizes the distance from a given point of the whole space. Von Neumann, in 1933 [18], found an iterative approach to solve this problem: he found that the projection onto the intersection of two subspaces can be found alternating projections onto the single subspaces. Starting from this discovery, Halperin [14] extended Von Neumann's theorem for the case of N subspaces with non empty intersection. Using this alternating projections approach, different algorithms where invented. Among these, we recall Kaczmarz method in 1937 [16], in order to find the solution of a linear system, MAMS in 1954 [1][17], useful to find a feasible solution for system inequalities, and Dykstra's alternating projections algorithm for general convex sets in 1986 [11] (etc...). This algorithms are used in order to solve different problems like constrained least-squares matrix minimization problems, matrix model updating problem in order to adapt a given model to measured data , control design, etc...

In 1967 Bregman [5] extended the concept of distance defining what now are called Bregman divergencies. With this extended concept, he defined the corresponding generalized projections and he gave an iterative Bregman projection theorem. Similar results were developed by Csiszar [8] in the field of information theory, using relative entropy as a pseudo-distance.

In this work some classical results of alternating projections and Bregman's theory will be presented (Chapters 2 and 4). In Chapter 3 we will propose an application of Von Neumann-Halperin's theory quantum maps.

1.2 Mathematical background

In this section we present some definitions that will be essentials to this work. Details can be found in specific algebra books.

Definition 1.1 *A **vector space** V is a set of elements (vectors) with two operations: addition $(+)$ and scalar multiplication. They satisfy the following properties:*

- $u + v = v + u$, $\forall\ u, v \in V$;
- $u + (v + w) = (u + v) + w$, $\forall\ u, v, w \in V$;
- *it exists a null vector* e *such that* $e + v = v$, $\forall\ v \in V$
- $\lambda(\mu v) = (\lambda\mu)v$, $\forall\ v \in V$ *and* λ, μ *scalars;*
- $(\lambda + \mu)v = \lambda v + \mu v$, $\forall\ v \in V$ *and* λ, μ *scalars;*
- $\lambda(u + v) = \lambda u + \lambda v$, , $\forall\ v, u \in V$ *and* λ *scalar;*
- $0v = e$, $\forall\ v \in V$;
- $1v = v$, $\forall\ v \in V$.

Definition 1.2 *A set* $\mathcal{C}$ *in a vector space is said to be* ***convex*** *if*

$$(1 - \alpha)x + \alpha y \in \mathcal{C}$$

for all $x, y \in \mathcal{C}$, *and* $0 \leq \alpha \leq 1$

Definition 1.3 *A* ***metric space*** X *is a vector space where it is defined a distance function* $d : X \times X \to \mathbb{R}_+$ *with the following properties for all* $x, y, z \in X$:

- $d(x, y) = 0 \iff x = y$;
- $d(x, y) = d(y, x)$;
- $d(x, z) \leq d(x, y) + d(y, z)$.

Definition 1.4 *Let* V *be a vector space over* F $(\mathbb{R}, \mathbb{C})$. *The* ***inner product*** *is a map*

$$\langle \cdot, \cdot \rangle : V \times V \to F$$

that satisfies:

1. $\langle x, x \rangle \geq 0$ $(= 0 \iff x = 0)$;
2. $\langle x, y \rangle = \langle y, x \rangle^*$;
3. $\langle \alpha_1 x + \alpha_2 y, z \rangle = \alpha_1 \langle x, z \rangle + \alpha_2 \langle y, z \rangle$ *where* $\alpha_1, \alpha_2 \in F$

Examples:

- In the Euclidean space $\mathbb{R}^n$ the inner product is given by:

$$\langle x, y \rangle = x_1 y_1 + ... + x_n y_n;$$

- In $\mathbb{C}^n$, the inner product is given by:

$$\langle x, y \rangle = \sum_{j=1}^{n} x_j y_j^*$$

- In L^2, set of square integrable functions, the inner product is given by:

$$\langle g, f \rangle = \int g^*(x) f(x) dx$$

- In $l^2 = \{\{x_j\} \in \mathbb{C} \quad s.t. \quad \sum_{j=1}^{\infty} |x_j|^2 < \infty\}$, the inner product is given by:

$$\langle x, y \rangle = \sum_{j=1}^{\infty} x_j y_j^*$$

Definition 1.5 *An* ***inner product space*** *is a linear space in which there is defined an inner product between pairs of elements of the space.*

Definition 1.6 *A sequence of points $\{x_i\}$ in a metric space with metric d is called a* ***Cauchy sequence*** *if $\forall\ \epsilon > 0$ there exists an integer N_ϵ such that $d(x_i, x_j) < \epsilon$ when $i > N_\epsilon$ and $j > \mathbb{N}_\epsilon$.*

Definition 1.7 *A metric space is called* ***complete*** *in the norm induced by its inner product if every Cauchy sequence of points in it converges to a point in the space.*

Definition 1.8 *An inner product space, which is also complete, is called* ***Hilbert space****.*

Definition 1.9 *A* ***projection*** *is a linear transformation P from a vector space to itself such that $P^2 = P$.*
Let V be an inner product space and consider a subspace $W \subset V$. Every vector $v \in V$ can be uniquely written as $v = w_1 + w_2$, $w_1 \in W$ and $w_2 \in W^\perp$. The ***orthogonal projection*** *of v onto W is defined as $P_W(v) = w_1$.*

It satisfies the following properties:

1. It is linear: $P_W(v) = P_W v$;
2. It is idempotent; $P^2 = P$;
3. It is self-adjoint: $\forall v_1, v_2 \in V\ \langle P_W(v_1), v_2 \rangle = \langle v_1, P_W(v_2) \rangle$

If $\omega_1, ..., \omega_n$ is an orthonormal basis of W, the orthonormal projection can be written as:

$$P_W(v) = \sum_{i=1}^{n} \langle v, \omega_i \rangle \omega_i$$

Before presenting the algorithms that are built upon iterated methods, we recall an important theorem for Hilbert spaces and projections. In the next sections $\mathcal{H}$ denotes a general Hilbert space.

Theorem 1.1 (Kolmogorov's criterion) *Let x be a vector in $\mathcal{H}$ and $\mathcal{C}$ be a closed convex subset of $\mathcal{H}$. Then $\exists!\ c_0 \in \mathcal{C}$ such that $\|x - c_0\| \leq \|x - c\|$, $\forall\ c \in \mathcal{C}$.*
Moreover, c_0 is the unique minimizing vector if and only if $\langle x - c_0, c - c_0 \rangle \leq 0$, $\forall \in C$.

Note: *Let $\mathcal{X}$ be a closed subset of $\mathcal{H}$ and let x be a generic point in $\mathcal{H}$. We will denote with $P_{\mathcal{X}}(x)$ the orthogonal projection of x onto the subset $\mathcal{X}$.*

Chapter 2

Basic Theory

2.1 Iterated Projection Theorem

In this Section we are going to see the original theorems that brought to the development of more advanced algorithms, which will be discussed in Section 2.3. These algorithms are used to solve linear systems ($Ax = b$), linear feasibility problems (i.e. find $x \in \mathbb{R}^n$ s.t. $Ax \leq b$), or, in general, convex feasibility problems (find $x \in \bigcap \mathcal{C}_i$ where $\mathcal{C}_i$ is closed, convex for $1 \leq i \leq m$). In Section 2.4 we will show some application of these alternating projections methods. As said in the history section, Von Neumann was interested to find the projection of a given point in $\mathcal{H}$ onto the intersection of two closed subspaces. Before seeing his theorem let us give the following definition:

Theorem 2.1 (Von Neumann's alternating projections theorem) *Let $\mathcal{N}$, $\mathcal{M}$ be closed subspaces of $\mathcal{H}$. Then for each $x \in \mathcal{H}$:*

$$\lim_{n\to\infty} (P_{\mathcal{N}} P_{\mathcal{M}})^n x = P_{\mathcal{M}\cap\mathcal{N}} x$$

Proof. Let us consider the sequences Σ_1 and Σ_2 of operators $P_{\mathcal{M}}$, $P_{\mathcal{N}}P_{\mathcal{M}}$, $P_{\mathcal{M}}P_{\mathcal{N}}P_{\mathcal{M}}$,..., and $P_{\mathcal{N}}$, $P_{\mathcal{M}}P_{\mathcal{N}}$, $P_{\mathcal{N}}P_{\mathcal{M}}P_{\mathcal{N}}$,..., respectively. We have to show that both sequences have same limit T and that it is $T = P_{\mathcal{M}\cap\mathcal{N}}$.
Let T_n be the n-th operator of Σ_1. It holds:

$$\langle T_m x, T_n y\rangle = \langle T_{m+n-\delta} x, y\rangle,$$

where $\delta = 1$ if m and n have the same parity, it is 0 otherwise.
We need to show that if $x \in \mathcal{H}$, then $\lim_{n\to\infty} T_n x$ exists. It holds:

$$\begin{aligned}
\|T_m x - T_n x\|^2 &= \langle T_m x - T_n x, T_m x - T_n x\rangle \\
&= \langle T_m x, T_m x\rangle - \langle T_m x, T_n x\rangle - \langle T_n x, T_m x\rangle + \langle T_n x, T_n x\rangle \\
&= \langle T_{2m-1} x, x\rangle + \langle T_{2n-1} x, x\rangle - 2\langle T_{m+n-\delta} x, x\rangle \\
&= \langle T_{2m-1} x, x\rangle + \langle T_{2n-1} x, x\rangle - 2\langle T_{2k-1} x, x\rangle.
\end{aligned}$$

$m+n-k$ is always odd, so the last term has been rewritten with k an integer number. Now

$$\langle T_{2i-1} x, x\rangle = \langle T_i x, T_i x\rangle = \|T_i x\|^2,$$

we have that

$$\|T_{i+1}x\|^2 = \langle T_{2i+1}x, x\rangle.$$

$T_{i+1}x$ is either $P_{\mathcal{M}}T_ix$ or $P_{\mathcal{N}}T_ix$. So, it holds that

$$\|T_{i+1}x\|^2 \le \|T_ix\|^2.$$

So, for all i, it holds:

$$\langle T_{2i-1}x, x\rangle \ge \langle T_{2i+1}x, x\rangle,$$

therefore $\lim_{i\to\infty}\langle T_{2i-1}x, x\rangle$ exists and it implies

$$\lim_{m,n\to\infty} \|T_mx - T_nx\| = 0.$$

Let us denote by x^* the limit for T_nx. If T is defined by the condition $Tx = x^*$, then $dom(T) = \mathcal{H}$ and T is singular valued, T is linear and by

$$\lim_{m,n\to\infty}\langle T_mx, T_ny\rangle = \lim_{m,n\to\infty}\langle T_{m+n-\delta}x, y\rangle,$$

it follows:

$$\langle Tx, Ty\rangle = \langle Tx, y\rangle.$$

So, T is a projection $P_{\mathcal{L}}$. Now, if $x \in \mathcal{M}\cap\mathcal{N}$, then $P_{\mathcal{M}}x = P_{\mathcal{N}}x = x$, $T_nx = x$ and $Tx = x$. So, $x \in \mathcal{L}$ and, by that, $\mathcal{M}\cap\mathcal{N} \subseteq \mathcal{L}$. Now, it holds that $P_{\mathcal{M}}T = P_{\mathcal{N}}T = T$, let $y \in \mathcal{H}$ and $Ty = x \in \mathcal{L}$. Then $P_{\mathcal{M}}x = P_{\mathcal{M}}Ty = Ty = x \in \mathcal{M}$, and $P_{\mathcal{N}}x = P_{\mathcal{N}}Ty = Ty = x \in \mathcal{N}$, which implies $\mathcal{L} \subseteq \mathcal{M}\cap\mathcal{N}$.
Now, making the same for Σ_2 it is clear that its limit $T' = P_{\mathcal{M}\cap\mathcal{N}}$, so $T = T'$ and the proof is complete. □

The generalization to the intersection of multiple subspaces was given by Halperin:

Theorem 2.2 (Halperin) *If $\mathcal{M}_1$,...,$\mathcal{M}_r$ are closed subspaces in $\mathcal{H}$, then $\forall x \in \mathcal{H}$*

$$\lim_{n\to\infty}(P_{\mathcal{M}_1}...P_{\mathcal{M}_r})^n x = P_{\bigcap_{i=1}^r \mathcal{M}_i}x$$

A proof for this theorem can be found in Halperin's original work [14].

2.1.1 Rate of Convergence

These two theorems gave the basis to the development of different algorithms. For this reason we are interested in the convergence of alternating projections. The rate is linked to the angle between the subspaces. We recall their definitions and properties.
Define the function arccos : $[-1, 1] \to [-\frac{\pi}{2}, \frac{\pi}{2}]$. We will use only the elements in interval $[0, 1]$. Then the *angle* $\theta(\mathcal{M}, \mathcal{N})$ between the closed subspaces $\mathcal{M}$ and $\mathcal{N}$ of $\mathcal{H}$ is the element of $[0, \frac{\pi}{2}]$.

Definition 2.1 (Friedrichs) *Define the cosine $c(\mathcal{M}, \mathcal{N})$ between the closed subspaces $\mathcal{M}$ and $\mathcal{N}$ of $\mathcal{H}$ as:*

$$c(\mathcal{M}, \mathcal{N}) = \sup\{|\langle x, y\rangle| : x \in \mathcal{M}\cap(\mathcal{M}\cap\mathcal{N})^\perp, \|x\| \le 1, y \in \mathcal{N}\cap(\mathcal{M}\cap\mathcal{N})^\perp, \|y\| \le 1\}.$$

Then the angle is given by:

$$\theta(\mathcal{M}, \mathcal{N}) = \arccos(c(\mathcal{M}, \mathcal{N})).$$

Definition 2.2 (Dixmier) *Define the cosine* $c_0(\mathcal{M},\mathcal{N})$ *as:*

$$c_0(\mathcal{M},\mathcal{N}) = \sup\{|\langle x,y\rangle| : x\in\mathcal{M}, \|x\|\le 1, y\in\mathcal{N}, \|y\|\le 1\}.$$

Then the minimal angle *is given by:*

$$\theta_0(\mathcal{M},\mathcal{N}) = \arccos(c_0(\mathcal{M},\mathcal{N})).$$

Properties:

1. if $\mathcal{M}\cap\mathcal{N}=\{0\}$ then $c_0(\mathcal{M},\mathcal{N}) = c(\mathcal{M},\mathcal{N})$;
2. Some consequences of definitions are:
   i) $0\le c(\mathcal{M},\mathcal{N})\le c_0(\mathcal{M},\mathcal{N})\le 1$;
   ii) $c(\mathcal{M},\mathcal{N}) = c(\mathcal{N},\mathcal{M})$ and $c_0(\mathcal{M},\mathcal{N}) = c_0(\mathcal{N},\mathcal{M})$;
   iii) $c_0(\mathcal{M},\mathcal{N}) = c_0(\mathcal{M}\cap(\mathcal{M}\cap\mathcal{N})^\perp, \mathcal{N}\cap(\mathcal{M}\cap\mathcal{N})^\perp)$;
   iv) $|\langle x,y\rangle|\le c_0(\mathcal{M},\mathcal{N})\|x\|\|y\|$ for all $x\in\mathcal{M}$, $y\in\mathcal{N}$.

