Université catholique de Louvain
Institute of Communication Technologies, Electronics and Applied Mathematics
Department of Applied Mathematics

Convergent Products of Stochastic Matrices: Complexity and Algorithms

Pierre-Yves Chevalier

Thesis submitted in partial fulfillment of the requirements for the Ph.D. Degree in Engineering Sciences

Thesis Committee

Advisors -	Julien Hendrickx	*Université catholique de Louvain*
	Raphaël Jungers	*Université catholique de Louvain*
Jury -	Jean-Pierre Tignol	*Université catholique de Louvain*
	Paolo Frasca	*CNRS, GIPSA-lab*
	Rodolphe Sepulchre	*University of Cambridge*
Chairman -	Philippe Lefèvre	*Université catholique de Louvain*

Louvain-la-Neuve, 2018.

Acknowledgements

I am very grateful to all the people without whom this thesis would not have been possible.

First of all, I would like to thank my advisors Julien and Raphaël who gave me the opportunity to start a thesis in INMA and who gave me guidance throughout this PhD. I would also like to thank Rodolphe Sepulchre, Jean-Pierre Tignol, Paolo Frasca and Philippe Lefèvre who have accepted to be members of my thesis committee. My thesis has also benefited from interactions with researchers including Vladimir Gusev, François Gonze, Alex Olshevsky, Julien Dewez, Jürgen Eckhoff, Jean-Paul Doignon, Samuel Fiorini, Julien Leroy.

I would like to thank my successive office mates: Adrien, Gianni, Nikos and François. I am also thankful to all the people who made of the Euler building a pleasing work environment: Corentin, Pierre-Alexandre, Maxime, Luc, Benoît, Matthew, Romain, Nicolas and to the administrative and technical people: Nathalie, Marie-Christine, Isabelle, Étienne, François W. and Pascale.

Finally, I would like to thank my family, my friends and of course my partner Anaïs.

Contents

1 Introduction **9**
1.1 Consensus Systems . . . 9
1.2 Consensus and Products of Stochastic Matrices . . . 10
1.3 State of the Art and Contributions . . . 11
1.4 Additional Application: Markov Chains . . . 14
1.5 Stochastic Matrices and Generalizations . . . 15
1.6 Organization of the Thesis . . . 17
1.7 List of Publications . . . 19

2 Matrix Classes Related to Convergence **21**
2.1 Definition and Properties of the Matrix Classes . . . 22
2.2 From SIA to a Positive Column . . . 27

I Contractive Sets **33**

3 Invariant Polyhedra **35**
3.1 Polyhedra and Faces . . . 36
3.1.1 Polyhedra and Polytopes . . . 36
3.1.2 Faces, Open Faces and Facets . . . 36
3.1.3 Invariant Polyhedra . . . 39
3.1.4 Polarity . . . 41
3.1.5 Centrally Symmetric Polyhedra and Seminorms . . . 42
3.2 Invariant Polyhedron for Stochastic Matrices . . . 43
3.2.1 Combinatorial Structure of $\mathcal{P}$ . . . 45
3.2.2 Other Invariant Polyhedra for Stochastic Matrices . . . 49

4 Complexity of Deciding Consensus **51**
4.1 Contractive Sets . . . 52
4.1.1 Properties of Contractive Sets . . . 54
4.2 Algorithmic Decision of Consensus . . . 57
4.2.1 Graph Representation of Products . . . 58
4.3 Computational aspects . . . 61
4.3.1 Discussion . . . 62
4.4 Sets of two stochastic undirected matrices . . . 63
4.5 Conclusion . . . 65

5 Tight Bound for Consensus **67**
5.1 Introduction . . . 68
5.2 The General Case . . . 70
5.3 The case of the Polyhedron $\mathcal{P}$ . . . 75
5.3.1 A New Finiteness Bound for Contractive Sets . . . 78
5.3.2 Stochastic Matrices . . . 79
5.3.3 Relation with SIA Matrices . . . 81
5.4 Conclusion . . . 81

6 Application to Other Polyhedra **83**
6.1 Primitivity of Matrix Sets . . . 84
6.2 A Face Lattice without the Sperner Property . . . 87
6.2.1 Existence of a Suitable Projective Transformation . . . 89

II Almost Contractive Sets **97**

7 Synchronizing Automata **99**
7.1 Automata and Synchronization . . . 100
7.2 Reset Threshold and Černý Conjecture . . . 102
7.3 Deciding Synchronization . . . 103
7.4 Negative Complexity Results . . . 105
7.5 Conclusion . . . 106

8 Almost Contractive Sets of Matrices **107**
8.1 Introduction . . . 109
8.2 Indices and Bounds . . . 110
8.3 Relation between Indices . . . 112
8.4 Bounds on the Positive-Column Index . . . 119

8.5 Bounds on the SIA index 119
8.5.1 Upper bounds 120
8.5.2 Numerical results 120
8.5.3 Lower bounds 124
8.6 Complexity Results 125
8.6.1 Deciding whether a Set is Almost Contractive 126
8.6.2 Complexity of Approximating the Indices 126
8.7 Conclusion 129

III Conclusion **133**

Chapter 1

Introduction

1.1 Consensus Systems

The problem of how a group of agents can reach agreement on some value has attracted a lot of attention in recent years. The need for coordination schemes is present in many applications: autonomous platoons of vehicles [BJMP12], data fusion in systems with distributed measurements [OSS05, XBK07], distributed optimization [NO09] or coordination of groups of mobile agents (see [JLM03] and references therein). *Consensus systems* describe the dynamics of these coordination procedures. They have also been used as models for natural phenomena such as flocking [VCBJ$^+$95] or opinion dynamics [BHT09]. See also [OSFM07, ME10] for a survey.

In many of these models, each agent has a value x_i and it updates this value by taking a weighted average of the values of agents with which it can communicate:

$$x_i(t+1) = \sum_j a_{ij}(t)x_j(t) \tag{1.1}$$

where a_{ij} is the weight of the value of agent j in the computation of the new value of agent i. Because the agents compute their new value as a weighted average of the values of other agents, the weights satisfy

$$a_{ij}(t) \geq 0 \text{ and } \sum_j a_{ij}(t) = 1. \tag{1.2}$$

The values of the agents can be put in a single state vector x and Equation (1.1) becomes the following update equation.

$$x(t+1) = A(t)x(t), \tag{1.3}$$

for which the conditions in Equation (1.2) become

$$A(t) \geq 0 \text{ (elementwise), and } A(t)\mathbf{1} = \mathbf{1}, \tag{1.4}$$

with $\mathbf{1} = \begin{pmatrix} 1 & \ldots & 1 \end{pmatrix}^\top$ being the all-one vector. Matrices whose elements satisfy the conditions of Equation (1.4) are called *stochastic*.

1.2 Consensus and Products of Stochastic Matrices

Agents following these dynamics tend to be more and more in agreement, in the sense that their values generally get closer to each other. A general question is whether System (1.3) converges to a state of consensus, i.e., a state in which all agents have the same value or, equivalently, a multiple of the vector $\mathbf{1}$. We can see that system (1.3) converges for any initial condition $x(0)$ if and only if the limit

$$\lim_{t\to\infty} A(t) \ldots A(1)A(0). \tag{1.5}$$

exists and is a rank-one matrix, i.e., all its rows are equal. We will therefore study the convergence of the products of the transition matrices $A(t) \ldots A(1)A(0)$.

In a large class of systems, the transition matrix $A(t)$ depends on the state x, making the system nonlinear [JLM03, BHT09]. Deciding whether the system converges to consensus is therefore a hard problem. See for example [BHT07] that presents a relatively simple model for which no conditions for convergence to consensus are known. In some situations, even if it is hard to explicit the complete sequence of matrices $A(t)$ corresponding to System (1.1), it may be possible to guarantee that these matrices remain in some set S. In this thesis, we study convergence conditions based only on the knowledge of the set S. We will study two types of question.

- In Part I, we study questions regarding *all sequences* of transition matrices taken from the set S, for example: *does System* (1.3) *converge for every sequence of transition matrices taken from the set* S*?*
- In Part II, we will study questions regarding *one sequence* of transition matrices taken from S, for example: *is there one sequence of transition matrices from* S *such that System* (1.3) *converge?*

These questions can be formulated as questions on the set S, because of the equivalence between the convergence of the trajectories and convergence of the products. The first question becomes: *do all left-infinite products of matrices* *from S converge to a rank-one matrix?* and we will call *contractive* [PV12] the sets that have a positive answer to this question. These sets are called contractive because for a contractive set all infinite products of matrices taken from the set contract the space $\mathbb{R}^n$ to the subspace span$\{\mathbf{1}\}$.

These sets have also been called *quasidefinite* [Paz65], *weakly-ergodic* [CW08] in the context of Markov chains and *consensus sets* [BO14] in the context of consensus. We will study the properties of contractive sets and we will develop an algorithm to recognize them. The second question becomes *is there one left-infinite product of matrices from S that converges to a rank-one matrix?* and we will call *almost contractive* [PV12] the sets for which the answer to this question is positive.

A third important question is whether all products converge to a rank-one matrix with probability one, in the case where the transition matrix is chosen randomly at each step. Under the assumption that each matrix has a nonzero probability to be chosen a each step, this question is equivalent to that of the existence of a left-infinite product that converges to a rank-one matrix, as explained in [PV12, Section 5]. This is the reason why we denote *almost contractive* the sets for which one infinite product converges to a rank-one matrix. Our Part II will therefore have consequences for systems with random switching.

In light of this, one could wonder whether studying convergence for *all* sequences of transition matrices is useful at all and why convergence for *almost all* sequences is not always satisfying. If fact, in many situations the sequence of transition matrices of System 1.3 is deterministic but unknown (as in, for example, [BHT09, JLM03]) and the convergence with probability one does not guarantee anything about the convergence of the system for that particular deterministic sequence.

We will see that these questions about infinite products are in fact equivalent to simpler questions on finite products. Contractive sets are the sets for which all sufficiently long products have a positive column, while almost contractive sets are the sets for there is a product that has a positive column.

1.3 State of the Art and Contributions

A characterization of contractive sets of stochastic matrices has been known since [Wol63], in which the author shows that a set of stochastic matrix is contractive if and only if every finite product of matrices from S is *SIA* [Wol63], where a matrix is called *SIA* – Stochastic Indecomposable Aperiodic – if it is stochastic and the limit $\lim_{n\to\infty} A^n$ exists and is a rank-one matrix. Said differently all infinite products converge to a rank-one matrix if and only if all infinite *periodic*[1] products converge to a rank-one matrix. A more workable characterization was established by V.D. Blondel and A. Olshevsky who studied the complexity of deciding whether a set is a consensus set [BO14], or equivalently a contractive set. They prove that a set of stochastic matrices is contractive if and only if a certain condition is met for every product of length

$$B = \frac{1}{2}(3^n - 2^{n+1} + 1), \tag{1.6}$$

from which they can conclude that the problem of deciding if a set is contractive is algorithmically decidable. They also prove that the problem is NP-hard for sets of at least two matrices. They based some of their results on earlier work on inhomogeneous Markov chains which also involve long products of stochastic matrices. In particular, they make use of a result by A. Paz, that states that, for a given set of stochastic matrices, all sufficiently long products are *scrambling* if and only if all products of length at most B (as defined in (1.6)) are scrambling, where a matrix is called *scrambling* if the supports[2] of any two rows intersect [Paz71, Section A.4 of Chapter II].

The first part of this thesis will build on these results. In Chapter 4, we will propose a singly exponential algorithm to decide whether a set is contractive, and in Chapter 5, we will improve the bound (1.6). To obtain these results, we use the fact that System (1.3) is a *switching system*. In particular we use the fact that stochastic matrices have a *common invariant polyhedron* and we make use of techniques developed for switching systems that have an invariant polyhedron [Bar88, Mar06, LW95]. Invariant polyhedra have been used to study the *joint spectral radius* of matrix sets. The joint spectral radius of a set is the maximal asymptotic growth rate among all possible infinite products of matrices taken from the set [Jun09]. This quantity is closely related to convergence to consensus. Indeed, it has been shown [BHOT05, Section IV] that convergence to consensus of all products taken from a set S is equivalent to the joint spectral radius of an associated S' being strictly smaller than 1. The idea is that stochastic matrices have **1** as a common eigenvector and that the convergence to consensus corresponds to the convergence to this common eigenvector. The convergence to consensus is therefore equivalent to the convergence

[1] We say here that an infinite product is periodic if it is the infinite power of a finite product.

[2] We call the support of a row, the set of columns in which the elements of the row are positive.

to zero once this component in **1** has been projected. Finally, the convergence to zero of all infinite products of matrices taken from a set is equivalent to the joint spectral radius of this set being smaller than one. Considering this close relation between consensus and joint spectral radius, we will be able to use techniques developed for the joint spectral radius [LW95] to study contractive sets.

Almost contractive sets of matrices have been analyzed in [PV12, Section 5], where the authors develop a polynomial-time algorithm to decide whether a set of stochastic matrices is almost contractive. A particular subclass of almost contractive sets of matrices has also been studied in the context of automata theory. Indeed, the *synchronization* of *automata* – that we will introduce in Chapter 7 – is equivalent to the almost contractivity for the subclass of *automaton matrices*, that is, matrices that have one 1 on each row and 0 everywhere else. Synchronization of automata has been studied intensively in an attempt to prove the Černý conjecture that states that a set of automaton matrices has either a product of length at most $(n-1)^2$ that has a positive column or no product that has a positive column[3]. We will see that many results obtained in the context of automata can be generalized to sets of stochastic matrices. For instance, the algorithm that decides whether a set of stochastic matrices is almost contractive is an extension of a classical algorithm that decides whether automaton is synchronizing [PV12]. We will define the *SIA index* of a contractive set of stochastic matrices as the length of the shortest SIA product of the set. The SIA index is related to the well-studied notion of *exponent* – the length of the shortest (entrywise) positive product, if one exists (see [BR91, Section 3.5] for a survey of the single matrix case, and [GGJ18, PV12] for more recent work on matrix sets). Similar quantities have been defined for different matrix classes. For instance, the scrambling index is defined as the length of the shortest scrambling product. We will make use of results from automata theory and we will also show that the study of this quantity could have consequences in automata theory and for the state of the art of the Černý conjecture. We will also consider some related quantities in Chapter 8: the *positive-column index*, the *scrambling index* and the *Sarymsakov index*. These different indices come from different matrix classes that have been used to study the convergence of products of stochastic matrices. We will define these classes in Chapter 2 and we will study the relation between them.

[3]This formulation is equivalent to the classical formulation of the conjecture, that uses the notions of *automaton* and *synchronizing word*, that we will define only in Chapter 7.

Our work also is closely related to the theory of positive systems. Positive systems are linear systems that leave a cone invariant [Smi95, AS03]. In particular, systems whose state variables are nonnegative are positive systems. They appear in different contexts such as economics or biology. An important result in the theory of positive systems is the Birkhoff Theorem [Bir57] that uses the contraction of the Hilbert projective metric to establish the convergence to a ray in the interior of the cone. The well-known Perron-Frobenius Theorem is a particular case of this theorem. This approach has been applied to prove the convergence of consensus systems [SSR10]. In this thesis, however, we will mainly consider sets of matrices *that do not contract the Hilbert metric*. A contractive set of matrices is in fact a set of stochastic matrices from which all sufficiently long products contract the Hilbert metric. Moreover, if all matrices in a set contract the Hilbert metric, the set is trivially contractive and if a given matrix A_1 contracts the Hilbert metric, then $\{A_1, A_2 \ldots, A_m\}$ is contractive if and only if $\{A_2 \ldots, A_m\}$ is contractive. The results of Chapter 4 can thus be seen as a generalization of these results to sets of stochastic matrices that do not contract the Hilbert metric. In Chapter 8, we have a similar situation: if a set contains a matrix that contracts the Hilbert metric, then this set is almost contractive.

1.4 Additional Application: Markov Chains

Until now, we have used consensus systems as our main motivation for the study of products of stochastic matrices. We have also mentioned that some of our results are closely related to automata theory, in particular to synchronizing automata and to the Černý conjecture. In fact the study of products of stochastic matrices is also relevant in the context of *inhomogeneous Markov chains* [Str05]. Inhomogeneous Markov Chains are Markov chains in which the transition matrix can change with time. They have a set of states and a vector $x(t)$ whose components $x_i(t)$ are the probabilities to be in each state i. At each step t, there is a probability $a_{ij}(t)$ to transition from a state i to a state j. We thus have the following update equation for the vector of probabilities:

$$x^\top(t+1) = x(t)^\top A(t), \tag{1.7}$$

where the elements of $A(t)$ are the transition probabilities a_{ij}. The chain is called inhomogeneous if $A(t)$ can change with time. When the transition matrices come from a finite set S, the setting becomes similar to that of consensus, Equation (1.7) being the transpose of Equation (1.3). Deciding whether the set S is contractive is equivalent to deciding whether the chain is mixing, i.e., it forgets its initial condition while deciding whether S is almost contractive is equivalent to deciding whether the chain can be mixing (or is mixing with probability one in the case of random switching).

1.5 Stochastic Matrices and Generalizations

Our primary focus is on stochastic matrices, due to their applications in consensus and Markov chains. However, many results that we will obtain hold for larger classes of matrices, that we present here. First, let us recall the definition of a stochastic matrix.