Lemma 2.1 *The following relations hold:*

1. $c(\mathcal{M},\mathcal{N}) = c_0(\mathcal{M},\mathcal{N}\cap(\mathcal{M}\cap\mathcal{N})^\perp) = c_0(\mathcal{M}\cap(\mathcal{M}\cap\mathcal{N})^\perp,\mathcal{N})$;
2. $c_0(\mathcal{N},\mathcal{M}) = \|P_\mathcal{M}P_\mathcal{N}\| = \|P_\mathcal{M}P_\mathcal{N}P_\mathcal{M}\|^{\frac{1}{2}}$;
3. $c(\mathcal{M},\mathcal{N}) = \|P_\mathcal{M}P_\mathcal{N} - P_{\mathcal{M}\cap\mathcal{N}}\| = \|P_\mathcal{M}P_\mathcal{N}P_{(\mathcal{M}\cap\mathcal{N})^\perp}\|$.

Having the definition of the angle, we next state the theorem that gives the exact rate in case of projection onto two subspaces.

Theorem 2.3

$$\|(P_{\mathcal{M}_2}P_{\mathcal{M}_1})^n - P_{\mathcal{M}_1\cap\mathcal{M}_2}\| = c(\mathcal{M}_1,\mathcal{M}_2)^{2n-1}$$

$(n=1,2,...)$.

In case of projecting onto more subspaces we cannot give an exact expression but we give an upper bound:

Theorem 2.4 *For each* $i=1,2,...,r$, *let* $\mathcal{M}_i$ *be a closed subspace of* $\mathcal{H}$. *Then, for each* $x\in\mathcal{H}$, *and for any integer* $n\ge 1$ *it holds:*

$$\|(P_{\mathcal{M}_r}...P_{\mathcal{M}_1})^n x - P_{\bigcap_{i=1}^r\mathcal{M}_i}x\| \le c^{\frac{n}{2}}\|x - P_{\bigcap_{i=1}^r\mathcal{M}_i}x\|,$$

where

$$c = 1 - \prod_{i=1}^{r-1}\sin^2\theta_i,$$

and θ_i *is the angle between* $\mathcal{M}_i$ *and* $\bigcap_{j=i+1}^r\mathcal{M}_j$.

Remark: By Theorem 2.4 we can see that a condition to finite time convergence of iterated projection is given by $c=0$ which is satisfied if $c(\mathcal{M}_i,\mathcal{M}_j)=0$ for all $1\le i,j\le r$, in other words: $\left[\mathcal{M}_i\cap\left(\bigcap_{t=1}^r\mathcal{M}_t\right)^\perp\right]\perp\left[\mathcal{M}_j\cap\left(\bigcap_{t=1}^r\mathcal{M}_t\right)^\perp\right]$ for every $i,j=i+1,...,r$.

The proofs of theorems and of the lemma can be found in [12].

2.2 Extensions

2.2.1 Row-Action Methods

The followings are iterative methods developed to solve large and sparse systems, linear and non-linear, equalities (i.e $Ax = b$) and inequalities (i.e. find $x \in \mathbb{R}^n$ s.t. $Ax \leq b$) in a finite dimensional space as said in Section 2.1.
Typically, row action methods involve alternating projections in hyperplanes, linear varieties or closed and convex sets and have these properties:

1. No changes or operations are made on the original matrix A;
2. They only use one row per iteration;
3. At every iteration, the computation of x_{k+1} requires only the value of x_k;
4. For finite dimensional problems, they only require vector arithmetic such as inner products and vector sums.

Definition 2.3 *A sequence of indices $\{i_k\}$ is called a* ***control sequence*** *of a row-action method if at the k-th iteration the convex set $\mathcal{C}_{i_k}$ is used.*

Here are some type of control:

- **Cyclic Control:** $i_k = k \mod n + 1$, where m is the number of convex sets involved in the problem;
- **Almost Cyclic Control:** $i_k \in M = \{1, 2, ..., m\}$ $\forall$ $k \geq 0$ and $\exists \bar{M}$ integer s.t. $\forall k$ $M \subset \{i_{k+1}, ..., i_{k+\bar{M}}\}$
- **Remotest Set Control:** i_k is chosen s.t. $d(x_k, \mathcal{C}_{i_k}) = \max_{i \in M} d(x_k, \mathcal{C}_i)$, x_k is the k-th iteration of the row-action method, $d(x_k, \mathcal{C}_i)$ is the distance from x_k to set $\mathcal{C}_i$;
- **Random Set Control:** i_k is chosen from set $\{1, 2, ..., m\}$ randomly with a probability function that guarantees that every set is chosen, with non zero probability, in every sweep of projection.

The followings are some most used row-action methods.

The relaxation method of Agmon, Motzkin, and Schönberg (MAMS) The problem to solve is the following:

$$\begin{gathered} Ax \leq b \\ A \in R^{m \times n} \\ x \in R^n \\ b \in R^m \end{gathered}$$

The problem can be generalized to any Hilbert space $\mathcal{H}$ to find x in the intersection of m closed half spaces given by $\mathcal{S}_i = \{x \in \mathcal{H} : \langle a_i, x \rangle \leq b_i\}$ $\forall i \in \mathcal{M}$. This is called **linear feasibility problem**.
Given an arbitrary $x_0 \in \mathcal{H}$, a typical step of this method can be described by:

$$x_{k+1} = x_k + \delta_k a_{i_k};$$

where:

$$\delta_k = \min(0, \omega_k \frac{b_{i_k} - < a_{i_k}, x_k}{< a_{i_k}, < a_{i_k} >});$$

where $0 < \epsilon \leq \omega_k \leq 2 - \epsilon < 2$ for all k, a small given ϵ and i_k is chosen by one of the control seen before.
This method does not guarantee the convergence to the nearest vector, in the feasible set, to x_0.

Hildreth's Method Let us consider the following problem:

$$\begin{aligned} &\text{minimize } \|x^2\| \\ &\text{s. t. } \langle a_i, x\rangle \leq b_i \ \forall i \in \mathcal{M} \end{aligned}$$

Let S_i indicate the subspace given by $\langle a_i, x\rangle \leq b_i$. Starting from $x_0 \notin \mathcal{S}_i$ $\forall i$ a typical step is given by:

$$\begin{aligned} x_{k+1} &= x_k + \delta_k a_{i_k}; \\ z^{k+1} &= z^k - \delta_k e_{i_k}; \end{aligned}$$

where:

$$\delta_k = \min\big(z^k_{i_k}, \omega_k \frac{b_{i_k} - \langle a_{i_k}, x_k\rangle}{\langle a_{i_k}, a_{i_k}\rangle}\big),$$

and e_{i_k} have all component zeros except the i_k-th component, which is one; any of the controls described in the beginning can be imposed and it holds $0 < \epsilon \leq \omega_k \leq 2 - \epsilon < 2$ for all k and a given small positive ϵ. Again, i_k follows one of the control introduced before. This algorithm converges to minimal norm. Other examples can be found in [12].

Now we are going to show an important algorithm to find the closest vector into an intersection of closed convex sets (**convex feasibility problem**).

2.2.2 Dykstra's Algorithm

Let $\mathcal{H}$ be a Hilbert space. For a given non empty, closed, convex set $\mathcal{C}$ of $\mathcal{H}$, and $x \in \mathcal{H}$, it exists a unique x^* that solves:

$$\min_{x\in\mathcal{C}} \|x_0 - x\|, \tag{2.2.1}$$

which satisfies the Kolmogorov criterion:

$$x^* \in \mathcal{C}, \quad \langle x_0 - x*, x - x*\rangle \leq 0, \quad \forall x \in \mathcal{C}.$$

Let us consider the case $\mathcal{C} = \bigcap_1^r \mathcal{C}_i$, where $\mathcal{C}_i$ is a closed, convex set in $\mathcal{H}$. Moreover, we will assume that $\forall y \in \mathcal{H}$, $P_{\mathcal{C}}(y)$ is not trivial, while $P_{\mathcal{C}_i}(y)$ is easy to calculate.
In order to solve the problem (2.2.1), this algorithm generates two sequences: the iterates $\{x_i^n\}$ and the increments $\{I_i^n\}$, with $n \in \mathbb{N}$ and $i = 1, ..., r$.

$$\begin{aligned} x_0^n &= x_r^{n-1} \\ x_i^n &= P_{\mathcal{C}_i}(x_{i-1}^n - I_i^{n-1}) \\ I_i^n &= x_i^n - (x_{i-1}^n - I_i^{n-1}) \end{aligned}$$

Where the initial values are $x_r^0 = x_0$, $I_i^0 = 0$.

Note:

- The increment I_i^{n-1} associated with C_i in the previous cycle is always subtracted before projecting into $\mathcal{C}_i$;
- If $\mathcal{C}_i$ is a subspace, then $P_{\mathcal{C}_i}$ is linear and it is not required , in the n-th cycle, to subtract the increment I_i^{n-1} before projectiong onto $\mathcal{C}_i$. So, in this case, Dykstra's algorithm reduces to MAP procedure.
- The following relations hold:

$$\begin{aligned} x_r^{n-1} - x_1^n &= I_1^{n-1} - I_1^n; \\ x_{i-1}^n - x_i^n &= I_i^{n-1} - I_i^n; \end{aligned}$$

  and

$$x_i^n = x_0 + I_1^n + ... + I_i^n + I_{i+1}^{n-1} + ... + I_r^{n-1}.$$

The next lemma proves the convergence of Dykstra's algorithm.

Lemma 2.2 *Let $\mathcal{C}_1, ..., \mathcal{C}_r$ be closed, convex subsets of a $\mathcal{H}$ and $\mathcal{C} = \bigcap_{i=1}^r \mathcal{C}_i \neq \emptyset$. The sequence $\{x_i^n\}$ generated by the algorithm 2.2.2 converges strongly to $x^* = P_{\mathcal{C}}(x_0)$, for every $x_o \in \mathcal{H}$.*

More details on Dykstra's and other alternating projections algorithms can be found in [12].

2.3 Typical applications

2.3.1 Solving Constrained L-S Matrix Problems

The task is to solve using Dykstra algorithm the following problem:

$$\begin{gathered} \min \|X - A\|_F^2 \\ \text{s.t. } X^T = X \\ L \leq X \leq U \\ \lambda_{min} \geq \epsilon > 0 \\ X \in \mathrm{P} \end{gathered}$$

where $\|M\|_F = \sqrt{\mathrm{tr}(MM^\dagger)}$ is the Frobenius norm, $A, L, U \in R^{n\times n}$; $A \leq B$ means $A_{ij} \leq B_{ij}$ with $1 \leq i, j \leq n$.

The constraints define sets whose intersection identifies a feasibility problem. Those sets are:

$$\mathcal{B} = \{X \in R^{n\times n} : L \leq X \leq U\};$$

$$\epsilon_{pd} = \{X \in R^{n\times n} : X^T = X, \lambda_{\min}(X) \geq \epsilon > 0\};$$

$$\mathrm{P} = \{X \in R^{n\times n} : X = \textstyle\sum_{i=1}^m \alpha_i G_i \text{ for some } \alpha_i \in R, 1 \leq i \leq m\} ;$$

with $1 \leq m \leq \frac{n(n+1)}{2}$.

Property: *In the definition of* P, $G_1, ..., G_m$ *are given* $n \times n$ *non-zero symmetric matrices whose entries are either* 0 *or* 1 *and have the following property: for all st-entry* $1 \leq s$, $t \leq n$, *it exists one and only one* k $(1 \leq k \leq m)$ *s.t.* $(G_k)_{st} = 1$.

Now the problem can be stated as:

$$\min\{\|X - A\|_F^2 : X \in B \cap \epsilon_{pd} \cap \mathrm{P}\}$$

Let us see how the projections onto the singular sets can be found.

Theorem 2.5 *If* $A \in R^{n \times n}$, *then the unique solution to* $\min_{X \in \mathcal{B}} \|X - A\|_F$ *is given by* $P_{\mathcal{B}}(A)$ *defined as:*

$$[P_{\mathcal{B}}(A)]_{ij} = \begin{cases} A_{ij} & if L_{ij} \leq A_{ij} \leq U_{ij}, \\ U_{ij} & if A_{ij} > U_{ij}, \\ L_{ij} & if A_{ij} < L_{ij}. \end{cases}$$

Theorem 2.6 *If* $A \in R^{n \times n}$, *then the unique solution to* $\min_{X \in \mathrm{P}} \|X - A\|_F$ *is given by* $P_{\mathrm{P}}(A) = \sum \bar{\alpha}_k G_k$ *where*

$$\bar{\alpha}_k = \frac{\sum_{i,j=1}^n A_{ij}[G_k]_{ij}}{\sum_{i,j=1}^n [G_k]_{ij}}$$

for $1 \leq k \leq m$.

Theorem 2.7 *Define* $B = \frac{A+A^T}{2}$, *then the unique solution to* $\min_{X \in \epsilon_{pd}} \|X - A\|_F$ *is given by* $P_{\epsilon_{pd}}(A) = Z\mathrm{diag}(d_i)Z^T$ *where*

$$d_i = \begin{cases} \lambda_i(B) & if \lambda_i(B) \geq \epsilon \\ \epsilon & if \lambda < \epsilon \end{cases}$$

and Z *is s.t.* $B = Z\Delta Z^T$ *is a spectral decomposition.*

We can now apply Dykstra's algorithm, which, in this particular case, becomes:

$$\text{Set: } A_0 = A;\ I_{\epsilon_{pd}}^0 = I_B^0 = 0.$$

$$\begin{aligned} \text{For } i = 0, 1, ..: \quad A_i &= P_P(A_i) - I_{\epsilon_{pd}}^i \\ I_{\epsilon_{pd}}^{i+1} &= P_{\epsilon_{pd}}(A_i) - A_i \\ A_i &= P_{\epsilon_{pd}}(A_i) - I_B^i \\ I_B^{i+1} &= P_B(A_i) - A_i \\ A_{i+1} &= P_B(A_i) \end{aligned}$$

Theorem 2.8 *If the closed convex set* $\mathcal{B} \cap \epsilon_{pd} \cap \mathrm{P}$ *is not empty, then for any* $A \in R^{n \times n}$ *the sequences* $\{P_{\mathrm{P}}(A_i)\}$, $\{P_{\epsilon_{pd}}(A_i)\}$ *and* $\{P_{\mathcal{B}}(A_i)\}$ *generated by the previous algorithm converge in the Frobenius norm to the unique solution of the problem of minimum.*
The proofs of the previous theorems can be found in [12].