Definition 1.1 (Stochastic matrix)**.** *A matrix A is called* stochastic *if*

$$\forall i, j,\ a_{ij} \geq 0$$

and

$$A\mathbf{1} = \mathbf{1}.$$

As observed in [ATB86], any stochastic matrix A satisfies

$$\max_i (Ax)_i - \min_i (Ax)_i \leq \max_i x_i - \min_i x_i.$$

Geometrically, this corresponds to the following polyhedron being invariant:

$$\mathcal{P} = \left\{ x \mid \frac{1}{2}(\max_i x_i - \min_i x_i) \leq 1 \right\}. \tag{1.8}$$

Most of the results that we will develop in Part I will hold for all sets of matrices that leave this polyhedron invariant. It is represented in dimension 2 and 3 in Figures 1.1 and 1.2.

Definition 1.2 (P-preserving matrix)**.** *A matrix A is call* P-preserving *if it leaves $\mathcal{P}$ invariant, that is if*

$$A\mathcal{P} \subseteq \mathcal{P}.$$

We will see in Section 3.2 that stochastic matrices are P-preserving but that not all P-preserving matrices are stochastic.

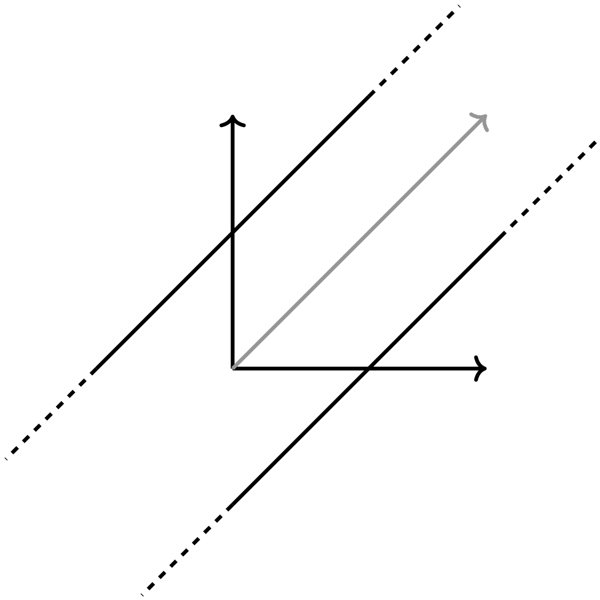


Figure 1.1: The polyhedron $\mathcal{P}$ (Equation 1.8) in dimension 2.

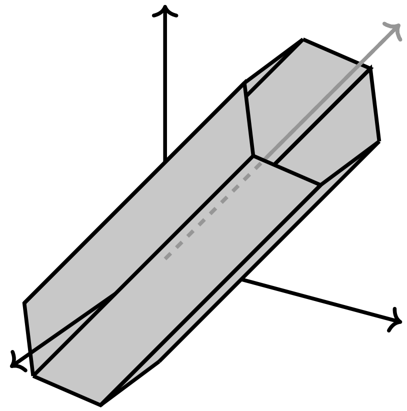


Figure 1.2: The polyhedron $\mathcal{P}$ in dimension 3.

Many of the results that we will develop depend only on the pattern of positive elements in the matrices and not on the values of these elements. Therefore, these results will hold not only for stochastic matrices but for matrices that share the same pattern of zero/nonzero elements. Stochastic matrices are *nonnegative* and they have at least one positive element on each row. Matrices that satisfy these conditions are called *row allowable*.

Definition 1.3 (Row allowable matrix [Har02]). *A matrix is called* row allowable *if it is nonnegative and it has a positive element on each row.*

The relation between row allowable, stochastic and P-preserving matrices is illustrated in Figure 1.3. Matrices that are row allowable and P-preserving are not necessarily stochastic. An example is given by

$$\begin{pmatrix} 2 & 0 \\ 1 & 1 \end{pmatrix},$$

which is row allowable and P-preserving but not stochastic.

We will see in Part II that the question of whether a set of stochastic matrices is almost contractive only depends on the pattern of nonzero elements in the matrices of the set and not on the value of these elements. Similarly, the SIA index of a set – that characterizes the length of the shortest SIA product – only depends on the pattern of positive elements. The techniques that we develop in this part could therefore be used to answer similar questions about sets of row allowable matrices instead of sets of stochastic matrices. However, the associated convergence result will not hold, since the convergence does depend on the values of the elements on the matrices and not only on the pattern of positive elements.

In Part I, we have a similar situation. The question of whether, for a set of stochastic matrices, all infinite products converge to a rank-one matrix is equivalent to all sufficiently long products being scrambling (see discussion in Section 4.1) which depends only on the pattern of positive elements of the matrices of S. If one wants to know whether all sufficiently long products of matrices from a set of *row allowable* matrices S' are scrambling, one could construct a set of *stochastic* matrices that have the same pattern of positive elements as those of S' and use the techniques developed in Chapter 4.

1.6 Organization of the Thesis

As we have already mentioned, we address two types of problems:

- questions about all products such as "do all left-infinite products converge to a rank-one matrix?"
- and questions about the existence of a product "is there at least one left-infinite product that converges to a rank-one matrix?".

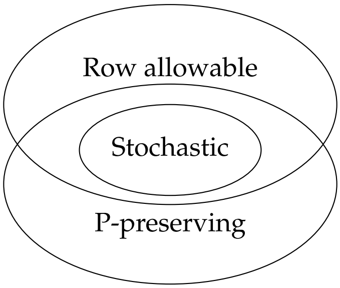


Figure 1.3: Relation between stochastic, P-preserving and row allowable matrices.

These two types of problems will form the two parts of this thesis. In each of these parts, we mainly ask two type of questions:

- complexity questions such as the complexity of deciding, given a set of P-preserving matrices, if all sufficiently long products contract $\mathcal{P}$,
- bounds or finiteness questions such as a quantity ℓ such that all sufficiently long products are SIA if and only if all products of length at least ℓ are SIA.

Table 1.1 summarizes our main results.

The first part deals with questions about all products of matrices taken from a given set.

- In Chapter 3, we introduce polyhedra and we study the properties of a particular polyhedron $\mathcal{P}$ that we use in the next chapters.
- In Chapter 4, we analyze the complexity of deciding whether a set of P-preserving matrices is contractive.
- In Chapter 5, we prove that a set is contractive if and only if all products of length $\leq p^* \triangleq \binom{n}{\lfloor n/3 \rfloor}(2^{n-\lfloor n/3 \rfloor -1} - 1)$ have power that contracts $\mathcal{P}$. In the case of stochastic matrices, this means that a set is contractive if and only if all products of length at most p^* are SIA.
- In Chapter 6, we show how the same techniques can be applied to other settings and invariant polyhedra. First, we show that our techniques can be used to analyze primitive sets of matrices and recover known results from the literature in this case. Second, we construct a polytope whose face lattice does not have the *Sperner property*. The Sperner property –

that we define in Chapter 5 – of $\mathcal{P}$ plays a key role in Chapter 5 and this example that we construct shows that not all polyhedra have the Sperner property.

The second part deals with questions about the existence of at least one product with a given property.

- In Chapter 7, we introduce some basic notions about automata and synchronization.
- In Chapter 8, we analyze almost contractive sets. We study mainly two quantities of almost contractive sets: the SIA index and the positive-column index, that are defined as the length of the shortest SIA and positive-column products. For the positive-column index, we prove an upper bound of $\frac{n^3-n}{6}$ and we show that any improvement of this bound would automatically improve the state of the art of the Černý conjecture. For the SIA index, we give the same upper bound but we conjecture that a much smaller upper bound of $2n$ holds. We support this conjecture with computational experiment. We also analyze the related notions of Sarymsakov and scrambling indices. Finally, we show that computing and even approximating these indices is NP-hard.

1.7 List of Publications

This thesis has led to the following publications.

- P-Y Chevalier, J. M. Hendrickx and R. M. Jungers, *A switched system approach to the decidability of consensus*, 21st International Symposium on the Mathematical Theory of Networks and Systems, 103-109, 2014.
- P-Y Chevalier, J. M. Hendrickx and R. M. Jungers, *Efficient algorithms for the consensus decision problem*, SIAM Journal on Control and Optimization 53 (5), 3104-3119, 2015.
- P-Y Chevalier, J. M. Hendrickx and R. M. Jungers, *Reachability of consensus and synchronizing automata*, 54th IEEE Control and Decision Conference, 2015.
- P-Y Chevalier, J. M. Hendrickx and R. M. Jungers, *Tight bound for deciding convergence of consensus systems*, Systems and Control Letters 105, 78-83, 2016.