2.3.2 MMUP: Matrix Model Updating Problem

Consider the following finite element model of a vibrating structure:

$$M\ddot{x}(t) + D\dot{x}(t) + Kx(t) = 0$$

M, D, K are $n \times n$ matrices that denote mass, damping and stiffness of the structure. M is symmetric and positive definite; D, K are symmetric.
For several reasons (modeling errors, etc...) the finite elements data do not agree with measured data and the required structure of the matrices is lost.
It is required to update an analytic finite element model such that the updated model reproduces the measured data while preserving the structure of the matrices. This problem is called **MMUP**(Matrix Model Updating Problem).
By FEM modal analysis, the solutions are of the form $x(t) = ve^{\lambda t}$, λ and v solve the quadratic eigenvalue problem (QEP):

$$(\lambda^2 M + \lambda D + K)v = 0$$

$P(\lambda) = \lambda^2 M + \lambda D + K$ is called *quadratic pencil* and the eigenvalues are given by the roots of $\det(P(\lambda)) = 0$. Eigenvalues and eigenvectors describe the dynamic of the system linking natural frequencies and mode shapes of the structure.

The solutions lead to the following inverse eigenvalue problem for $P(\lambda)$:
Given:

- real $n \times n$ matrices M, K, D ($M = M^T > 0$, $D = D^T$, $K = K^T$) with $\Lambda(P) = \{\lambda_1, ..., \lambda_{2n}\}$ and eigenvectors $\{x_1, ..., x_{2n}\}$;
- a set of p self-conjugate numbers $\{\mu_1, ..., \mu_p\}$, p vectors $\{y_1, ..., y_p\}$, with $p < 2n$.

Find: $\tilde{K}$,$\tilde{D} \in R^{n\times n}$ (both symmetric) s.t. $\Lambda(\tilde{P}(\lambda) = \lambda^2 M + \lambda\tilde{D} + \tilde{K}) = \{\mu_1, ..., \mu_p, \lambda_{p+1}, ..., \lambda_{2n}\}$ and the eigenvectors are $\{y_1, ..., y_p, x_{p+1}, ..., x_{2n})\}$.

The problem can be reformulated as follows:

$$\begin{gathered}
\text{find } \min \|K - \tilde{K}\|_F^2 + \|D - \tilde{D}\|_F^2 \\
\text{such that:} \\
\tilde{K} = \tilde{K}^T; \\
\tilde{D} = \tilde{D}^T; \\
M(\Lambda_1^*)^2 Y_1 + \tilde{D}(\Lambda_1^*)Y_1 + \tilde{K}Y_1 = 0
\end{gathered}$$

where $\Lambda_1^* = \text{diag}(\mu_1, ..., \mu_p)$, $Y_1 = [y_1, ..., y_p]$ are the desired matrices.
In order to simplify the problem, let us define the following matrices:

$$A = M(\Lambda_1^*)^2 Y_1; \qquad B = (\Lambda_1^*)Y_1; \qquad C = Y_1;$$

$$X = \begin{bmatrix} K & 0 \\ 0 & D \end{bmatrix}; \qquad \tilde{X} = \begin{bmatrix} \tilde{K} & 0 \\ 0 & \tilde{D} \end{bmatrix}.$$

The problem can be rewritten as:

$$\min \|X - \tilde{X}\|_F^2$$
$$\text{s.t. } \tilde{X} = \tilde{X}^T;$$
$$A + \tilde{X}_{22}B + \tilde{X}_{11}C = 0$$

Now, defining $\hat{I} = \begin{bmatrix} I_{n\times n} \\ I_{n\times m} \end{bmatrix}$ and $W = \begin{bmatrix} C \\ B \end{bmatrix}$ it holds:

$$A + \hat{I}^T\tilde{X}W = A + \tilde{K}C + \tilde{D}B = A + \tilde{X}_{22}B + \tilde{X}_{11}C$$

So the constraints become:

$$\tilde{X} = \tilde{X}^T \tag{2.3.1}$$
$$A + \hat{I}^T\tilde{X}W = 0 \tag{2.3.2}$$

We will project onto the subspace $\mathcal{S}$ of symmetric matrices, defined by constraint 2.3.1, whose projection is given by

$$P_{\mathcal{S}}(X) = \frac{X + X^T}{2};$$

and onto the linear variety $\mathcal{V} = \{x \in R^{2n\times 2n} : A + Z^T XW = 0\}$ whose projection is given by the following result.

Theorem 2.9 *If $X \in R^{2n\times 2n}$ is any given matrix, the the projection onto the linear variety V is given by $P_{\mathcal{V}}(X) = X + Z\Sigma W^T$, where Σ satisfies:*

$$W^T W\Sigma^T = -\frac{1}{2}(A^T + W^T X^T Z)$$

The solution can now be found using MAP on $\mathcal{S}$ and $\mathcal{V}$.

2.3.3 Projection Methods on Quantum Information Science

A natural problem in quantum information science is to construct, if it exists, a quantum operation sending a given set of quantum states $\{\rho_1, .., \rho_k\}$ to another set of quantum states $\{\bar{\rho}_1, ..., \bar{\rho}_k\}$. Quantum states are mathematically represented as density matrices (positive semi-definite, Hermitian matrices with unitary trace) while quantum operations are represented by trace preserving, completely positive maps (CPTP maps) T that maps $n \times n$ density matrices to $m \times m$ density matrices, having the form:

$$\mathrm{T}(X) = \sum_{j=1}^{r} F_j X F_j^\dagger$$

where it holds $\sum_{j=1}^{r} F_j^* F_j = I_n$. More details on quantum formalism will be given in Chapter 3.
Given some density matrices $A_1, ..., A_k$ and $B_1, ..., B_k$ our task is to find a CPTP map which satisfies $\mathrm{T}(A_i) = B_i$. If we denote with $E_{11}, E_{12}, ..., E_{nn}$ standard orthonormal basis, T is a CPTP map if and only if the Choi matrix $C(\mathrm{T})$

$$C(\mathrm{T}) = \begin{bmatrix} P_{11} & ... & P_{1n} \\ ... & P_{ij} & ... \\ P_{n1} & ... & P_{nn} \end{bmatrix} := \begin{bmatrix} \mathrm{T}(E_{11}) = & ... & \mathrm{T}(E_{1n}) \\ ... & \mathrm{T}(E_{ij}) & ... \\ \mathrm{T}(E_{n1}) & ... & \mathrm{T}(E_{nn}) \end{bmatrix}$$

is positive semi-definite and $\mathrm{tr}(P_{ij}) = \delta_{ij}$.
Our problem is equivalent to the positive semi-definite feasibility problem for $P = (P_{ij})$:

$$\begin{cases} \sum_{ij}(A_l)_{ij}P_{ij} = B_l & l = 1, ..., k \\ \mathrm{tr}(P_{ij}) = \delta_{ij} & 1 \leq i \leq j \leq n \\ P \in H_+^{nm} \end{cases}$$

It is easy to see that the first condition is true: we can write $A_l = \sum_{i,j}(A_l)_{i,i}E_{i,j}$. Then, by linearity of T, we have $\mathrm{T}(A_l) = B_l$ that is equivalent to $B_l = \mathrm{T}(A_l) = \mathrm{T}(\sum_{i,j}(A_l)_{i,i}E_{i,j}) = \sum_{i,j}(A_l)_{i,j}\mathrm{T}(E_{i,j}) = \sum_{i,j}(A_l)_{i,j}P_{i,j}$. Let us define:

$$\mathcal{L}_A(P) = (\textstyle\sum_{ij}(A_l)_{ij}P_{ij})_l;$$
$$\mathcal{L}_T(P) = (\mathrm{tr}(P_{ij}))_{i,j}$$
$$\mathcal{L}(P) = (\mathcal{L}_A(P), \mathcal{L}_T(P));$$
$$B = [B_1...B_k];$$
$$\Delta = (\delta_{ij})_{i,j}$$

We want to find a matrix P in the intersection of $\mathcal{H}_+^{nm}$ with the affine subspace $A = \{P : \mathcal{L}(P) = (B, \Delta)\}$.
If $P = U^\dagger \mathrm{diag}(\lambda_1, ..., \lambda_{mn})U$, then the projection onto $\mathcal{H}_+^{nm}$ is given by: $P_{\mathcal{H}_+^{nm}}(P) = U^\dagger \mathrm{diag}(\lambda_1^+, ..., \lambda_{mn}^+)U$, where $r^+ = \max\{0, r\}$.
The projection onto A is given by $P_A(P) = P + \mathcal{L}^\dagger R$ where $\mathcal{L}^\dagger$ is the Moore-Penrose general inverse while $R = (B, \Delta) - \mathcal{L}(P)$ is the residual. Using MAP we can find the solution to our problem. More details can be found in [7].

2.3.4 Fixed-Order Control Design for LMI Control problems using MAP

Consider a linear, TI, continuous dynamic system with the follow state-space representation:

$$\begin{aligned} \dot{x} &= Ax + B_1 w + B_2 u \\ z &= C_1 x + D_{11} w + D_{12} u \\ y &= C_2 x + D_{21} w \end{aligned}$$

Where $x(t) \in \mathbb{R}^{n_p}$ is the state, $u(t) \in \mathbb{R}^{n_u}$ is the control. $w(t) \in \mathbb{R}^{n_w}$ is an external input (i.e. noise), $z(t) \in \mathbb{R}^{n_z}$ is the regulated output and $y(t) \in \mathbb{R}^{n_y}$ is the measured output.
We seek a linear, TI, continuous time controller of order n_c with state-space representation:

$$\begin{aligned} \dot{x}_c &= A_c x_c + B_c y \\ u &= C_c x_c + D_c y \end{aligned}$$

We will consider $w(t) = 0$: *noise free stabilization problem.*

Theorem 2.10 *The following statements are equivalent:*

a) it exists a stabilizing dynamic output-feedback controller of order n_c;

b) there exist matrices $X > 0$, $Y > 0$ s.t.

$$\begin{aligned} (B_2)^{\perp}[AX + XA^T](B_2)^{\perp T} &< 0; \\ (C_2)^{T\perp}[YA + A^T Y](C_2)^{T\perp T} &< 0; \\ \begin{bmatrix} X & I \\ I & Y \end{bmatrix} &\geq 0; \\ \operatorname{rank} \begin{bmatrix} X & I \\ I & Y \end{bmatrix} &\leq n_p + n_c. \end{aligned}$$

Theorem 2.11 *The following statements are equivalent:*

a) it exists an H_∞ sub-optimal controller of order n_c;

b) there exist matrices $X > 0$, $Y > 0$ s.t.

$$\begin{aligned} \begin{bmatrix} B_2 \\ D_{12} \end{bmatrix}^{\perp} \begin{bmatrix} AX + XA^T + B_1 B_1^T & XC_1^T + B_1 D_{11}^T \\ C_1 X + D_{11} B_1^T & D_{11} D_{11}^T - I \end{bmatrix} \begin{bmatrix} B_2 \\ D_{12} \end{bmatrix}^{\perp T} &< 0; \\ \begin{bmatrix} C_2^T \\ D_{21}^T \end{bmatrix}^{\perp} \begin{bmatrix} YA + A^T Y + C_1^T C_1 & YB_1 + C_1^T D_{11} \\ B_1^T Y + D_{11}^T C_1 & D_{11}^T D_{11} - I \end{bmatrix} \begin{bmatrix} C_2^T \\ D_{21}^T \end{bmatrix} &< 0; \\ \begin{bmatrix} X & I \\ I & Y \end{bmatrix} &\geq 0; \\ \operatorname{rank} \begin{bmatrix} X & I \\ I & Y \end{bmatrix} &\leq n_p + n_c. \end{aligned}$$

For $n_c = n_p$ the relations of Theorems 2.10 and 2.11 are convex: it becomes a convex feasibility problem.
Now we need to find a projection formulation for the problem.
Let S_n be the set of real, symmetric $n \times n$ matrices equipped with the Frobenius norm and the inner product: $\langle x, y \rangle = \operatorname{tr}(xy)$.
Consider the set $\mathcal{L} = \{X \in S_n : EXF + F^T X E^T + Q < 0\}$ where $E, F, Q \in S_n$ are of compatible dimensions. $\mathcal{L}$ is a convex set of S_n. Every LMI constraint seen in the previous two theorems can be written as in $\mathcal{L}$(an example can be seen in [13]).
We will consider the closed ϵ-approximation of $\mathcal{L}$: $\mathcal{L}_\epsilon = \{X \in S_n : EXF + F^T X E^T + Q \leq \epsilon I\}$ with $\epsilon > 0$

Proposition 2.1 *Let us define the following sets in S_{2n}:*

$$\mathcal{J}_\epsilon \doteq \{W \in S_{2n} : \begin{bmatrix} E & F^T \end{bmatrix} W \begin{bmatrix} E^T \\ F \end{bmatrix} \leq -Q_\epsilon\}$$

$$\mathcal{T} \doteq \{W \in S_{2n} : W = \begin{bmatrix} W_{11} & W_{12} \\ W_{12}^T & W_{22} \end{bmatrix}, W_{11} = W_{22} = 0, W_{12} \in S_n\}$$

where $Q_\epsilon = Q + \epsilon I$. Then the following statements are equivalent:

a) $X \in \mathcal{L}_\epsilon$;

b) $X = W_{12}$

where $W \in \mathcal{J}_\epsilon \cap \mathcal{T}$

Proposition 2.2 *Let $W \in S_{2n}$. Consider the SVD:*

$$\begin{bmatrix} E & F^T \end{bmatrix} = U \begin{bmatrix} \Sigma & 0 \end{bmatrix} V^T$$

and define

$$\bar{W} \doteq V^T W V = \begin{bmatrix} \bar{W}_{11} & \bar{W}_{12} \\ \bar{W}_{12}^T & \bar{W}_{22} \end{bmatrix}$$

with $\bar{W}_{11} \in S_n$.
Consider the eigenvalue-eigenvector decomposition:

$$\bar{W}_{11} + \Sigma^{-1} U^T Q_\epsilon U \Sigma^{-1} = L \Lambda L^T$$

The projection $W^* = P_{J_\epsilon}(W)$ of W onto $\mathcal{J}_\epsilon$ is

$$W^* = V \begin{bmatrix} \bar{W}_{11}^* & \bar{W}_{12} \\ \bar{W}_{12}^T & \bar{W}_{22} \end{bmatrix} V^T$$

where $\bar{W}_{11}^* = L\Lambda_- L^T - \Sigma^{-1} U^T Q_\epsilon U \Sigma^{-1}$, Λ_- is the diagonal matrix obtained by replacing the positive eigenvalues of Λ by 0.

Proposition 2.3 *Let $W \in S_{2n}$. The orthogonal projection $W^* = P_{\mathcal{T}}(W)$ of W in $\mathcal{T}$ is*

$$W^* = \begin{bmatrix} 0 & X^* \\ X^* & 0 \end{bmatrix}$$

where $X^* = \frac{1}{2}(W_{12} + W_{12}^T)$
In addiction to the previous constraints, we need to derive the expression of the orthogonal projection onto the positivity and rank constraints sets. In order to achieve that, we define the following sets:

Definition 2.4 *Let us define:*

$$\begin{aligned}
\mathcal{D} &\doteq \{Z \in S_{2n} : Z = \begin{bmatrix} X & 0 \\ 0 & Y \end{bmatrix}, X, Y \in S_n\} \\
\mathcal{P} &\doteq \{Z \in S_{2n} : Z \geq -J\} \\
\mathcal{R} &\doteq \{Z \in S_{2n} : \operatorname{rank}(Z + J) \leq k\}; n \leq k \leq 2n \\
J &\doteq \begin{bmatrix} 0 & I_n \\ I_n & 0 \end{bmatrix} \in S_{2n}
\end{aligned}$$

We next provide the expression of the orthogonal projection onto the previous sets.

Proposition 2.4 *Let $Z = \begin{bmatrix} Z_{11} & Z_{12} \\ Z_{12}^T & Z_{22} \end{bmatrix} \in S_{2n}$. The projection $Z^* = P_{\mathcal{D}}(Z)$ of Z onto $\mathcal{D}$ is*

$$Z^* = \begin{bmatrix} Z_{11} & 0 \\ 0 & Z_{22} \end{bmatrix}$$

Proposition 2.5 *Let $Z \in S_n$ and $Z + J = L\Lambda L^T$. $Z^* = P_{\mathcal{P}}(Z)$, the projection of Z onto $\mathcal{P}$, is given by*

$$Z^* = L\Lambda_+ L^T - J$$

where Λ_+ is the diagonal matrix obtained by replacing the negative eigenvalues of Λ by zero.

Proposition 2.6 *$Z \in S_{2n}$ and $Z + J = U\Sigma V^T$. The projection $Z^* = P_{\mathcal{R}}(Z)$ of Z onto $\mathcal{R}$ is given by:*

$$Z^* = U\Sigma_k V^T - J$$

where Σ_k is the diagonal matrix obtained by replacing the $2n-k$ smallest singular values of $Z + J$ by zero.

Summarizing, we have decomposed the constraints onto simpler sets in order to reduce our problem into a feasibility problem and we have found the explicit expressions of the orthogonal projections onto each set. Now, applying alternating projection methods, we can find the desired solution. The original article with additional details can be found in [13].

Chapter 3

Iterated Quantum Maps

3.1 Statistical Description of Finite-Dimensional Quantum Systems

To every quantum physical system $\mathcal{Q}$ is associated a complex Hilbert space $\mathcal{H}$ whose dimension depends on the observable quantities we want to describe. If the system is finite-dimensional, namely the quantities of interest admit a finite set of outcomes, the Hilbert space is isomorphic to C^N. $\mathcal{H}$ is associated to the inner product:

$$\langle x, y \rangle = \sum_j x_j^* y_j = x^\dagger y.$$

where the x_j represent the components of $x \in \mathbb{C}^N$.

Postulate 3.1 *A physical quantity relative to the system of interest that can (in principle) be measured is called* ***observable****.*

In quantum mechanics any observable is associated to an Hermitian operator $A \in \mathfrak{h}(\mathcal{H})$ ($\mathfrak{h}(\mathcal{H})$ indicates the set of all hermitian operators in $\mathcal{H}$). The operator A can be written, by the spectral theorem, as $A = \sum_j a_j \Pi_j$ where a_j are the eigenvalues of A and Π_j the respective orthogonal projectors ($\Pi_j \Pi_k = \delta_{jk} \Pi_j$, $\sum_j \Pi_j = I$). The eigenvalues represent the possible outcomes of A and the projectors the quantum events.

Postulate 3.2 *A state of maximal information for the system is associated to a state vector $|\psi\rangle$, which is a norm-1 vector in $\mathcal{H}$.*

It is very difficult to know exactly the state of the system, more often there is some uncertainty. Let us suppose that $\mathcal{H}$ is composed of states $\{|\psi_j\rangle\}$ and let p_j be the corresponding probability of being in that state.

Definition 3.1 *The* ***density operator*** *for the system is defined by*

$$\rho = \sum_j p_j |\psi_j\rangle \langle \psi_j| ;$$

where $p_j \geq 0$ and $\sum_j p_j = 1$. They have, in quantum mechanics, the role of probability densities. The density operator is often called density matrix or, simply, state.

Hereafter we will consider only density operators that correspond to complex $N \times N$ matrices such that

$$\rho = \rho^\dagger; \qquad \text{tr}(\rho) = 1; \qquad \text{tr}(\rho^2) \leq 1.$$

If $p_1 = 1$, $\rho = |\psi_1\rangle \langle\psi_1|$. In this case ρ is called *pure state* and it has the same meaning of the state vector $|\psi\rangle$ provided before. In case of two or more $p_j > 0$, ρ is called *mixed state* and it cannot be described by a single state vector.

In case we want to calculate the probability P_{a_j} of observing the value a_j as an outcome of the observable A, we can calculate it as:

$$P_{a_j} = \text{tr}(\rho\Pi_j)$$

while the conditional density operator after the measurement of a_j is:

$$\rho_{A=a_j} = \frac{\Pi_j \rho \Pi_j}{\text{tr}(\Pi_j \rho \Pi_j)}$$

If we want to compute the expectation of A assuming that the state is ρ:

$$\mathbb{E}_\rho(A) = \text{tr}(A\rho) = \langle A, \rho\rangle_{H.S.}$$

where $\langle \cdot, \cdot\rangle_{H.S.}$ is an inner product for the space of operators in $\mathcal{H}$, the called *Hilbert-Schmidt inner product.*

Systems composed by different subsystems are described as tensor products of different subsystems, i.e. $\mathcal{H} = \mathcal{H}_A \otimes \mathcal{H}_B$.
Tensor product is a way to assemble vector space.

Definition 3.2 *Let V and W be Hilbert spaces of dimensions n and m respectively. Then the tensor product $V \otimes W$ is an Hilbert space of dimension nm which elements are linear combination of $|v\rangle \otimes |w\rangle$ where $|v\rangle \in V$ and $|w\rangle \in W$.*

A state $\rho \in \mathcal{H}_A \otimes \mathcal{H}_B$ that can be decomposed as $\rho = \rho_A \otimes \rho_B$ where $\rho_A \in \mathcal{H}_A$ and $\rho_B \in \mathcal{H}_B$ is called *uncorrelated.*
A state ρ that can be decomposed as $\rho = \sum_k \lambda_k \rho_A^k \otimes \rho_B^k$ is called *classically correlated.*
A state that cannot be decomposed as before is called *entangled.*

Proposition 3.1 ***(Schmidt decomposition)*** *For every $|\psi\rangle \in \mathcal{H} = \mathcal{H}_A \otimes \mathcal{H}_B$ there exists orthonormal bases $\{e_j \in \mathcal{H}_A\}$ and $\{f_j \in \mathcal{H}_B\}$ such that*

$$|\psi\rangle = \textstyle\sum_{j=1}^d \sqrt{\lambda_j}\, |e_j\rangle \otimes |f_j\rangle;$$

with $\lambda_j \geq 0$, $\sum_j \lambda_j = \| \, |\psi\rangle \, \|^2$ and $d = \min\{dim(\mathcal{H}_A), dim(\mathcal{H}_B)\}$. □

Proposition 3.2 ***Operator Schmidt decomposition*** *For every operator $\rho \in \mathcal{H} = \mathcal{H}_A \otimes \mathcal{H}_B$ there exist orthonormal bases $\{A_j \in \mathcal{H}_A\}$ and $\{B_j \in \mathcal{H}_B\}$ such that*

$$\rho = \sum_{j=1}^d \sqrt{\lambda_j} A_j \otimes B_j$$

with $\lambda_j \geq 0$, $\sum_j \lambda_j = \| \, |\psi\rangle \, \|^2$ and $d = \min\{dim(\mathcal{H}_A), dim(\mathcal{H}_B)\}$. □

If there are a composite system described by the density operator $\rho \in \mathcal{H}_A \otimes \mathcal{H}_B$, the reduced density operator for the subsystem A can be computed as:

$$\rho_A = \mathrm{tr}_B(\rho);$$

where tr_B is the *partial trace* over the system B.

Definition 3.3 *Let $\rho^{AB} = \rho \otimes \sigma \in \mathcal{H} = \mathcal{H}_A \otimes \mathcal{H}_B$, then the* ***partial trace*** *over B is defined as:*

$$\mathrm{tr}_B(\rho^{AB}) = \mathrm{tr}_B(\rho \otimes \sigma) = \rho \mathrm{tr}(\sigma);$$

By linearity, for general $\rho = \sum_{ij} c_{ij} X_i \otimes Y_j$ we define:

$$\mathrm{tr}_B(\rho) = \sum_{ij} c_{ij} \mathrm{tr}(Y_j) X_i;$$

More details about tensor product and partial trace can be found in Appendix A. More details about statistical description in quantum systems can be found in [2],[19],[23].

3.2 Open System Dynamics

Quantum dynamics are typically studied in two different scenarios: **Closed systems** and **Open systems**.
A system is called open if it has non trivial interaction with the environment to which it belongs, while closed systems are isolated.

Postulate 3.3 *The state vector $|\psi\rangle$ of a closed quantum system obeys to the Schrödinger equation:*

$$\begin{cases} \hbar|\dot{\psi}\rangle = -iH_0 |\psi\rangle \\ |\psi(0)\rangle = |\psi_0\rangle \end{cases};$$

where H_0 is the Hamiltonian of the system which, like classical mechanic, depends by the energy of the system, and $\hbar$ is the Planck constant (we will consider $\hbar = 1$).

Quantum states belong to complex sphere $\mathbb{S}^{2N-1}$ which can be lifted to the Lie group $SU(N) = \{U \in \mathbb{C}^{N \times N} : U^\dagger U = UU^\dagger = I \text{ , } \det(U) = 1\}$.
So the *Schrödinger equations for the unitary propagator* can be obtained:

$$\begin{cases} \dot{U} = -iH_0 U \\ U(0) = I \end{cases};$$

where $U \in SU(N)$. The solution for the state is given by:

$$|\psi(t)\rangle = U(t) |\psi_0\rangle ;$$

where

$$U(t) = e^{-iH_0 t};$$

Therefore for any pure state we obtain:

$$|\psi_j(t)\rangle \langle\psi_j(t)| = U(t) |\psi_j(0)\rangle \langle\psi_j(0)| U^\dagger(t).$$

This implies, by linearity, that for a general state $\rho(t)$ we have, in terms of unitary propagator:

$$\rho(t) = U(t)\rho(0)U^{\dagger}(t);$$

A real physical quantum system, in general, have interaction with the environment in which it belongs. Le us consider a finite-dimensional quantum system $\mathcal{S}$ coupled to the environment $\mathcal{E}$, chosen so that $\mathcal{S}$ and $\mathcal{E}$ together can be considered isolated, and let $\mathcal{H}_{\mathcal{S}}$ and $\mathcal{H}_{\mathcal{E}}$ be the system and environment Hilbert spaces, with $\dim(\mathcal{H}_{\mathcal{S}}) = N < \infty$. The total Hamiltonian for the composite system is given by:

$$H_{tot} = H_{\mathcal{S}} \otimes I_{\mathcal{E}} + I_{\mathcal{S}} \otimes H_{\mathcal{E}} + H_{\mathcal{S}\mathcal{E}}; \tag{3.2.1}$$

where $H_{\mathcal{S}}$,$H_{\mathcal{E}}$,$H_{\mathcal{S}\mathcal{E}}$ are the system, environment and interaction Hamiltonian, respectively. On the joint space $\mathcal{H}_{\mathcal{S}} \otimes \mathcal{H}_{\mathcal{E}}$ the dynamics is unitary by Postulate 3.3, since $\mathcal{S} \otimes \mathcal{E}$ is isolated by construction. Let us assume that the initial state is $\rho_{\mathcal{S}\mathcal{E},0} = \rho_0 \otimes \rho_{\mathcal{E}}$. By previous considerations we have:

$$\rho_{\mathcal{S}\mathcal{E},t} = U_{\mathcal{S}\mathcal{E},t}(\rho_0 \otimes \rho_{\mathcal{E}})U^{\dagger}_{\mathcal{S}\mathcal{E},t};$$

where $U_{\mathcal{S}\mathcal{E},t} = e^{-iH_{tot}t}$. In order to obtain the state of the system we need the partial trace which gives:

$$\rho_{\mathcal{S},t} = \mathcal{T}(\rho_0) = \mathrm{tr}_{\mathcal{E}}(U_{\mathcal{S}\mathcal{E},t}(\rho_0 \otimes \rho_{\mathcal{E}})U^{\dagger}_{\mathcal{S}\mathcal{E},t}). \tag{3.2.2}$$

Definition 3.4 *A map $\mathcal{E}(\cdot)$ is a **Completely Positive(CP)** map if for every R being an auxiliary system of arbitrary, finite dimension, $(\mathcal{I} \otimes \mathcal{E})(A) \geq 0$ for every operator $A \geq 0$ on the combined system $R \otimes \mathcal{H}$, where $\mathcal{I}$ denotes the identity map on $\mathfrak{h}(R)$.*
Clearly CP implies positivity when $\dim(R) = 0$.

Definition 3.5 *A map $\mathcal{E}(\cdot)$ is said **Trace Preserving(TP)** if $\mathrm{tr}(\mathcal{E}(A)) = \mathrm{tr}(A)$.*

The map (3.2.2) is a CPTP map [2],[19],[21].
Note. CPTP maps are also called *quantum channels* or *Kraus maps.*

Properties:

- CPTP map can be written as

$$\mathcal{E}(\rho) = \sum_k M_k \rho M_k^{\dagger}$$

where M_k is such that $\sum_k M_k M_k^{\dagger} = I$;

- Any CPTP map is non-expansive: let ρ and σ be two states, then

$$\|\mathcal{E}(\rho) - \mathcal{E}(\sigma) \leq \|\rho - \sigma\|$$

where $\|A\| = \mathrm{tr}(\sqrt{A^{\dagger}A})$ is the trace norm;

- Since a CPTP maps the state of density operators in itself and is a contraction, by (Brouwer's) fixed point theorem it admits at least a fixed state ρ_0 such that $\mathcal{E}(\rho_0) = \rho_0$.