	Complexity	Bounds
Contractive sets	**State of the art:** Decidable but NP-hard to decide whether a set is contractive [BO14]. **Contribution:** Singly exponential algorithm that decides if a set is contractive [CHJ15a]. **Results apply to:** sets of P-preserving matrices. **See:** Chapter 4.	**State of the art:** A set is contractive if and only if all products of length $B = \frac{1}{2}(3^n - 2^{n+1} + 1)$ are scrambling [Paz71, BO14]. **Contribution:** a set is contactive if and only if all products of length $\leq p^* = \binom{n}{\lfloor n/3 \rfloor}(2^{n-\lfloor n/3 \rfloor - 1} - 1) \approx \frac{3}{2\sqrt{\pi n}} B$ have a power that contracts $\mathcal{P}$ [CHJ16]. **Results apply to:** sets of P-preserving matrices. **See:** Chapter 5.
Almost contractive sets	**State of the art:** Existence of a positive-column product can be decided in polynomial time [PV12]. **Contribution:** NP-hard to approximate the length of the shortest positive-column or SIA product. This remains NP-hard even when restricted to automaton matrices or to matrices with positive diagonals [CHJ15b]. **See:** Section 8.6.	**Contribution:** bound on the positive-column index is the same as the bound on the reset threshold of synchronizing automata. Any improvement of the bound on the reset threshold of automata would translate into an improvement of the bound on positive-column index, and conversely [CHJ15b]. We conjecture that the bound on the SIA index is linear in n and support this conjecture [CGHJ17]. **See:** Sections 8.4 and 8.5.

Table 1.1: Main results of the thesis

- P-Y Chevalier, V. V. Gusev, J. M. Hendrickx and R. M. Jungers, *Sets of Stochastic Matrices with Converging Products: Bounds and Complexity*, submitted.

Chapter 2

Matrix Classes Related to Convergence

Many types of matrices have been studied in the context convergence of inhomogeneous products of stochastic matrices. They will be used throughout the thesis. We present here these classes and their relation to one another.

This chapter is divided in two sections.

- In Section 2.1, we define the classes of *positive-column*, *scrambling*, *Sarymsakov SIA* matrices. We introduce basic properties of these classes and we show that these classes are included in one another.
- In Section 2.2 we provide an upper bound on the power at which an SIA matrix has a positive column. From this upper bound we will derive a simple method to check whether a stochastic matrix is SIA. We also provide a tight upper bound in the case of stochastic matrices that have one 1 on each row and 0 everywhere else. These matrices, that we call *automaton matrices*, will play a central role in Chapter 8.

2.1 Definition and Properties of the Matrix Classes

We start by defining the rather intuitive notion of a positive-column matrix.

Definition 2.1 (Positive-column matrix)**.** *We call* positive-column matrix *a stochastic matrix that has a positive column.*

Positive-column matrices are called *Markov matrices* in [Sen81] but we will avoid this term since it is also used to denote stochastic matrices. Positive-column matrices have a simple interpretation in the context of consensus systems. They correspond to consensus iterations in which one agent j influences all the others, in the sense that any agent i uses the value of agent j in the computation of its new value.

Definition 2.2 (Scrambling matrix [Har02])**.** *A stochastic matrix is called* scrambling *if for any pair of rows (i, j), there is a column k such that $a_{ik} > 0$ and $a_{jk} > 0$.*

Scrambling matrices have been studied in the context of inhomogeneous Markov chains (see for example [Paz71] and [Sen81]) because of their contraction properties. It has been observed that a stochastic matrix A is scrambling if and only if, for any vector, A contracts the *Hilbert projective metric* [Sen81, Lemma 3.2]:

$$\forall x, y > 0, \quad d(Ax, Ay) < d(x, y),$$

where the Hilbert metric in the nonnegative orthant is defined as

$$d(x, y) = \log \left(\frac{\max(x_i/y_i)}{\min(x_i/y_i)} \right)$$

and can be seen as a measure of the angle between two vectors. Equivalently, a stochastic matrix is scrambling if and only if its *Birkhoff contraction coefficient* (or *coefficient of ergodicity*)

$$\tau(A) = \sup_{x,y>0,\ x \neq ay} \frac{d(Ax, Ay)}{d(x, y)}$$

is strictly smaller than one. The Birkhoff contraction coefficient is a great tool to study the convergence rate of inhomogeneous products of stochastic matrices. The Birkhoff contraction coefficient of a stochastic matrix A is equal to zero if and only if A is rank-one. It is an upper bound on the convergence rate of sequence of powers of A. Contrary to eigenvalues, the Birkhoff contraction coefficient is submultiplicative, that is, for any two stochastic matrices A and B

$$\tau(AB) \leq \tau(A)\tau(B).$$

This implies that the product of two scrambling matrices is scrambling and that any infinite product of matrices taken from a compact set of scrambling matrices converges to a matrix that has a Birkhoff coefficient of zero, i.e., to a rank-one matrix.

Scrambling matrices also have a consensus interpretation. If a matrix is scrambling, then in the corresponding consensus iteration, for any pair of agents i and j, there is an agent k that influences both agents i and j. This agent k will help the two agents i and j to converge to the same value and iterating a scrambling matrix will lead to consensus.

It is well known that any sufficiently long product of scrambling matrices has a positive column [Har02, Theorem 4.6]. In the next proposition, we quantify what sufficiently long means. We did not find this result in the literature.

Proposition 2.1. *Any product of $\ell = \lceil \log_2(n) \rceil$ scrambling $n \times n$ matrices has a positive column.*

Proof. Let A be an $n \times n$ scrambling matrix. For any pair of rows $i_1, i_2 \in \{1, \ldots, n\}$, there are $j \in \{1, \ldots, n\}$ and $a \in \mathbb{R}$ such that $Aae^j \geq e^{i_1} + e^{i_2}$, where e^i is the vector in which the ith element is 1 and all other elements are 0. Therefore, for any nonnegative vector v with exactly k positive elements, there exists a vector w with no more than $\lceil \frac{k}{2} \rceil$ positive elements such that, $Aw \geq v$.

Now let $A_1 A_2 \ldots A_\ell$ be any product of ℓ $n \times n$ scrambling matrices. There is $v_1 \in \mathbb{R}^n$ with no more than $\lceil \frac{n}{2} \rceil$ positive elements such that $A_1 v_1$ has a positive column. There is $v_2 \in \mathbb{R}^n$ with no more than

$$\left\lceil \frac{\lceil \frac{n}{2} \rceil}{2} \right\rceil = \left\lceil \frac{n}{4} \right\rceil$$

positive elements such that $A_1 A_2 v_2$ has a positive column. By repeating the same argument, we obtain that there is a vector v_ℓ with only $\lceil \frac{n}{2^\ell} \rceil = 1$ positive element and such that $A_1 A_2 \ldots A_\ell$ has a positive column. ■

Scrambling matrices have the following properties:

- the class is closed under multiplication: the product of two scrambling matrices is scrambling [Sen81, Corollary that follows Theorem 4.11],

- for any matrix size n, there is a length ℓ such that the product of ℓ $(n \times n)$ scrambling matrices has a positive column (Proposition 2.1).

There is another class of matrices that has these two interesting properties: the class of Sarymsakov matrices. Furthermore, this class is larger in the sense that it contains all scrambling matrices. To define, we will use the *consequent function* defined as follows. For any $S \subseteq \{1, \dots, n\}$, the consequent function F of an $n \times n$ matrix A is defined as

$$F(S) = \{j : \exists i \in S \text{ s.t. } a_{ij} > 0\}.$$

Definition 2.3 (Sarymsakov matrix [Har02])**.** *We call a* Sarymsakov matrix *a stochastic matrix such that for any two disjoint nonempty subsets of* $\{1, \dots, n\}$*,* S *and* S' *either*

$$F(S) \cap F(S') \neq \emptyset \tag{2.1}$$

or

$$|F(S) \cup F(S')| > |S \cup S'| \tag{2.2}$$

The consequent function can again be interpreted in terms of influence in a consensus system. The consequent function of a set of agents S is the set of agents that influence the value of at least one agent of S. Condition (2.1) therefore means that there is an agent influencing at least one agent in S and one in S'. This condition helps the values of agents in S to get closer to the values of agents in S'. Condition (2.2) means that there are more than $|S \cup S'|$ agents that influence S or S'. As we will see in Proposition 2.4, the sequence of powers of a Sarymsakov matrix converges to a matrix that has all its rows equal, and lead to consensus.

Scrambling matrices can also be defined using the consequent function. By definition, a matrix is scrambling if and only if for any two singleton S and S'

$$F(S) \cap F(S') \neq \emptyset. \tag{2.3}$$

By extension, Equation (2.3) also holds for any nonempty subsets of $\{1, \dots, n\}$, S and S' and not only singletons. From this, we can see that scrambling matrices are Sarymsakov. The converse is not true and the matrix

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & \frac{1}{2} & \frac{1}{2} \\ \frac{1}{3} & \frac{1}{3} & \frac{1}{3} \end{pmatrix}$$

is an example of a Sarymsakov non-scrambling matrix.