3.3 CPTP Projections

Let us consider a CPTP map $\mathcal{E}$ such that $\mathcal{E}^2 = \mathcal{E}$, which we shall call a CPTP projection. Then it maps any operator X onto the set of fixed points $Fix(\mathcal{E}) = \{X \in \mathcal{H} : \mathcal{E}(X) = X\}$. It has been proved [4] that the fixed points of a CPTP map form an algebra with respect to a weighed product what has been called a *distorted algebra* [10]. More precisely, for any CPTP map there exists a decomposition of the Hilbert space $\mathcal{H}$ such that:

$$\mathcal{H} = (\bigoplus_i \mathcal{H}_{Si} \otimes \mathcal{H}_{Fi}) \oplus \mathcal{H}_R;$$

By that it results that:

$$Fix(\mathcal{E}) = (\bigoplus_i \mathcal{B}(\mathcal{H}_{Si}) \otimes \tau_i) \oplus \emptyset_R;$$

where $\emptyset_R$ is the null operator on $\mathcal{H}_R$.
We will consider, for sake of simplicity, only maps that admit at least one full rank state, so the decomposition is given by:

$$\mathcal{H} = \bigoplus_i \mathcal{H}_{Si} \otimes \mathcal{H}_{Fi}; \tag{3.3.1}$$

and

$$Fix(\mathcal{E}) = \bigoplus_i \mathcal{B}(\mathcal{H}_{Si}) \otimes \tau_i; \tag{3.3.2}$$

A CPTP projection is then a CPTP map $\mathcal{E}$ that projects a state ρ into the set $Fix(\mathcal{E})$. We next provide its structure with respect to $Fix(\mathcal{E}$ as in 3.3.2.
In general:

Proposition 3.3 *If $\hat{\mathcal{E}}^2 = \hat{\mathcal{E}}$ and there exists a full rank $\bar{\rho}$ such that $\mathcal{E}(\bar{\rho}) = \bar{\rho}$, then*

$$\hat{\mathcal{E}}(\rho) = \bigoplus_i \mathrm{tr}_{Fi}(\Pi_{SFi}\rho\Pi_{SFi}) \otimes \tau_i; \tag{3.3.3}$$

where Π_{SFi} is the projection onto $\mathcal{H}_{Si} \otimes \mathcal{H}_{Fi}$ as in 3.3.1.
It easy to see that $\hat{\mathcal{E}}(\rho) \in Fix(\hat{\mathcal{E}})$.
Let us call $\mathcal{A} = \bigoplus_i \mathcal{A}_i = \bigoplus_i \mathcal{B}(\mathcal{H}_{Si}) \otimes \tau_i$. The orthogonal projection of $\rho \in \mathcal{H}$ onto $\mathcal{A}$ is given, by definition (1.9):

$$\rho_{\mathcal{A}} = \sum_{l,j} \langle \sigma_l \otimes \tau_j, \rho \rangle_{HS} \sigma_l \otimes \tau_j;$$

where $\sigma_l \otimes \tau_j$ is an orthonormal basis for $\mathcal{A}_i$. Decomposing $\rho = \sum_k A_k \otimes B_k$ we have:

$$\begin{aligned}
\rho_{\mathcal{A}} &= \bigoplus_i \sum_{l,j} (\sum_k \mathrm{tr}[(\sigma_l \otimes \tau_j)(A_k \otimes B_k)] \sigma_l \otimes \tau_j) \\
&= \bigoplus_i \sum_{l,j} (\sum_k \mathrm{tr}(\sigma_l A_k) \mathrm{tr}(\tau_j B_k) \sigma_l \otimes \tau_j) \\
&= \bigoplus_i \sum_{l,j} (\mathrm{tr}[\sigma_l \sum_k (A_k \mathrm{tr}(\tau_j B_k))] \sigma_l \otimes \tau_j))
\end{aligned}$$

Note: $\sum_k (A_k \mathrm{tr}(\tau_j B_k)) = \mathrm{tr}_{Fi}(\rho) \iff \tau_j = I$ so, in general, $\hat{\mathcal{E}}(\rho)$ of 3.3.3 is not an orthogonal projection with respect to Hilbert Schmidt inner product.
In order to obtain an orthogonal projection we can define a new inner product.

Definition 3.6 *Let ξ be a positive definite operator. We define the* ***modified ξ-inner product*** *as:*

$$\langle X, Y\rangle_\xi = \mathrm{tr}(X\xi Y); \tag{3.3.4}$$

We define the ***modified symmetric ξ-inner product*** *as:*

$$\langle X, Y\rangle_{\xi,s} = \mathrm{tr}(X\xi^{\frac{1}{2}}Y\xi^{\frac{1}{2}}). \tag{3.3.5}$$

Proposition 3.4 *(3.3.4) is a valid inner product.*
Proof. We have to show that $\langle\cdot,\cdot\rangle_\xi$ satisfies the properties of definition (1.4):

1. $\langle X, X\rangle_\xi = \mathrm{tr}(X\xi X) = \mathrm{tr}(\xi X^2) \geq 0$ clearly $= 0 \iff X = 0$;

2. $\langle X, Y\rangle_\xi = \mathrm{tr}(X\xi Y) = \mathrm{tr}(Y^\dagger\xi^\dagger X^\dagger)^\dagger = \mathrm{tr}(Y\xi X)^* = \langle Y, X\rangle_\xi^*$;

3. $\langle\alpha_1 X + \alpha_2 Y, Z\rangle_\xi = \mathrm{tr}((\alpha_1 X + \alpha_2 Y)\xi Z) = \alpha_1 \mathrm{tr}(X\xi Z) + \alpha_2 \mathrm{tr}(Y\xi Z) = \alpha_1\langle X, Z\rangle_\xi + \alpha_2\langle Y, Z\rangle_\xi$.

Similarly for (3.3.5). □

Note. Changing the inner product for the Hilbert space ($\langle\cdot,\cdot\rangle_{HS} \to \langle\cdot,\cdot\rangle_{\xi,s}$) is equivalent to a change of measure in a classical probability space.
In fact it holds:

$$\mathbb{E}_\rho(X) = \langle\rho, X\rangle_{HS}$$

$$\mathbb{E}_{\tilde{\rho}}(X) = \langle\rho, X\rangle_{\xi,s}$$

where $\tilde{\rho} = \xi^{-\frac{1}{2}}\rho\xi^{-\frac{1}{2}}$ is the "new" unnormalized state.

In order to show that $\mathcal{E}$ is an orthogonal projection with reference to (3.3.4), we will need the following lemma. With $W = \bigoplus W_i$ we will denote an operator that acts as W_i on $\mathcal{H}_i$ for a decomposition of $\mathcal{H} = \bigoplus_i \mathcal{H}_i$.

Lemma 3.1 *Let $W = \bigoplus W_i$ and let Y be an operator. Then $\mathrm{tr}(WY) = \sum_i \mathrm{tr}(W_i Y_i)$, where $Y_i = \Pi_i Y \Pi_i$*
Proof. Let Π_i be the projector onto $\mathcal{H}_i$. Remembering that $\bigoplus_i \Pi_i = I$ and $\Pi_i = \Pi_i^2$, it holds:

$$\begin{aligned}
\mathrm{tr}(X) &= \mathrm{tr}(\sum_i \Pi_i X) \\
&= \sum_i \mathrm{tr}(\Pi_i X) \\
&= \sum_i \mathrm{tr}(\Pi_i^2 X) \\
&= \sum_i \mathrm{tr}(\Pi_i X \Pi_i)
\end{aligned}$$

Therefore we obtain:

$$\begin{aligned}
\mathrm{tr}(WY) &= \mathrm{tr}(\bigoplus_j W_j Y) \\
&= \mathrm{tr}((\sum_i \Pi_i)(\bigoplus_j W_j)Y) \\
&= \sum_i \mathrm{tr}(\Pi_i \bigoplus_j W_j Y) \\
&= \sum_i \mathrm{tr}(\Pi_i W_i Y) \\
&= \sum_i \mathrm{tr}(\Pi_i W_i \Pi_i Y) \\
&= \sum_i \mathrm{tr}(W_i \Pi_i Y \Pi_i) \\
&= \sum_i \mathrm{tr}(W_i Y_i)
\end{aligned}$$

□

Proposition 3.5 *Let $\xi = \rho^{-1}$, where ρ is a full-rank fixed state in $\mathcal{H} = \bigoplus_i \mathcal{H}_{Si} \otimes \mathcal{H}_{Fi}$. Then (3.3.3) is an orthogonal projection with the reference to the modified inner product (3.3.4).*

Proof. We already know that $\mathcal{E}$ is linear and idempotent. In order to show that $\mathcal{E}$ is an orthogonal projection we need to show that it is self-adjoint (Definition 1.9).

Let us decompose X, Y and ρ as:

$$X = \bigoplus_i X_i;$$
$$Y = \bigoplus_i Y_I;$$
$$\rho = \bigoplus \rho_i \otimes \tau_i$$
where
$$\Pi_i X \Pi_i = X_i = \sum_k A_{k,i} \otimes B_{k,i}$$
$$\Pi_i Y \Pi_i = Y_i = \sum_l C_{l,i} \otimes D_{l,i}$$

We can consider

$$W = \mathcal{E}(X)\rho^{-1} = \bigoplus_i ([\mathrm{tr}_{Fi}(X_i) \otimes \tau_i](\rho_i^{-1} \otimes \tau_i^{-1})) = \bigoplus_i W_i$$

and by Lemma (3.1):

$$\begin{aligned}
< \mathcal{E}(X), Y >_\xi &= \mathrm{tr}(\mathcal{E}(X)\rho^{-1}Y) \\
&= \mathrm{tr}(\bigoplus_i \mathrm{tr}_{Fi}(X_i) \otimes \tau_i(\rho_i^{-1} \otimes \tau_i^{-1})Y_i) \\
&= \sum_{i,k} \mathrm{tr}(A_{k,i}\mathrm{tr}(B_{k,i}) \otimes \tau_i(\rho_i^{-1} \otimes \tau_i^{-1})Y_i) \\
&= \sum_{i,k} \mathrm{tr}((A_{k,i}\mathrm{tr}(B_{k,i})\rho_i^{-1} \otimes I)Y_i) \\
&= \sum_{i,k,l} \mathrm{tr}([A_{k,i}\mathrm{tr}(B_{k,i})\rho_i^{-1} \otimes I][C_{l,i} \otimes D_{l,i}]) \\
&= \sum_{i,k,l} \mathrm{tr}(A_{k,i}\mathrm{tr}(B_{k,i})\rho_i^{-1}C_{l,i} \otimes D_{l,i}) \\
&= \sum_{i,k,l} \mathrm{tr}(B_{k,i})\mathrm{tr}(A_{k,i}\rho_i^{-1}C_{l,i})\mathrm{tr}(D_{l,i})
\end{aligned}$$

By similar calculation:

$$\begin{aligned}
< X, \mathcal{E}(Y) >_\xi &= \mathrm{tr}(X\rho^{-1}\mathcal{E}(Y)) \\
&= \mathrm{tr}(X_i(\rho_i^{-1} \otimes \tau_i^{-1})\mathrm{tr}_{Fi}(Y_i) \otimes \tau_i) \\
&= \sum_{i,l} \mathrm{tr}(X_i\rho_i^{-1}C_{l,i}\mathrm{tr}(D_{l,i}) \otimes I) \\
&= \sum_{i,k,l} \mathrm{tr}([A_{k,i} \otimes B_{k,i}][\rho_i^{-1}C_{l,i}\mathrm{tr}(D_{l,i}) \otimes I]) \\
&= \sum_{i,k,l} \mathrm{tr}(B_{k,i})\mathrm{tr}(A_{k,i}\rho_i^{-1}C_{l,i})\mathrm{tr}(D_{l,i})
\end{aligned}$$

So, $< \mathcal{E}(X), Y >_\xi = < X, \mathcal{E}(Y) >_\xi$ wich proves self-adjointness. $\square$

The same properties still hold if we define the modified symmetric ξ-inner product (3.3.5).

3.4 Iterated CPTP Map Theorem

Thanks to the previous results we can now apply alternating projections theorem to iterated CPTP maps.

Theorem 3.1 *Let $\mathcal{H}$ be an Hilbert space and $\hat{\mathcal{E}}_1,...,\hat{\mathcal{E}}_r$ maps that project onto $Fix(\hat{\mathcal{E}}_i) \subset \mathcal{H}$, $i = 1, ..., r$ and let $Fix(\hat{\mathcal{E}}) = \bigcap_{i=1}^r Fix(\hat{\mathcal{E}}_i) \neq \emptyset$. If there exists a full-rank state $\rho_0 \in Fix(\mathcal{E})$, then $\forall x \in \mathcal{H}$:*

$$\lim_{n\to\infty} (\hat{\mathcal{E}}_r...\hat{\mathcal{E}}_1)^n x = \hat{\mathcal{E}}x$$

where $\hat{\mathcal{E}}$ is the projection onto $Fix(\hat{\mathcal{E}}) = \bigcap_{i=1}^r Fix(\hat{\mathcal{E}}_i)$

Proof. Let us consider $\xi = \rho_0^{-1}$, then $\rho_0 \in Fix(\hat{\mathcal{E}})$ implies that the maps $\hat{\mathcal{E}}_i$ are all orthogonal projection with respect to the modified inner product (Propositions 3.3, 3.5).

By Halperin classical theorem (Theorem 2.2) the limit of the cyclic orthogonal projections onto subsets converges to the projection onto the intersection of the subsets; by that, because $\hat{\mathcal{E}}_i$ is the orthogonal projection onto $Fix(\hat{\mathcal{E}}_i)$, the cyclic projections converge to the projection of x onto the intersection of the subsets, that is $Fix(\hat{\mathcal{E}})$, which proves the theorem. $\square$

3.5 Quantum Application

In this Section we will focus on open quantum systems composed by a finite number n of distinguishable systems, defined on

$$\mathcal{H} = \bigotimes_{a=1}^{n} \mathcal{H}_a, \qquad \dim(\mathcal{H}_a) = d_a, \qquad \dim(\mathcal{H}) = D$$

We are interested in studying dynamics in the case of quasi-locality constraints which specify a list of neighborhood: groups of subsystems on which operators can act simultaneously. Our goal is to understand if a given full rank state $\rho \in \mathcal{D}(\mathcal{H})$ can be prepared or, equally, it exists a dynamic for which ρ is *global asymptotically stable*(GAS).

We will next indicate with $\mathcal{N}_j$ the list of subsystems' indicies that compose th $j-$th neighborhood. By definition, a neighborhood induce a bipartite structure of $\mathcal{H}$:

$$\mathcal{H} = \mathcal{H}_{\mathcal{N}_j} \otimes \mathcal{H}_{\overline{\mathcal{N}}_j}$$

By this structure we can define the *reduced neighborhood state* ρ_j:

$$\rho_j = \mathrm{tr}_{\overline{\mathcal{N}}_j}(\rho).$$

Definition 3.7 *An operator M_j is a* neighborhood operator *if its action is non trivial only on $\mathcal{N}_j$. It can be decomposed as:*

$$M_j = M_{\mathcal{N}_j} \otimes I_{\overline{\mathcal{N}}_j};$$

where I is the identity operator for the complement of $\mathcal{N}_j$.

Definition 3.8 *A CPTP map $\mathcal{E}(\cdot)$ is said* Quasi local(QL) *if it can be written as:*

$$\mathcal{E}(\rho) = \sum_k M_k \rho M_k^\dagger;$$

where M_k is a neighborhood operator for the same $\mathcal{N}_j$. These maps are called neighborhood maps.

By this definition $\mathcal{E}(\cdot)$ can be decomposed as:

$$\mathcal{E}(\cdot) = \mathcal{E}_{\mathcal{N}_k}(\cdot) \otimes Id_{\overline{\mathcal{N}}_k}(\cdot).$$

Now we can describe our dynamics of interest:

i) It must exist a sequence of p CPTP maps, $p \to \infty$, such that $\mathcal{E}_p \circ ...\mathcal{E}_1(\rho) = \rho_d$ where ρ_d is the state to be prepared and ρ is any state in $\mathcal{H}$.

ii) For every $t = 1, ..., p$ it must hold: $\mathcal{E}_t(\rho_d) = \rho_d$;

iii) For every $t = 1, ..., p$ it exists a neighborhood $\mathcal{N}_j(t)$ such that $\mathcal{E}_t(\cdot) = \mathcal{E}_{\mathcal{N}_j(t)} \otimes Id_{\overline{\mathcal{N}}_j(t)}$

A dynamics that satisfies conditions i), ii), iii) is a *Quasi-local stabilizing dynamics.*

Let us recall that given an $X \in \mathcal{B}(\mathcal{H}_A \otimes \mathcal{H}_B)$, we can write its operator Schmidt decomposition (Proposition 3.2) as:

$$X = \sum_j A_j \otimes B_j$$

By this notion we state the following definition:

Definition 3.9 *Given an $X \in \mathcal{B}(\mathcal{H}_A \otimes \mathcal{H}_B)$, we define the* Schmidt span *as:*

$$\Sigma_A(X) = \text{span}(\{A_j\})$$

The Schmidt span is important because, if we want to leave an operator invariant with a neighborhood map, it also imposes the invariance of its Schmidt span.

Lemma 3.2 *Given a $\rho \in \mathcal{D}(\mathcal{H} = \mathcal{H}_{\mathcal{N}_|} \otimes \mathcal{H}_{\overline{\mathcal{N}}_|})$ and $\mathcal{E} = \mathcal{E}_{\mathcal{N}_j} \otimes I_{\overline{\mathcal{N}}_j} \in \mathcal{B}(\mathcal{H})$, it holds:*

$$\text{span}(\rho) \subseteq Fix(\mathcal{E}_{\mathcal{N}_j} \otimes I_{\overline{\mathcal{N}}_j}) \Rightarrow \Sigma_{\mathcal{N}_j} \otimes \mathcal{B}(\mathcal{H}_{\overline{\mathcal{N}}_j}) \subseteq Fix(\mathcal{E}_{\mathcal{N}_j} \otimes I_{\overline{\mathcal{N}}_j})$$

Given a $\rho > 0 \in Fix(\mathcal{E})$ the following properties hold:

I. $Fix(\mathcal{E})$ is a $*$-algebra w.r.t. the modified product $A \times_{\rho^{-1}} B = A\rho^{-1}B$;

II. $Fix(\mathcal{E})$ is invariant for the modular product $\Delta_\rho(X) = \rho^{\frac{1}{2}} X \rho^{-\frac{1}{2}}$.

In general $\Sigma_{\mathcal{N}_k} \subseteq Fix(\mathcal{E}_{\mathcal{N}_k})$; we need to enlarge $\Sigma_{\mathcal{N}_k}$ in order to satisfy properties I.,II.

Definition 3.10 *Let $\rho \in \mathcal{D}(\mathcal{H})$ and $W \in \mathcal{B}(\mathcal{H})$. The* minimal modular-invariant distorted algebra generated by W *is the smallest ρ-distorted algebra generated by W which is invariant w.r.t $\Delta_\rho(\cdot)$. In our case $W = \Sigma_{\mathcal{N}_k}(\rho)$ and we will call $\mathcal{F}_{\rho_{\mathcal{N}_k}}$ the minimal $\Delta_{\rho_{\mathcal{N}_k}}$ invariant, distorted algebra w.r.t. $\rho_{\mathcal{N}_k}$ modified product.*

Note: $\mathcal{F}_{\rho_{\mathcal{N}_k}}$ can be constructed iteratively starting from $\mathcal{F}^0 = alg_{\rho_{\mathcal{N}_k}}(\Sigma_{\mathcal{N}_k}(\rho))$ with k-th step given by:

$$\mathcal{F}^{k+1} = alg_{\rho_{\mathcal{N}_k}}(\Delta_{\rho_{\mathcal{N}_k}}(\Sigma_{\mathcal{N}_k}(\rho)))$$

It runs until $\mathcal{F}^{k+1} = \mathcal{F}^k = \mathcal{F}_{\rho_{\mathcal{N}_k}}$.

We next present a key result by Takesaki (in a finite-dimensional version given by Petz [21]).

Theorem 3.2 *Let $\mathcal{A}$ be a $\dagger$-closed subalgebra of $\mathcal{B}(\mathcal{H})$, and ρ a full rank state. Then the following are equivalent:*

(i) There exists a unital CP map $\mathcal{E}^\dagger$ such that $Fix(\mathcal{E}^\dagger) = \mathcal{A}$, $(\mathcal{E}^\dagger)^2 = \mathcal{E}^\dagger$ and $\mathcal{E}(\rho) = \rho$;

(ii) $\mathcal{A}$ is invariant w.r.t. $\Delta_\rho(\cdot)$.

The previous theorem allows to provide a characterization of distorted algebras that contain a given full-rank fixed state and are fixed point of some CPTP map.

Theorem 3.3 *Let ρ be a full-rank state. A distorted algebra $\mathcal{A}_\rho$ admits a CPTP map $\mathcal{E}(\cdot)$ such that $Fix(\mathcal{E}) = \mathcal{A}_\rho$ if and only if it is invariant for Δ_ρ.*

In our case, by the two previous theorems, it exists a CPTP map $\mathcal{E}_{\mathcal{N}_k}$ such that $Fix(\mathcal{E}_{\mathcal{N}_k}) = \mathcal{F}_{\rho_{\mathcal{N}_k}}$ and $\mathcal{E}^2_{\mathcal{N}_k} = \mathcal{E}_{\mathcal{N}_k}$. For $\mathcal{E}_k = \mathcal{E}_{\mathcal{N}_k} \otimes Id_{\overline{\mathcal{N}}_k}$ it holds that:

$$Fix(\mathcal{E}_k) = \mathcal{F}_{\rho_{\mathcal{N}_k}} \otimes \mathcal{B}(\mathcal{H}_{\overline{\mathcal{N}}_k}) = \mathcal{F}_k$$

By Takesaki's theorem and its extension to CPTP maps (theorems 3.2,3.3), it is clear what we have done: we need to find a QL stabilizing map; by Takesaki's theorem we know it exists always a CPTP map whose fixed point set is a modular invariant distorted algebra. Thanks to this information, we constructed those algebras, for neighborhood maps, by finding the minimal modular-invariant algebras. Thanks to these steps, we can state the main result of the section:

Proposition 3.6 *Let ρ_d be a full-rank state. There exists a Quasi-local stabilizing dynamics satisfying i), ii), iii) if and only if*

$$\operatorname{span}(\rho_d) = \bigcap_k \mathcal{F}_k$$

Proof. ($\Rightarrow$) By contradiction, let us suppose it exists $\rho_2 \in \bigcap_k \mathcal{F}_k$ such that $\rho_2 \neq \rho_d$. It clearly implies that ρ cannot be GAS because there exist an other state invariant for some maps, which implies that not all the maps "guide" to ρ_d.
($\Leftarrow$) By Theorems 3.2, 3.3 and what we have seen in Section 3.3, we can see that $\mathcal{E}_k$ is a CPTP projection. By Theorem 3.1, we already know that for every ρ, $(\mathcal{E}_q...\mathcal{E}_1)^k(\rho) \to \bigcap_k \mathcal{F}_k$ for $k \to \infty$. Now, by hypothesis, $\bigcap_k \mathcal{F}_k = \operatorname{span}(\rho_d)$ and, being ρ_d the only state in his span, we can conclude that ρ_d is GAS. $\square$

3.5.1 Example

A Gibbs state is defined as:

$$\rho_\beta = \frac{e^{\beta H}}{\operatorname{tr}(e^{\beta H})}, \qquad H = \sum_j H_j, \qquad \beta \in \mathbb{R}^+,$$

where each H_j is a neighborhood operator relative to $\mathcal{N}_j$. ρ is called *commuting Gibbs state* if it holds $[H_j, H_k]$ for all j, k.
Every 1D lattice can be associated to indexed subsystems (i.e. $i = 1, 2, 3, ...$). On those systems, we define the neighborhood structure *nearest neighborhood* (NN) as the one where $\mathcal{N}_i = \{i, i+1\}$, and the *next nearest neighbor* (NNN) as the one given by $\mathcal{N}_i = \{i, i+1, i+2\}$. We will solve our examples thanks the following proposition [15].

Proposition 3.7 *Gibbs states of 1D NN commuting Hamiltonians can be prepared with Quasi-local stabilizing dynamics acting on the NNN neighborhood structure (see Figure 3.5.1).*

We will solve our example following the proof of the previous proposition, which can be found in [15].

Let see our first example in the commuting case.
Let us consider the following Hamiltonian:

$$H = \sum_i \sigma_z^{(i)} \otimes \sigma_z^{(i+1)},$$

where σ_z is the Pauli matrix $\begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$. The Gibbs state is given by:

$$\rho = e^{\beta H} = \sum_i \sigma_{i,i+1},$$

where $\sigma_{i,i+1} = e^{\beta_{i,i+1}\sigma_z^{(i)} \otimes \sigma_z^{(i+1)}}$. For sake of simplicity, we will consider only the case of 4 subsystems. We need to find the minimal modular-invariant algebras for the subsystems $\{1, 2, 3\}$ and $\{2, 3, 4\}$. First, we need to find $\Sigma_{123}(\rho)$ which is given by:

$$\Sigma_{123}(\rho) = \sigma_{12}\sigma_{23}[I_{12} \otimes \Sigma_3(\sigma_{34})];$$

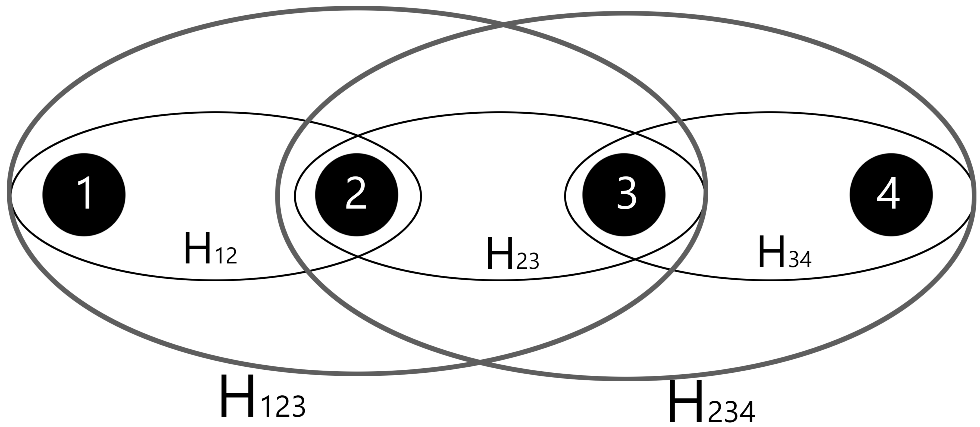

Figure 3.1: A visual example of a 1D lattice system with NN and NNN neighborhood structure .

Now, by definition, it holds:

$$
\begin{aligned}
\Sigma_3(\sigma_{34}) &= \operatorname{span}\{\operatorname{tr}_4([I_3 \otimes B]\sigma_{34}), \quad \forall B \in \mathcal{B}(\mathcal{H}_4)\} \\
&= \operatorname{span}\{\operatorname{tr}_4\Big(\begin{bmatrix} e^{\beta} & 0 & 0 & 0 \\ 0 & \frac{1}{e^{\beta}} & 0 & 0 \\ 0 & 0 & \frac{1}{e^{\beta}} & 0 \\ 0 & 0 & 0 & e^{\beta} \end{bmatrix} \begin{bmatrix} B & 0_{2\times 2} \\ 0_{2\times 2} & B \end{bmatrix}\Big)\} \\
&= \operatorname{span}\{\operatorname{tr}_4\Big(\begin{bmatrix} Be^{\beta\sigma_z} & 0_{2\times 2} \\ 0_{2\times 2} & Be^{-\beta\sigma_z} \end{bmatrix}\Big)\} \\
&= \operatorname{span}\{\begin{bmatrix} \operatorname{tr}(Be^{\beta\sigma_z}) & 0_{2\times 2} \\ 0_{2\times 2} & \operatorname{tr}(Be^{-\beta\sigma_z}) \end{bmatrix}\} \\
&= \operatorname{span}\{I, \sigma_z\}
\end{aligned}
$$

By that, we obtain $\Sigma_{123}(\rho) = \sigma_{12}\sigma_{23}[I_{12} \otimes \operatorname{span}\{I, \sigma_z\}]$. By symmetry, the same result can be obtained for $\Sigma_{234}(\rho) = \sigma_{23}\sigma_{34}[\operatorname{span}\{I, \sigma_z\} \otimes I_{34}]$. Now, it is evident that Σ_{123} and Σ_{234} are invariant for the distorted and the modular product, so we can immediately write:

$$
\begin{aligned}
\mathcal{F}_{123} &= \sigma_{12}\sigma_{23}[I_{12} \otimes \operatorname{span}\{I, \sigma_z\}] \otimes \mathcal{B}(\mathcal{H}_4) \\
\mathcal{F}_{234} &= \sigma_{23}\sigma_{34}[\operatorname{span}\{I, \sigma_z\} \otimes I_{34}] \otimes \mathcal{B}(\mathcal{H}_1)
\end{aligned}
$$

We need to find $\mathcal{F}_{123} \cap \mathcal{F}_{234}$. First of all, notice that $\sigma_{12} = \hat{\sigma}_{12} \otimes I_{34}$ (analogously for σ_{34}). Consider τ_{123} and τ_{234} operators in $\mathcal{F}_{123}$, $\mathcal{F}_{234}$ respectively. They can be expressed as following:

$$\begin{aligned} \tau_{123} &= \sigma_{23} \sum_i \hat{\sigma}_{12} \otimes D_{3i} \otimes W_{4i} \\ \tau_{234} &= \sigma_{23} \sum_j V_j \otimes C_{2j} \otimes \hat{\sigma}_{34} \end{aligned}$$

To find the intersection we need to impose $\tau_{123} = \tau_{234}$ and it is possible if and only if

$$\begin{cases} \sum_i D_{3i} \otimes W_{4i} = \hat{\sigma}_{34} \\ \sum_j V_j \otimes C_{2j} = \hat{\sigma}_{12} \end{cases}$$

So, finally, we get that $\mathcal{F}_{123} \cap \mathcal{F}_{234} = \sigma_{12}\sigma_{23}\sigma_{34}$.

As second example, we are considering a non-commuting Gibbs state. In this case, the Hamiltonian is given by:

$$H = \sigma_x \otimes \sigma_x \otimes I_{34} + I \otimes \sigma_z \otimes \sigma_z \otimes I + I_{12} \otimes \sigma_x \otimes \sigma_x$$

where $\sigma_x = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$ and $\sigma_z = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$ are Pauli matrices.