The product of any $n-1$ Sarymsakov matrices is scrambling [Sen92, Section 4]. Therefore, Sarymsakov matrices have the following properties:

- the class is closed under multiplication: the product of two Sarymsakov matrices is Sarymsakov [Sen92, Section 4],
- for any matrix size n, there is a length ℓ such that the product of ℓ $(n \times n)$ Sarymsakov matrices has a positive column.

We could wonder whether the Sarymsakov class is the largest that is closed under multiplication and such that any sufficiently long product has a positive column. The answer is known to be negative. In [XLC$^+$15, Section III, Subsection B], the authors construct slightly larger classes (the union of the Sarymsakov class and one single matrix) that have these two properties.

Proposition 2.2 (Section III, Subsection B of [XLC$^+$15])**.** *Let A be an $n \times n$ stochastic matrix such that A^2 is Sarymsakov and for any two disjoint nonempty subsets S and S' either*

$$F(S) \cap F(S') \neq \emptyset \tag{2.4}$$

or

$$|F(S) \cup F(S')| \geq |S \cup S'|.^1 \tag{2.5}$$

Then $\{S_{SAR,n} \cup A\}$ – where $S_{SAR,n}$ is the set of $n \times n$ Sarymsakov matrices – is closed under multiplication and there exists ℓ such that any product of length ℓ of matrices from $\{S_{SAR,n} \cup A\}$ has a positive column. Furthermore, there exist non-Sarymsakov matrices satisfying these conditions, so that $\{S_{SAR,n} \cup A\}$ is strictly larger than $S_{SAR,n}$. For example, the matrix

$$\begin{pmatrix} \frac{1}{2} & \frac{1}{2} & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 1 \end{pmatrix}$$

is not Sarymsakov but it satisfies conditions (2.4) *and* (2.5) *and*

$$A^2 = \begin{pmatrix} \frac{1}{4} & \frac{1}{4} & \frac{1}{2} \\ 0 & 0 & 1 \\ 0 & 0 & 1 \end{pmatrix}$$

is Sarymsakov.

[1] the inequality is not strict here, contrary to that in the definition of a Sarymsakov matrix

Finally, we introduce the notion of an SIA matrix. An SIA matrix is a matrix whose sequence of powers converge to a rank-one matrix. It is thus the equivalent of a contractive or almost contractive set in the case of a single matrix.

Definition 2.4 (SIA matrix [Wol63]). *A matrix A is* SIA *– Stochastic Indecomposable Aperiodic – if it is stochastic, the limit*

$$L = \lim_{n \to \infty} A^n$$

exists and all the rows of L are the same.

SIA matrices are sometimes called *regular* matrices [Sen81, Definition 4.3].

Proposition 2.3 (Characterization of SIA Matrices). *A stochastic matrix P is SIA if and only if there is p such that P^p has a positive column.*

Proof. Only if. Let P be SIA. We prove that there is p such that P^p has a positive column. The limit

$$\lim_{t \to +\infty} P^t$$

exists by definition of an SIA matrix. If there is no p such that P^p has a positive column, then $\lim_{t \to +\infty} P^t$ has no positive column. But, at the same time, all the rows of $\lim_{t \to +\infty} P^t$ are the same and the sum of the elements on each row is equal to 1, so that $\lim_{t \to +\infty} P^t$ has a positive column and we have a contradiction.

If. Let p be such that P^p is positive-column. By inclusions (2.6), P^p is SIA because positive-column matrices are SIA. Hence

$$L \triangleq \lim_{t \to +\infty} (P^p)^t$$

exists and has all rows equal. By stochasticity of P and the fact that L has all its rows equal,

$$L = LP = LP^2 = \cdots = LP^{p-1}$$

and therefore

$$\lim_{t \to +\infty} P^t = L$$

exists and has all rows equal and P is SIA. ∎

This simple characterization has several consequences. It means that only the pattern of positive elements determines if a matrix is SIA or not. As we discussed in Section 1.5, this means that many of the results developed in Chapter 8 apply in fact to row allowable matrices (Definition 1.3) and not only stochastic matrices.

The four matrix classes that we have defined are included in one another.

Proposition 2.4. *A positive-column matrix is scrambling, a scrambling matrix is Sarymsakov and a Sarymsakov matrix is SIA:*

$$S_{PC} \subset S_{SCR} \subset S_{SAR} \subset S_{SIA}, \tag{2.6}$$

where S_{PC}, S_{SCR}, S_{SAR} and S_{SIA} denote the sets of positive-column, scrambling, Sarymsakov and SIA matrices.

Proof. The first inclusion follows from the definitions and the second and third inclusions are proved in [Sen81, Section *Bibliography and Discussion to* §§*4.3–4.4*] and [XLC+15, Section II]. These inclusions are also depicted in Figure 2.1. ∎

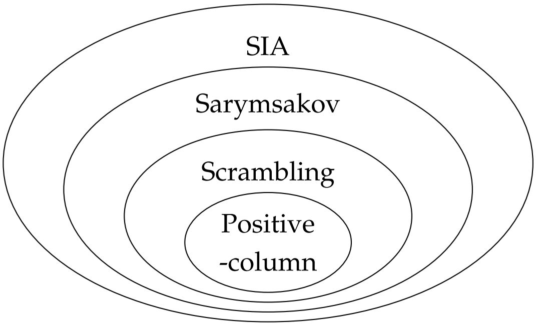


Figure 2.1: Relation between different classes of nonnegative matrices.

2.2 From SIA to a Positive Column

In Proposition 2.3, we have seen that a matrix is SIA if and only if there exists a power p such that P^p has a positive column. In this section we give an upper bound on p. Indeed, we show that the $n^2 - 3n + 3$th power of any arbitrary SIA matrix has a positive column and present a finer bound that depends on the number of columns that are eventually positive (Theorem 2.2). In Corollary 2.2, we observe that this result provides a method to check whether a stochastic matrix is SIA. We also prove (Corollary 2.3) a stronger bound in the case of matrices that have one 1 on each row and 0 everywhere else. We call these matrices *automaton matrices*.

We will that a matrix and a graph are *associated* if the matrix is the adjacency matrix of the graph.

We will rely on the *local exponents* of primitive matrices. Recall that a square matrix P is primitive if $P^t > 0$ (entrywise) for some natural t. The smallest such t is known as the *exponent* $\exp(P)$ of P [BR91, Section 3.5]. The local exponents are refinements of this characteristic : the kth *local exponent* $\exp_k(P)$ (with $1 \leq k \leq n$) of a primitive matrix P is the smallest power having at least k positive rows [BL90]. Observe that the first local exponent of P is the positive-column index of the transpose of P and the nth local exponent is $\exp(P)$. We will make use of the following theorem:

Theorem 2.1 ([BL90, Theorem 3.4])**.** *The largest value of the kth local exponent among primitive $n \times n$ matrices, with $n \geq 2$, is equal to $n^2 - 3n + k + 2$.*

With this theorem and the obvious local exponent in the case $n = 1$, we obtain that the first local exponent is

$$\exp(n, 1) = \begin{cases} n^2 - 3n + 3 & \text{if } n \geq 2 \\ 0 & \text{if } n = 1 \end{cases} . \tag{2.7}$$

Before stating our result, we define the notion of a column that is positive in sufficiently large powers of the matrix.

Definition 2.5 (Eventually positive columns)**.** *Let P be a stochastic matrix. We say that a column i is* eventually positive *if there is a power p such that the ith column of P^p is positive.*

We can already notice that the if P^p has a positive ith column $(P^p)_i$, then the ith column of P^{p+1} is equal to $P(P^p)_i$ and is also positive. Therefore, P^t has a positive ith column for any $t \geq p$.

Our bound on the power at which an SIA matrix has a positive column depends on the number of eventually positive columns. Essentially, if P has c eventually positive columns, it has a $c \times c$ primitive submatrix whose $\exp(c, 1)$th power has a positive column. This positive column then propagates to the remaining rows of the matrix $n - c$ steps ($n - c$ is an upper bound on the length of the shortest path from any node to any node of the primitive submatrix). The bound therefore *grows* with the number of eventually positive columns: it is the sum of a quadratic term in c and a linear term in $n - c$.

When considering the graph associated to an SIA matrix, we can observe that the number of eventually positive columns of the matrix is equal to the number of nodes that can be reached from all nodes. When the matrix is not SIA the number of eventually positive columns is zero.