In this case manual calculations are not easy, so we have used a mathematical software (MATLAB) in order to find if the Gibbs state can be prepared with NNN operators. We implemented the following algorithm:

Algorithm 1 Pseudo-code for the non commuting case:

find $\Sigma_{123}(\rho)$;
$\mathcal{F}_{123} \Leftarrow \Sigma_{123}$;
while rank($\mathcal{F}_{123}$) **is not** stable **do**
    find $alg(\Sigma_{123})$ closing Σ_{123} w.r.t. the distorted product;
    find $\Delta(alg(\Sigma_{123}))$ closing $alg(\Sigma_{123})$ w.r.t. modular product;
    $\mathcal{F}_{123} \Leftarrow \Delta(alg(\Sigma_{123}))$;
end while
$\mathcal{F}_{123} \Leftarrow \mathcal{F}_{123} \otimes \mathcal{B}(\mathcal{H}_4)$
Repeat for neighborhood $\{2, 3, 4\}$;
Calculate a basis for $\mathcal{F} = \mathcal{F}_{123} \cap \mathcal{F}_{234}$;
if rank($\begin{bmatrix} \mathcal{F} & \rho \end{bmatrix}$)==1 **then**
    ρ can be prepared;
else
    ρ cannot be prepared;
end if

Also in this case, thanks to the software, we found that the state can be prepared. The code used can be found in Appendix

Chapter 4

Bregman's Theory

4.1 Bregman's Divergences and their Properties

In this chapter it will be shown an extension of Halperin (Von Neumann) theorem using Bregman divergences and projections.

Definition 4.1 *Let $\mathcal{A}_i \in \mathcal{X}$ a family of closed convex sets; $\mathcal{R} = \bigcap_{i \in I} \mathcal{A}_i \neq \emptyset$ and $\mathcal{S} \subset \mathcal{X}$ a convex set such that $\mathcal{S} \cap \mathcal{R} \neq \emptyset$.*
The function $D(x, y)$ defined over $\mathcal{S} \times \mathcal{S}$ is a ***Bregman divergence*** *if satisfies the following conditions:*

I. $D(x, y) \geq 0$ *and* $D(x, y) = 0 \iff x = y$;

II. For every $y \in \mathcal{S}$, $i \in I$ it exists $x \in \mathcal{A}_i \cap \mathcal{S}$ such that:

$$D(x, y) = \arg\min_{z \in \mathcal{A}_i \cap \mathcal{S}} D(z, y)$$

x is called ***Bregman projection*** *of y onto $\mathcal{A}_i$ and it will be indicated by $P_i(y)$;*

III. For every index $i \in I$ and $y \in \mathcal{S}$ the function $G(z) = D(z, y) - D(z, P_i(y))$ is convex over $\mathcal{A}_i \bigcap \mathcal{S}$;

IV. A derivative $D'_x(x, y)$ of $D(x, y)$ exists when $x = y$ and $D'_x(y, y) = 0$ (i.e. $\lim_{t \to 0} \frac{D(y+tz, y)}{t} = 0$ for all $z \in \mathcal{X}$);

V. For every $z \in \mathcal{R} \cap \mathcal{S}$ and for every real number L, the set $T = \{x \in \mathcal{S} : D(z, x) \leq L\}$ is compact;

VI. If $D(x^n, y^n) \to 0$, $y^n \to y^* \in \bar{S}$ and the set of elements of $\{x^n\}$ is compact, $x^n \to y^*$.

The following proposition gives a useful tool to verify if a function is a Bregman divergence.

Proposition 4.1 *If $f(x)$ is a strictly convex and differentiable function, then*

$$D_f(x, y) = f(x) - f(y) - \langle \nabla f(y), x - y \rangle \tag{4.1.1}$$

satisfies conditions I-IV.

Proof.

I. It derives from the property of the convex functions:

$$f(x) \geq f(y) + \langle \nabla f(y), x - y \rangle$$

From which we obtain that $D_f(x, y) \geq 0$. And, for the strictly convexity, it is clear that $D_f(x, y) = 0 \Leftrightarrow x = y$.

II. This condition is satisfied because $\arg\min_{x \in \mathcal{S}} D(x, y) = 0$ exists, ergo $\arg\min_{x \in \mathcal{A}_i \cap \mathcal{S}} D(x, y)$ exists for every closed convex set $\mathcal{A}_i$.

III. It results

$$\begin{aligned} G(z) &= -f(y) + f(P_i(y)) - \langle \nabla f(y), y \rangle + \langle \nabla f(P_i(y)), P_i(y) \rangle - \langle \nabla f(y) - \nabla f(P_i(y)), z \rangle \\ &= cost. - \langle \nabla f(y) - \nabla f(P_i(y)), z \rangle \end{aligned}$$

which is convex.

IV. $D'_x(y, y) = \nabla f(y) - \nabla f(y) = 0$

□

Note. A "Bregman Distance" as (4.1.1) does not guarantee conditions V,VI. More assumptions are needed.
Here are two classical examples used by Bregman himself in his work.

Proposition 4.2 *Let $\mathcal{X}$ be an Hilbert space, $\mathcal{S} = \mathcal{X}$ and*

$$D(x, y) = \langle x - y, x - y \rangle$$

Then it is a Bregman Divergence.
Proof. Let us see that $D(x, y)$ satisfies the condition in Definition 4.1.

I. It is obvious by Definition 1.4;

II. In this case Bregman Projection is the classical orthogonal projection, so it is satisfied;

III. It holds:

$$\begin{aligned} G(z) &= D(z, y) - D(z, P_i(y)) \\ &= \langle z - y, z - y \rangle - \langle z - P_i(y), z - P_i(y) \rangle \\ &= 2\langle z, P_i(y) - y \rangle + \langle y, y \rangle - \langle P_i(y), P_i(y) \rangle \end{aligned}$$

which is linear in z, so the condition is satisfied;

IV. $D'_x(y, y) = \lim_{t \to 0} \frac{\langle y + tz - y, y + tz - y \rangle}{t} = \lim_{t \to 0} t\langle z, z \rangle = 0$;

V. $T = \{y \in \mathcal{S} : \langle x - y, x - y \rangle \leq L\}$ is not generally compact but it is bounded. A weak topology can be introduced in order to achieve compactness (i.e. $\mathbb{R}^n$, $\mathbb{C}^n$ have a strong topology, so this property is often satisfied);

VI. Let $\langle x^n - y^n, x^n - y^n \rangle \to 0$, $y^n \to y^*$ and let $\{x^n\}$ be compact. Let $x^{n_k} \to x^*$. Then for every $u \in \mathcal{X}$:

$$|\langle u, x^{n_k} - y^{n_k} \rangle| \leq \|u\| \|x^{n_k} - y^{n_k}\| \to 0$$

That implies $x^* = y^*$.

$\square$

Proposition 4.3 *Let $f(x) = \sum_{i=1}^{n} x_i \ln x_i$ be the entropy function, where x_i is a probability measure.*
Then

$$D_f(x, y) = \sum_{i=1}^{n} x_i \ln \frac{x_i}{y_i}$$

is a Bregman divergence.
A proof can be found in [5].

4.2 Generalized Pythagorean Theorem

In this paragraph we will give a generalization of Pythagorean theorem for Bregman divergence.

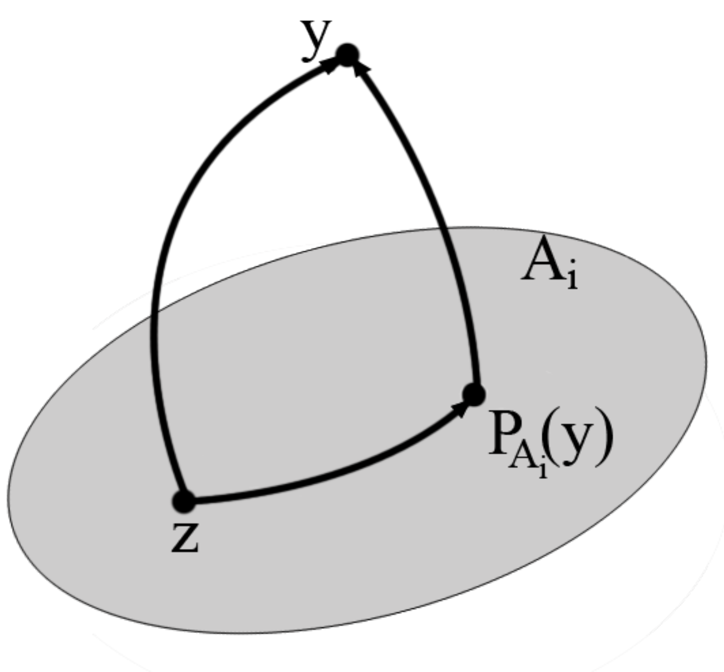

Figure 4.1: A visual representation of generalized Pythagorean theorem.

Lemma 4.1 *Let $z \in \mathcal{A}_i \cap \mathcal{S}$. Then for any $y \in \mathcal{S}$*

$$D(P_i(y), y) \leq D(z, y) - D(z, P_i(y))$$

is valid.

Proof. By condition III, $\lambda \in [0,1]$

$$\begin{aligned} D(\lambda z + (1-\lambda)P_i y, y) - D(\lambda z + (1-\lambda)P_i(y), P_i(y)) \quad &\leq \quad \lambda(D(z,y) - D(z, P_i(y))) + \\ &\quad +(1-\lambda)D(P_i(y), y) \end{aligned}$$

When $\lambda > 0$:

$$\begin{aligned} D(z,y) - D(z,P_i(y)) - D(P_i(y),y) \quad &\geq \quad \frac{D(\lambda z + (1-\lambda)P_i(y), y) - D(P_i(y), y)}{\lambda} - \\ &\quad - \frac{D(\lambda z + (1-\lambda)P_i(y), y)}{\lambda} \end{aligned}$$

Now $\lambda z + (1-\lambda)P_i(y) \in \mathcal{A}_i \cap \mathcal{S}$, the first term in the right side of the previous inequality is non-negative (condition II) and the second term tend to 0 when $\lambda \to 0$ (condition IV). From which

$$D(z,y) - D(z,P_i(y)) - D(P_i(y),y) \geq 0$$

□

In Figure 4.2 an intuitive vision of Lemma is given. It can be seen that Bregman Projections behave "like" classical projections giving a property similar to classical Pythagorean Theorem, essential property in order to prove the convergence of the iterated method that will be given in the next paragraph.

4.3 Iterated Convergence Results

Let us consider the following iterative process:

- choose $x^0 \in \mathcal{S}$
- if $x^n \in \mathcal{S}$ is known, select an index (in some way) $i_n(x^n) \in I$ and find the point x^{n+1} wich is the Bregman projection of x^n onto $\mathcal{A}_{i_n(x^n)}$.

The series $\{x^n\}$ is called **relaxation sequence**.

Lemma 4.2 *For any relaxation sequence it holds:*

1. *The set of elements $\{x^n\}$ is compact;*
2. *It exists* $\lim_{n\to\infty} D(z, x^n), \quad \forall z \in R$;
3. $D(x^{n+1}, x^n) \to 0$ *when* $n \to \infty$.

Proof. Let $z \in \mathcal{R} \cap \mathcal{S}$. By Lemma 4.1

$$D(x^{n+1}, x^n) \leq D(z, x^n) - D(z, x^{n+1})$$

Because $D(x^{n+1}, x^n) \geq 0$, we have $D(z, x^n) \geq D(z, x^{n+1})$, so it exists the limit for $D(z, x^n)$ which, with Lemma 4.1, gives $D(x^{n+1}, x^n) \to 0$. That proves properties 2-3.

Now the set of element of $\{x^n\}$ is contained in $T = \{x \in \mathcal{S} : D(z,x) \leq D(z,x^0)\}$, compact for condition V, so 1. is proven. $\square$

Lemma 4.2 assures the convergence of the relaxation sequence and it gives the basis to prove the main Bregman result.

Theorem 4.1 (Bregman's iterative method) *Let $I = \{1, 2, ..., m\}$ and let the indices be chosen in cyclic order. Then any limiting point x^* of the relaxation sequence $\{x^n\}$ is a common point of the sets $\mathcal{A}_i$.*
Proof. Let x^* be the limiting point of $\{x^n\}$ and $x^{n_k} \to x^*$.
Let us separate from $\{x^{n_k}\}$ a sequence fully contained in one $\mathcal{A}_i$ (i.e. $\{x^{n_k}\} \in \mathcal{A}_1$) and separate out from $\{x^{n_k+i-1}\}$ the sequences which are convergent (we assume $\{x^{n_k+i-1}\}$ themselves convergent). We have:

$$\begin{gathered} x^{n_k} \to x^* = x_1^* \\ x^{n_k+1} \to x_2^* \\ \vdots \\ x^{n_k+m-1} \to x_m^* \end{gathered}$$

$\{x^{n_k+i-1}\} \in \mathcal{A}_i \Rightarrow x_i^* \in \mathcal{A}_i$.
From Lemma 4.2 ($D(x^{n_k+1}, x^{n_k}) \to 0$) and condition VI $\Rightarrow \lim x^{n_k+1} = \lim x^{n_k} = x_1^* = x_2^*$ which implies $x^* \in \mathcal{A}_2$. In the same way it holds $x^* \in \mathcal{A}_3$,...
Concluding we have:

$$x^* \in \bigcap_{i \in I} \mathcal{A}_i$$

$\square$

Theorem 4.2 *If $\forall y \in \mathcal{S}$ it exists $\max_{i \in I} \min_{x \in \mathcal{A}_i} D(x,y)$, let $i_n(x^n)$ be the index which realizes*

$$\max_{i \in I} \min_{x \in \mathcal{A}_i} D(x, x^n);$$

Then any limiting point of the relaxation sequence is a common point of the sets $\mathcal{A}_i$.

In other words Bregman gives a sort of generalization of Von Neumann and Halperin's methods using Bregman projections to extend iterative methods.
Bregman divergences have been studied in various cases, for istance, one of the most important is the case of Bregman divergences generated by particular functions: Legendre functions. Details can be found in [6].

4.4 Quantum Bregman's Divergences

In this paragraph we will show how some quantum functions can be seen as Bregman divergences.

Proposition 4.4 *Let x and y be quantum states. The quadratic distance $\|x - y\|_\xi^2$ induced by the modified ξ-inner product is a Bregman divergence.*
Proof. Immediate from Proposition 4.4 $\square$

In the next proposition we will show that the quantum relative entropy is a Bregman divergence.

Proposition 4.5 *Let x, y be strictly positive quantum states and let $f(x) = \mathrm{tr}(x \log x)$. Then*

$$D_f(x, y) = \mathrm{tr}(x \log x) - \mathrm{tr}(x \log y) \tag{4.4.1}$$

is a Bregman distance which also satisfies conditions V-VI. In other words it is a Bregman divergence.