Theorem 2.2. *Let P be an $n \times n$ SIA matrix, let c be the number of eventually positive columns of P and let*

$$\ell^* = exp(c, 1) + n - c = \begin{cases} c^2 - 4c + 3 + n & \text{if } c \geq 2 \\ n - 1 & \text{if } c = 1 \end{cases},$$

with $exp(c, 1)$ as defined in Equation (2.7) *Then P^{ℓ^*} has a positive column.*

Proof. We can assume that the first c columns are eventually positive, and let us partition P in blocks.

$$P = \begin{pmatrix} A & B \\ C & D \end{pmatrix},$$

with A having a size of $c \times c$ and blocks B, C and D having sizes $c \times (n - c)$, $(n - c) \times c$ and $(n - c) \times (n - c)$ respectively.

Claim: A is primitive and $B = 0$. Let p be such that P^p has c positive columns. If we partition the matrix P^p in blocks, as we did for P, we obtain

$$P^p = \begin{pmatrix} E & F \end{pmatrix},$$

with E being $n \times c$. We have that E is entrywise positive because we assumed that the first c columns of P^p are positive.

$$P^{p+1} = \begin{pmatrix} EA + FC & EB + FD \end{pmatrix}.$$

We prove now that $B = 0$. Assume to the contrary that b_{ij}, the element at position i, j in the matrix B, is positive. The jth column of EB is equal to

$$(EB)_j = \sum_i E_i b_{ij},$$

with E_i being the ith column of E. Because E is positive, the jth column of EB is positive, and similarly, the jth column of GB is positive. Therefore, the $(c + j)$th column of P^p is positive, and we have a contradiction because only the first c columns of P are eventually positive. Hence $B = 0$. Since $B = 0$, we have that $E = A^p$ and we can conclude that A is primitive.

Let us now compute $P^{\ell^*} = P^{n-c}P^{\exp(c,1)}$ Let us define G and H as below:

$$P^{n-c} = \begin{pmatrix} A^{n-c} & 0 \\ G & D^{n-c} \end{pmatrix}$$

and

$$P^{\exp(c,1)} = \begin{pmatrix} A^{\exp(c,1)} & 0 \\ H & D^{\exp(c,1)} \end{pmatrix}$$

The matrix $A^{\exp(c,1)}$ has a positive column thanks to Theorem 2.1 applied to the transpose of A. Let us assume that the ith column is positive.

We can notice that both G and A^{n-c} have a positive element on each row. A^{n-c} has a positive element on each row because P has a positive element on each row and the first rows of P are $\begin{pmatrix} A & 0 \end{pmatrix}$. And G has a positive element on each row because there is a path of length $n-c$ from each node $c+1, \ldots, n$ to some node $1, \ldots, c$ in the graph associated to P.

Therefore $GA^{\exp(c,1)}$ and $A^{n-c}A^{\exp(c,1)}$ have a positive ith column and therefore

$$P^{\ell^*} = \begin{pmatrix} A^{n-c}A^{\exp(c,1)} & 0 \\ GA^{\exp(c,1)} + D^{n-c}H & D^{\ell^*} \end{pmatrix}$$

has a positive column. ■

Example 2.1. *Let us consider the matrix*

$$P = \begin{pmatrix} 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \end{pmatrix}$$

This matrix has $c = 4$ eventually positive columns, and we can verify that its power $c^2 - 4c + 3 + n = 9$ is positive-column

$$P^9 = \begin{pmatrix} 0 & 1 & 2 & 1 & 0 & 0 \\ 1 & 1 & 1 & 2 & 0 & 0 \\ 2 & 3 & 1 & 1 & 0 & 0 \\ 1 & 3 & 3 & 1 & 0 & 0 \\ 1 & 2 & 1 & 0 & 0 & 0 \\ 1 & 1 & 0 & 1 & 0 & 0 \end{pmatrix}.$$

In this case, the bound of Theorem 2.2 is attained since the 8th power has no positive column:

$$A^8 = \begin{pmatrix} 1 & 2 & 1 & 0 & 0 & 0 \\ 0 & 1 & 2 & 1 & 0 & 0 \\ 1 & 1 & 1 & 2 & 0 & 0 \\ 2 & 3 & 1 & 1 & 0 & 0 \\ 1 & 1 & 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \end{pmatrix}.$$

This matrix can be partitioned in blocks as in the proof of Theorem 2.2:

$$P = \begin{pmatrix} A & 0 \\ C & D \end{pmatrix}$$

with

$$A = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 1 & 0 & 0 \end{pmatrix}.$$

A is a primitive matrix whose power exp$(c,1) = 7$ is positive-column and

$$P^{n-c} = P^2 = \begin{pmatrix} A^2 & 0 \\ CA + DC & D^2 \end{pmatrix} = \left(\begin{array}{cccc|cc} 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 & 0 & 0 \\ \hline 0 & 1 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 \end{array}\right)$$

is such that A^2 and $CA + DC$ have a positive element on each row. This is why P^9 is positive-column.

Corollary 2.1. *The $(n^2 - 3n + 3)$rd power of an $n \times n$ SIA matrix has a positive column.*

Proof. This is a consequence of the previous theorem and

$$\forall n \in \mathbb{N},\ k \in \{1, \ldots, n\},\ k^2 - 4k + 3 + n \leq n^2 - 3n + 3.$$

■

Corollary 2.2. *Let A be a stochastic matrix. Whether A is SIA can be decided in $\mathcal{O}(n^\omega \log(n))$ operations, where $\mathcal{O}(n^\omega)$ is the complexity of computing the product of two $n \times n$ matrices.*

Proof. A is SIA if and only if A^{n^2-3n+3} has a positive column and A^{n^2-3n+3} can be computed by performing $\mathcal{O}(\log(n))$ multiplications. ■

Corollary 2.3. *Let P be an $n \times n$ SIA automaton matrix. Then P^{n-1} has a positive column. Moreover, the value $n-1$ cannot be decreased (in general): for any $n \in \mathbb{N}$ there exists an $n \times n$ SIA automaton matrix whose $n-2$ power has no positive column.*

Proof. SIA automaton matrices have one eventually positive column. Therefore, Theorem 2.2 can be applied and $P^{\ell^*} = P^{n-1}$ has a positive column. The following matrix is an example of an SIA automaton matrix, whose $(n-2)$nd power has no positive column.

$$\begin{pmatrix} 1 & & & & \\ 1 & & & & \\ & 1 & & & \\ & & \ddots & & \\ & & & 1 & 0 \end{pmatrix}.$$

■

In Chapter 4 we will develop an algorithm to decide whether all infinite products of matrices taken from a set S of P-preserving matrices (Definition 1.2) converge to a matrix that has all its rows the same. Corollary 2.2 already solves this problem in the particular case where the set contains only one matrix and this matrix is stochastic.

Part I

Contractive Sets

Do all left-infinite products converge to a rank-one matrix?

Chapter 3

Invariant Polyhedra

This chapter presents an introduction to polyhedra and their classical properties. We analyze in particular the properties of the polyhedron

$$\mathcal{P} = \left\{ x \in \mathbb{R}^n \mid \frac{1}{2}(\max_i x_i - \min_i x_i) \leq 1 \right\},$$

that is invariant for stochastic matrices. This polyhedron will play a key role in the development of Chapters 4 and 5.

- In Section 3.1, we introduce classical notions and results about polyhedra.
- In Section 3.2, we study a polyhedron that is invariant for any stochastic matrix. We analyze the combinatorial structure of this polyhedron. The results of this section are preliminary results of [CHJ14], [CHJ15a] and [CHJ16]. The main results of these articles can be found in Chapters 4 and 5.

3.1 Polyhedra and Faces

We define some notions related to polyhedra. Unless mentioned otherwise, we will use the definitions and notations of [Sch98].

3.1.1 Polyhedra and Polytopes

We begin with the notions of polyhedron and polytope.

Definition 3.1 (Polyhedron). *We call a* polyhedron *a subset $\mathcal{Q}$ of $\mathbb{R}^n$ that is the intersection of a finite number of halfspaces or equivalently that can be defined by*

$$\mathcal{Q} = \{x \mid Ax \leq b\}.$$

We will use the letter $\mathcal{Q}$ when referring to any polyhedron and the letter $\mathcal{P}$ for the particular polyhedron $\mathcal{P} = \left\{x \mid \frac{1}{2}(\max_i x_i - \min_i x_i) \leq 1\right\}$. Note that a polyhedron is not necessarily bounded. The term *polytope* denotes a set that is the convex hull of a finite number of points:

$$\mathcal{Q} = \text{conv}\{x_1, x_2, \ldots, x_k\} = \left\{\sum_{i=1}^{k} \alpha_i x_i \mid \forall i,\ \alpha_i > 0,\ \sum_i \alpha_i = 1\right\},$$

and it can be proven [Sch98, Corollary 7.1c] that a set is a polytope if and only if it is a bounded polyhedron.

3.1.2 Faces, Open Faces and Facets

The notion of *face* generalizes the notions of facet, edge or vertex of a polyhedron to arbitrary dimensions. The polyhedron itself is also one of its faces. We give here the classical definition. We will see in Lemma 3.2 that it is equivalent to define faces as intersections of facets.