Proof. First of all let us recall the following identity:

$$D_f(x, y) = f(x) - f(y) - \langle \nabla f(y), x - y \rangle = f(x) - f(y) - \lim_{x \to 0^+} t^{-1}[f(y + t(x - y)) - f(y)]$$

Now $f(x) = \mathrm{tr}[x \log x]$ from which:

$$D_f(x, y) = \mathrm{tr}[x \log x] - \mathrm{tr}[y \log y] - \lim_{x \to 0^+} t^{-1}\mathrm{tr}[(y + t(x - y)) \log(y + t(x - y)) - y \log y]$$

$$\begin{aligned}
\lim_{x \to 0^+} t^{-1}\mathrm{tr}[(y + t(x - y)) \log(y + t(x - y)) - y \log y] &= \\
\lim_{x \to 0^+} t^{-1}\mathrm{tr}[y \log(y + t(x - y)) - y \log y] + \mathrm{tr}[(x - y) \log y] &= \\
\lim_{x \to 0^+} t^{-1}\mathrm{tr}[y \log(y + t(x - y)) - y \log y] &= \\
\lim_{x \to 0^+} t^{-1}\mathrm{tr}[y(\log(y) + t \log' y(x - y) + o(t^2)) - y \log y] &= \\
\mathrm{tr}[x - y] &= \\
0 &
\end{aligned}$$

Finally:
$D_f(x, y) = \mathrm{tr}[x \log x] - \mathrm{tr}[y \log y] - \mathrm{tr}[x \log y] + \mathrm{tr}[y \log y] = \mathrm{tr}[x \log x] - \mathrm{tr}[x \log y]$
Conditions V-VI are clearly satisfied. $\square$

Appendix A

Tensor Product and Partial Trace

As said in the relative chapter, tensor product is a way to assemble vector spaces in order to obtain a larger vector space. Consider two Hilbert spaces $\mathcal{H}_1$ and $\mathcal{H}_2$ of dimensions n and m respectively, then $\mathcal{H}_1 \otimes \mathcal{H}_2$ is an nm dimensional vector space. If we consider the elements $v \in \mathcal{H}_1$ and $w \in \mathcal{H}_2$, then the elements in $\mathcal{H}_1 \otimes \mathcal{H}_2$ are of the kind $v \otimes w$, in addiction to that if $\{i\}$ and $\{j\}$ are orthonormal bases for $\mathcal{H}_1$ and $\mathcal{H}_2$, then $\{i\} \otimes \{j\}$ is a basis for $\mathcal{H}_1 \otimes \mathcal{H}_2$.

By definition, tensor product satisfies the following properties:

- let t be a scalar element, and $v \in \mathcal{H}_1$, $w \in \mathcal{H}_2$, then

$$t(v \otimes w) = tv \otimes w = v \otimes tw;$$

- let $v_1, v_2 \in \mathcal{H}_1$ and $w \in \mathcal{H}_2$, then

$$(v_1 + v_2) \otimes w = v_1 \otimes w + v_2 \otimes w;$$

- let $v \in \mathcal{H}_1$ and $w_1, w_2 \in \mathcal{H}_2$, then

$$v \otimes (w_1 + w_2) = v \otimes w_1 + v \otimes w_2.$$

Now, let us suppose that A is a linear operator in $\mathcal{H}_1$ and B is a linear operator acting on $\mathcal{H}_2$. We can produce an operator $A \otimes B$ acting on $\mathcal{H}_1 \otimes \mathcal{H}_2$ and the following relation is valid:

$$(A \otimes B)(v \otimes w) = Av \otimes Bw$$

In matrix representation, tensor product is known as the Kronecker product and the previous relation derives from the distributive property of the multiplication.
Given two matrices A of dimensions $m \times n$ and B $p \times q$, the Kronecker product is given by:

$$A \otimes B = \begin{bmatrix} A_{11}B & ... & A_{1n}B \\ \vdots & \ddots & \vdots \\ A_{m1}B & ... & A_{mn}B \end{bmatrix}$$

which is a $mp \otimes nq$ matrix.

Now, consider a composite system $\mathcal{H}_{AB} = \mathcal{H}_A \otimes \mathcal{H}_B$. The partial trace over the subsystem B is defined as:

$$\mathrm{tr}_B(\cdot) : \mathcal{H}_{AB} \to \mathcal{H}_A$$

Consider, for example, the state$\rho = \rho_a \otimes \rho_B$, we obtain $\mathrm{tr}_B(\rho) = \rho_A \mathrm{tr}(\rho_B)$. The partial trace can be thought as a way to obtain the information of a particular subsystem. For example, let us consider a bipartite qubit system $\mathcal{H}_A \otimes \mathcal{H}_B$; we have

$$\rho_{AB} = \begin{bmatrix} a & b & c & d \\ e & f & g & h \\ l & m & n & o \\ p & q & r & s \end{bmatrix} \Rightarrow \rho_A = \begin{bmatrix} \mathrm{tr}\begin{bmatrix} a & b \\ e & f \end{bmatrix} & \mathrm{tr}\begin{bmatrix} c & d \\ g & h \end{bmatrix} \\ \mathrm{tr}\begin{bmatrix} l & m \\ p & q \end{bmatrix} & \mathrm{tr}\begin{bmatrix} n & o \\ r & s \end{bmatrix} \end{bmatrix} = \begin{bmatrix} a+f & c+h \\ l+q & n+s \end{bmatrix}$$

while

$$\rho_B = \begin{bmatrix} a+n & b+o \\ e+r & f+s \end{bmatrix}$$

For more information about partial trace and tensor products refer to specific texts.

Appendix B

MATLAB Code

```
%%Thesis code for the thesis example.
%%Beware: not refined code!
close all;
clear all;

%def hamiltonian matrices
x=[0 1; 1 0]; %sigma_x
z=[1 0; 0 -1]; %sigma_z
e=[1 0; 0 1]; % id

%elementary matrices
%e_1=[1 0; 0 0];
%e_2=[0 1; 0 0];
%e_3=[0 0; 1 0];
%e_4=[0 0; 0 1];

%pauli matrices
e_1=[1 0; 0 1];
e_2=[0 1; 1 0];
e_3=[1 0; 0 -1];
e_4=[0 -1i; 1i 0];
%subsystem dimensions vector
dim=[2 2 2 2];

H=kron(kron(kron(x,x),e),e)+kron(kron(kron(e,z),z),e)+kron(kron(kron(e,e),x),x); %hamiltonian
rho=expm(-H)/trace(expm(-H)); %normalized state
rho_123=TrX(rho,[4], dim);  %neighborhood{1,2,3} reduced state
rho_234=TrX(rho,[1], dim);  %neighborhood{2,3,4} reduced state

%%Computing minimal modular invariant *algebra for 123
%Find Sigma_123, Sigma_123v is in vectorial form
Sigma1_123=TrX(kron(eye(8),e_1)*rho,[4], dim);
Sigma2_123=TrX(kron(eye(8),e_2)*rho,[4], dim);
```

```
Sigma3_123=TrX(kron(eye(8),e_3)*rho,[4], dim);
Sigma4_123=TrX(kron(eye(8),e_4)*rho,[4], dim);
Sigma1_123v=reshape(Sigma1_123,[],1);
Sigma2_123v=reshape(Sigma2_123,[],1);
Sigma3_123v=reshape(Sigma3_123,[],1);
Sigma4_123v=reshape(Sigma4_123,[],1);
Sigma_123v=[Sigma1_123v Sigma2_123v Sigma3_123v Sigma4_123v];
Sigma_123=[Sigma1_123 Sigma2_123 Sigma3_123 Sigma4_123];

%iteration in order to find F_123
oldrank=0;
oldrank2=1;
while oldrank2>oldrank
    %algebra cycle
   while rank(Sigma_123v)>oldrank
       oldrank=rank(Sigma_123v);
       for i=1:8:length(Sigma_123)
           for j=1:8:length(Sigma_123)
               p_j=Sigma_123(:,i:i+7)*pinv(rho_123)*Sigma_123(:,j:j+7);
               if rank([Sigma_123v reshape(p_j,[],1)])>rank(Sigma_123v)
                   Sigma_123v=[Sigma_123v reshape(p_j,[],1)];
                   Sigma_123=[Sigma_123 p_j];
               end
           end
       end
   end
   %modular invariancy cycle
   while rank(Sigma_123v)>oldrank2
       oldrank2=rank(Sigma_123v);
       for i=1:8:length(Sigma_123)
               p_j=(rho_123)^(1/2)*Sigma_123(:,i:i+7)*pinv(rho_123)^2;
               if rank([Sigma_123v reshape(p_j,[],1)])>rank(Sigma_123v)
                   Sigma_123v=[Sigma_123v reshape(p_j,[],1)];
                   Sigma_123=[Sigma_123 p_j];
               end

       end
   end
end
%Extending with B(H4)
F1_123=kron(Sigma_123,e_1);
F2_123=kron(Sigma_123,e_2);
F3_123=kron(Sigma_123,e_3);
F4_123=kron(Sigma_123,e_4);
F_123=[F1_123 F2_123 F3_123 F4_123];
F_123v=[reshape(F1_123,256,[]) reshape(F2_123,256,[]) reshape(F3_123,256,[]) reshape(F4_123,256,[]
```

```
%%Same thing for Subsystem 234
Sigma1_234=TrX(kron(e_1,eye(8))*rho,[1], dim);
Sigma2_234=TrX(kron(e_2,eye(8))*rho,[1], dim);
Sigma3_234=TrX(kron(e_3,eye(8))*rho,[1], dim);
Sigma4_234=TrX(kron(e_4,eye(8))*rho,[1], dim);
Sigma1_234v=reshape(Sigma1_234,[],1);
Sigma2_234v=reshape(Sigma2_234,[],1);
Sigma3_234v=reshape(Sigma3_234,[],1);
Sigma4_234v=reshape(Sigma4_234,[],1);
Sigma_234v=[Sigma1_234v Sigma2_234v Sigma3_234v Sigma4_234v];
Sigma_234=[Sigma1_234 Sigma2_234 Sigma3_234 Sigma4_234];
oldrank_0=0;
oldrank_2=1;
while oldrank_2>oldrank_0
   while rank(Sigma_234v)>oldrank_0
       oldrank_0=rank(Sigma_234v);
       for i=1:8:length(Sigma_234)
           for j=1:8:length(Sigma_234)
               p_j=Sigma_234(:,i:i+7)*pinv(rho_234)*Sigma_234(:,j:j+7);;
               if rank([Sigma_234v reshape(p_j,[],1)])>rank(Sigma_234v)
                   Sigma_234v=[Sigma_234v reshape(p_j,[],1)];
                   Sigma_234=[Sigma_234 p_j];
               end
           end
       end
   end
   while rank(Sigma_234v)>oldrank_2
       oldrank_2=rank(Sigma_234v);
       for i=1:8:length(Sigma_234)
               p_j=(rho_234)^(1/2)*Sigma_234(:,i:i+7)*pinv(rho_234)^2;
               if rank([Sigma_234v reshape(p_j,[],1)])>rank(Sigma_234v)
                   Sigma_234v=[Sigma_234v reshape(p_j,[],1)];
                   Sigma_234=[Sigma_234 p_j];
               end

       end
   end
end
F1_234=kron(e_1, Sigma_234);
F2_234=kron(e_2, Sigma_234);
F3_234=kron(e_3, Sigma_234);
F4_234=kron(e_4, Sigma_234);
F_234=[F1_234 F2_234 F3_234 F4_234];
F_234v=[reshape(F1_234,256,[]) reshape(F2_234,256,[]) reshape(F3_234,256,[]) reshape(F4_234,256,[]

%%finding intersection
%ints(): function of the Geometry Approach Toolbox
```

```
rank([ints(F_123v, F_234v) reshape(rho,[],1)])

%checking if spans are the same
rank([I reshape(rho,[],1)])
```

Bibliography

[1] Shmuel Agmon. The relaxation method for linear inequalities. *Canadian Journal of Mathematics*, 6(3):382–392, 1954.

[2] Claudio Altafini and Francesco Ticozzi. Modeling and control of quantum systems: an introduction. *Automatic Control, IEEE Transactions on*, 57(8):1898–1917, 2012.

[3] Heinz H Bauschke and Adrian S Lewis. Dykstras algorithm with bregman projections: A convergence proof. *Optimization*, 48(4):409–427, 2000.

[4] Robin Blume-Kohout, Hui Khoon Ng, David Poulin, and Lorenza Viola. Information-preserving structures: A general framework for quantum zero-error information. *Physical Review A*, 82(6):062306, 2010.

[5] Lev M Bregman. The relaxation method of finding the common point of convex sets and its application to the solution of problems in convex programming. *USSR computational mathematics and mathematical physics*, 7(3):200–217, 1967.

[6] Yair Censor and Arnold Lent. An iterative row-action method for interval convex programming. *Journal of Optimization theory and Applications*, 34(3):321–353, 1981.

[7] Yuen-Lam Cheung, Dmitriy Drusvyatskiy, Chi-Kwong Li, Diane Pelejo, and Henry Wolkowicz. Projection methods in quantum information science. *arXiv preprint arXiv:1407.6604*, 2014.

[8] Imre Csiszár. I-divergence geometry of probability distributions and minimization problems. *The Annals of Probability*, pages 146–158, 1975.

[9] Domenico d'Alessandro. *Introduction to quantum control and dynamics*. CRC press, 2007.

[10] Kenneth R. Davidson. *C*-algebras by example*, volume 6. American Mathematical Soc., 1996.

[11] Richard L Dykstra. An algorithm for restricted least squares regression. *Journal of the American Statistical Association*, 78(384):837–842, 1983.

[12] René Escalante and Marcos Raydan. *Alternating projection methods*, volume 8. SIAM, 2011.

[13] Karolas M Grigoriadis and Robert E Skelton. Fixed-order control design for lmi control problems using alternating projection methods. In *Decision and Control, 1994., Proceedings of the 33rd IEEE Conference on*, volume 3, pages 2003–2008. IEEE, 1994.

[14] Israel Halperin. The product of projection operators. *Acta Sci. Math.(Szeged)*, 23:96–99, 1962.

[15] Peter D Johnson, Francesco Ticozzi, and Lorenza Viola. General fixed points of quasi-local frustration-free quantum semigroups: from invariance to stabilization. *arXiv preprint arXiv:1506.07756*, 2015.

[16] Stefan Kaczmarz. Angenäherte auflösung von systemen linearer gleichungen. *Bulletin International de l'Academie Polonaise des Sciences et des Lettres*, 35:355–357, 1937.

[17] TS Motzkin and IJ Schoenberg. The relaxation method for linear inequalities. *Canadian Journal of Mathematics*, 6(3):393–404, 1954.

[18] J von Neumann. *Functional operators*, volume 2. Princeton Univ. Press Princeton, N. J, 1950.

[19] Michael A Nielsen and Isaac L Chuang. *Quantum computation and quantum information*. Cambridge university press, 2010.

[20] Denes Petz. Bregman divergence as relative operator entropy. *Acta Mathematica Hungarica*, 116(1-2):127–131, 2007.

[21] Dénes Petz. *Quantum information theory and quantum statistics*. Springer Science & Business Media, 2007.

[22] Francesco Ticozzi. Symmetrizing quantum dynamics beyond gossip-type algorithms. *arXiv preprint arXiv:1509.01621*, 2015.

[23] Michael M Wolf. Quantum channels & operations: Guided tour. *Lecture notes available at http://www-m5. ma. tum. de/foswiki/pub M*, 5, 2012.

Ringraziamenti

Desidero ringraziare immensamente il mio relatore, il professor Francesco Ticozzi per la pazienza, il tempo e la passione che mi ha dedicato per seguirmi in questo lavoro.
Voglio fortemente ringraziare la mia famiglia, sempre disponibile a tollerare i miei scleri universitari.
Ringrazio i miei amici, anch'essi sopravvissuti: Gabriele, compagno d'avventura in questi lunghi anni, Fabio ed Alessandro, sfaccendati sempre troppo indaffarati e tutti gli altri che sono in grado di sopportarmi (Damiano, Giulia, Lucrezia, "Quei del 90", etc...).

Grazie, grazie di cuore.