Definition 3.2 (Faces of a polyhedron). *A non-empty subset F of a polyhedron $\mathcal{Q}$ is called a* face *if $F = \mathcal{Q}$ or if it can be represented as $F = \mathcal{Q} \cap \left\{x \mid b^\top x = c\right\}$ where $b \in \mathbb{R}^n$ is non-zero, $c \in \mathbb{R}$ and*

$$\forall x \in \mathcal{Q},\ b^\top x \leq c.$$

We call a proper face a face that is not equal to $\mathcal{Q}$.

For example, the faces of a square are the square itself, the four corners and the four sides.

In order to define the notion of open face, we need to define the relative interior of a set. The concept of relative interior is a refinement of the concept of interior when dealing with low dimensional sets in higher dimensional spaces. As an example, the relative interior of the set

$$\{\alpha(1,0) + (1-\alpha)(0,1) \mid \alpha \in [0,1]\} \tag{3.1}$$

is

$$\{\alpha(1,0) + (1-\alpha)(0,1) \mid \alpha \in]0,1[\},$$

while the interior of (3.1) is the empty set. Formally, the *relative interior* $\mathrm{ri}(S)$ of a set S is the interior of this set in its affine hull, with the affine hull being the set of all affine combinations:

$$\mathrm{aff}(S) = \left\{ \sum_{i=1}^{k} \alpha_i x_i \mid k > 0,\ x_i \in S,\ \alpha_i \in \mathbb{R},\ \sum_{i=1}^{k} \alpha_i = 1 \right\}.$$

In our example the affine hull of the set (3.1) would be

$$\{\alpha(1,0) + (1-\alpha)(0,1) \mid \alpha \in \mathbb{R}\}.$$

We will also use the term *relative boundary* for the complement of the relative interior: $S = \mathrm{rb}(S) \cup \mathrm{ri}(S)$.

Definition 3.3 (Open face). *We call an* open face *the relative interior of a face. In particular, if the face is a single point, the corresponding open face is the face itself.*

The points $u_0, u_1, \ldots, u_n$ are said to be *affinely independent* if $u_1 - u_0, u_2 - u_0, \ldots, u_n - u_0$ are linearly independent. This definition does not depend on the choice of u_0.

Example 3.1. *The points* $(0,1), (1,0)$ *are affinely and linearly independent, while the points* $(0,1), (0,-1)$ *are affinely independent but not linearly independent.*

Definition 3.4 (Dimension of a face). *If a face contains* $d+1$ *affinely independent points, we call* d *the* dimension *of the face.*

In particular a face of dimension zero is called a *vertex*, a face of dimension $n-1$ is called a *facet*.

Definition 3.5 (Facet). *A face of dimension* $n-1$ *is a* facet. *For a facet, there is a unique hyperplane* $b^\top x = c$ *such that the facet is equal to* $\mathcal{Q} \cap \{x \mid b^\top x = c\}$. *We call* $b^\top x \leq c$ *the* facet constraint. *We say that the constraint is* active *at a point* x *when* $b^\top x = c$.

The next lemma shows how the facet inequalities define the polyhedron.

Lemma 3.1 (Theorem 8.1 in [Sch98]). *Let $\mathcal{Q} = \{x \mid Ax \leq b\}$ be a polyhedron and let $Ax \leq b$ be non-redundant constraints (no row of $Ax \leq b$ can be removed without changing $\{x \mid Ax \leq b\}$). A subset F of $\mathcal{Q}$ is a facet if and only if*

$$F = \{x \in \mathcal{Q} \mid a_i^\top x = b_i\}$$

for $a_i^\top x = b_i$ a row of $Ax = b$.

We now present a lemma that allows to represent the faces in terms of the inequalities that define the polyhedron. This also means the faces are intersections of facets, since the facets contraints are the nonredundant inequalities that define the polyhedron.

Lemma 3.2 (Section 8.3 in [Sch98]). *Let $\mathcal{Q} = \{x \mid Ax \leq b\}$ be a polyhedron. A non-empty subset F of $\mathcal{Q}$ is a face of $\mathcal{Q}$ if and only if it can be written as*

$$F = \{x \in \mathcal{Q} \mid A'x = b'\}, \tag{3.2}$$

where $A'x = b'$ is a subset of the rows of $Ax = b$.

Corollary 3.1. *Let $\mathcal{Q} = \{x \mid Ax \leq b\}$ be a polyhedron. A non-empty subset O of $\mathcal{Q}$ is an open face of $\mathcal{Q}$ if and only if it can be written as*

$$O = \{x \mid A'x = b',\ A''x < b''\},$$

where $A'x = b'$ is a subset of the rows of $Ax = b$ and $A''x = b''$ are the remaining rows.

Proof. Equation (3.2) can be rewritten

$$F = \{x \mid Ax \leq b,\ A'x = b'\},$$

or

$$F = \{x \mid A''x \leq b'',\ A'x = b'\}, \tag{3.3}$$

with $A''x = b''$ the rows of $Ax = b$ that are not in $A'x = b'$. With equation (3.3), we can see that the relative interior of F is

$$O = \mathrm{ri}(F) = \{x \mid Ax < b,\ A'x = b'\}.$$

■

The combination of Lemma 3.1 and Corollary 3.1 has interesting consequences. It means that two different open faces differ in at least one facet constraint. That is, there is a facet constraint $a_i^\top x \leq b_i$ such that points of one of the faces satisfy $a_i^\top x = b_i$ and points of the other satisfy $a_i^\top x < b_i$.

The second consequence is that a polyhedron $\mathcal{Q} = \{Ax \leq b\}$ decomposes into the disjoint union of its open faces: a point $x \in \mathcal{Q}$ is in exactly one open face. This face is given by $\{y \mid A'y = b',\ A''y < b''\}$ where A', b' is the largest subsystem of A, b such that $A'x = b'$ and A'', b'' are the remaining rows.

3.1.3 Invariant Polyhedra

We say that a matrix A leaves a polyhedron $\mathcal{Q}$ *invariant* if

$$A\mathcal{Q} \subseteq \mathcal{Q}$$

and we say that the matrix A *contracts* the polyhedron if

$$A\mathcal{Q} \subseteq \text{int}(\mathcal{Q}),$$

where $\text{int}(\mathcal{Q})$ denotes the interior of $\mathcal{Q}$.

Example 3.2. *The matrix*

$$A = \begin{pmatrix} \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & -\frac{1}{2} \end{pmatrix} \tag{3.4}$$

leaves the polyhedron $\mathcal{Q} = \{x \mid \|x\|_1 \leq 1\}$ invariant, while

$$A^2 = \frac{1}{2} \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

contracts it.

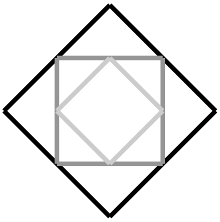


Figure 3.1: The polyhedron $\mathcal{Q}$ (in black) and its image by A (as defined in Equation 3.4, in gray) and A^2 (light gray).

In Chapter 4 we will show that iterative contraction of the polyhedron $\mathcal{P}$ leads to convergence to its *characteristic cone*, which is the set of directions in which the polyhedron is infinite. In order to formally define this notion of characteristic cone, we define a *cone* a set $\mathcal{C}$ that satisfies

$$\forall x \in \mathcal{C},\ \forall a \in \mathbb{R}^+,\ ax \in \mathcal{C}.$$

A *polyhedral cone* is a set that is both a cone and a polyhedron. Any polyhedron is the sum of a polyhedral cone and a polytope (a bounded polyhedron), as shown in the next theorem.

Theorem 3.1 (Decomposition Theorem for Polyhedra, Corollary 7.1b in [Sch98]). *A set $\mathcal{Q}$ is a polyhedron if and only if it can be written as*

$$\mathcal{Q} = \mathcal{S} + \mathcal{C},$$

where $\mathcal{C}$ is a polyhedral cone and $\mathcal{S}$ is a polytope.

The *characteristic cone* of a polyhedron $\mathcal{Q}$ is the set $\mathcal{S}$ of the previous theorem. It is defined as

$$\{y \mid \forall x \in \mathcal{Q},\ x + y \in \mathcal{Q}\}\,.$$

We define cone$\{x_1, \ldots, x_k\}$ as the set of *conic combinations* of a set $\{x_1, \ldots, x_k\}$

$$\text{cone}\{x_1, \ldots, x_k\} = \left\{ \sum_{i=1}^{k} \alpha_i x_i \mid \forall i,\ \alpha_i \geq 0 \right\}.$$

We end this section with a lemma that allows to check if a polyhedron is invariant by checking a finite number of vectors. It is a simple result that is not necessarily new but we did not find it in the literature.

Lemma 3.3. *Let*

$$\mathcal{Q} = \textit{conv}\{x_1, \ldots, x_k\} + \textit{cone}\{y_1, \ldots, y_t\}$$

be any polyhedron. A matrix A leaves $\mathcal{Q}$ invariant if and only if

$$Ax_1, \ldots, Ax_k \in \mathcal{Q} \tag{3.5}$$

and

$$Ay_1, \ldots, Ay_t \in \textit{cone}\{y_1, \ldots, y_t\}. \tag{3.6}$$

Proof. If: By definition any $x \in \mathcal{Q}$ can be written as

$$x = \sum_{i=1}^{k} \alpha_i x_i + \sum_{i=1}^{t} \beta_i y_i$$

for some $\alpha_i \geq 0$ satisfying $\sum_i \alpha_i = 1$ and some $\beta_i \geq 0$. Conditions (3.5) and (3.6) guarantee that

$$Ax = \sum_{i=1}^{k} \alpha_i A x_i + \sum_{i=1}^{t} \beta_i A y_i \in \mathcal{Q}.$$

Only if: $Ax_i \in \mathcal{Q}$ because $x_i \in \mathcal{Q}$ and $\mathcal{Q}$ is invariant.

We prove that for any i, $Ay_i \in \text{cone}\{y_1, \dots, y_t\}$. Let us take $x_j \in \{x_1, \dots, x_k\}$. We have

$$\forall c \geq 0,\ x_j + c y_i \in \mathcal{Q}$$

and, by invariance of $\mathcal{Q}$:

$$\forall c \geq 0,\ Ax_j + Acy_i \in \mathcal{Q}.$$

Hence

$$\forall c \geq 0,\ \frac{Ax_j}{c} + Ay_i \in \frac{\mathcal{Q}}{c} = \frac{\text{conv}\{x_1, \dots, x_k\}}{c} + \text{cone}\{y_1, \dots, y_t\}$$

and Ay_i is in the closure of $\text{cone}\{y_1, \dots, y_t\}$ which is equal to $\text{cone}\{y_1, \dots, y_t\}$. ■

3.1.4 Polarity

Definition 3.6 (Polar of a polyhedron)**.** *The polar of a polyhedron $\mathcal{Q} \subset \mathbb{R}^n$, denoted $\mathcal{Q}^*$ is the set*

$$\mathcal{Q}^* = \left\{ z \in \mathbb{R}^n \mid \forall x \in \mathcal{Q},\ z^\top x \leq 1 \right\}.$$

If a polyhedron is invariant for a matrix A, its polar $\mathcal{Q}^*$, which is also a polyhedron, is invariant for its transpose:

$$A\mathcal{Q} \subseteq \mathcal{Q} \Leftrightarrow A^\top \mathcal{Q}^* \subseteq \mathcal{Q}^*. \tag{3.7}$$

The polar can be computed using the following proposition.

Proposition 3.1 (Theorem 9.1 in [Sch98]). *Let $\mathcal{Q}$ be a polyhedron that contains the origin:*

$$\mathcal{Q} = conv\{0, x_1, \ldots, x_k\} + cone\{y_1, \ldots, y_t\}.$$

Then

$$\mathcal{Q}^* = \left\{ z \in \mathbb{R}^n \mid \forall i \in \{1, \ldots, k\},\ z^\top x_i \leq 1,\ and\ \forall j \in \{1, \ldots, t\},\ z^\top y_i \leq 0 \right\}.$$

3.1.5 Centrally Symmetric Polyhedra and Seminorms

When dealing with matrix products or linear systems it is sometimes natural to use an invariant polyhedron $\mathcal{Q}$ that is symmetric around the origin

$$-\mathcal{Q} = \mathcal{Q}.$$

Indeed, if a polyhedron $\mathcal{Q}$ is invariant with respect to a matrix A, then the polyhedron $\mathcal{Q} \cap -\mathcal{Q}$ is symmetric around the origin and it is also invariant with respect to the matrix A:

$$A(\mathcal{Q} \cap -\mathcal{Q}) \subseteq \mathcal{Q} \cap -\mathcal{Q}.$$

Therefore, a matrix (or a set of matrices) that has an invariant polyhedron has one that is symmetric around the origin. We will call *centrally symmetric* the polyhedra that are symmetric around the origin.

A centrally symmetric polyhedron $\mathcal{Q}$ that contains the origin in its interior can be seen as the unit ball of some seminorm $\|x\|_{\mathcal{Q}}$ that we define as the smallest nonnegative integer c such that $x \in c\mathcal{Q}$.

$$\|x\|_{\mathcal{Q}} = \min_{c \in \mathbb{R}^+} c \text{ such that } x \in c\mathcal{Q}$$

We can verify that the function $\|.\|_{\mathcal{Q}}$ satisfy the conditions that define a seminorm.

- $\forall x \in \mathbb{R}^n,\ a \in \mathbb{R},\ \|ax\| = |a|\|x\|$. This condition holds because of the symmetry of $\mathcal{Q}$.

- $\forall x, y \in \mathbb{R}^n,\ \|x + y\| \leq \|x\| + \|y\|$. This condition holds because of the convexity of $\mathcal{Q}$.

Invariance of the polyhedron corresponds to the seminorm being nonincreasing and contraction of the polyhedron corresponds to the seminorm being decreasing. The seminorm can be used as a Lyapunov function to prove convergence to the characteristic cone, while the unit ball will be the level-set of this Lyapunov function. In Chapters 4 and 5 we will deal with the polyhedron $\mathcal{P}$ that is the unit ball of a seminorm.

3.2 Invariant Polyhedron for Stochastic Matrices

In this section, we study the polyhedron $\mathcal{P} = \left\{x \mid \frac{1}{2}(\max_i x_i - \min_i x_i) \leq 1\right\}$. We show that it is invariant for all stochastic matrices. We characterize its faces and we count them. Finally, we discuss other invariant polyhedra that are invariant for stochastic matrices.

The structure of $\mathcal{P}$ and its number of faces will play a key role in the next chapters. In Chapter 4, we will see that the complexity of our algorithm to check if a set is contractive depends on the number of faces of $\mathcal{P}$, while in Chapter 5, the number of faces of a fixed dimension will appear instrumental.

The set

$$\mathcal{P} = \left\{x \mid \frac{1}{2}(\max_i x_i - \min_i x_i) \leq 1\right\} \tag{3.8}$$

that we have already introduced, is a polyhedron. This is because the constraint can be decomposed into multiple linear constraints:

$$\mathcal{P} = \bigcap_{ij} \left\{x \mid \frac{1}{2}(x_i - x_j) \leq 1\right\}. \tag{3.9}$$

Another way to write $\mathcal{P}$ is as the sum of a the unit ball of the infinity norm $\mathcal{B}_\infty = \{x \mid \|x\|_\infty \leq 1\}$ and the subspace spanned by $\mathbf{1}$.

$$\mathcal{P} = \mathcal{B}_\infty + \text{cone}\{\mathbf{1}, -\mathbf{1}\}. \tag{3.10}$$

This equality holds because

$$\mathcal{B}_\infty = \left\{x \mid \max_i x_i \leq 1,\ \min_i x_i \geq -1\right\}.$$

Equation (3.10) allows to compute easily the polar of $\mathcal{P}$. Indeed,

$$\mathcal{B}_\infty = \text{conv}(\{1, -1\}^n)$$

and

$$\begin{aligned} \mathcal{P} &= \text{conv}(\{1, -1\}^n) + \text{cone}\{\mathbf{1}, -\mathbf{1}\} \qquad (3.11)\\ &= \text{conv}(\{1, -1\}^n \backslash \{\mathbf{1}, -\mathbf{1}\}) + \text{cone}\{\mathbf{1}, -\mathbf{1}\}. \qquad (3.12)\end{aligned}$$

We can now compute the polar of $\mathcal{P}$. A direct application of Proposition 3.1 on $\mathcal{P} = \text{conv}(\{1,-1\}^n) + \text{cone}\{\mathbf{1}, -\mathbf{1}\}$ (Equation (3.11)) gives

$$\begin{aligned}\mathcal{P}^* &= \left\{z \mid \forall v \in \{1,-1\}^n,\ z^\top v \leq 1,\ z^\top \mathbf{1} \leq 0,\ z^\top(-\mathbf{1}) \leq 0\right\} \\ &= \left\{z \mid z^\top \mathbf{1} = 0,\ \|z\|_1 \leq 1\right\}.\end{aligned}$$

The polar $\mathcal{P}^*$ is represented in Figure 3.2. It is the intersection of the unit ball of the norm $\|.\|_1$ and a hyperplane. Its vertices are the vector

$$\begin{pmatrix}\frac{1}{2} & -\frac{1}{2} & 0 & \dots & 0\end{pmatrix}$$

and all the vectors obtained by permutations of its elements.

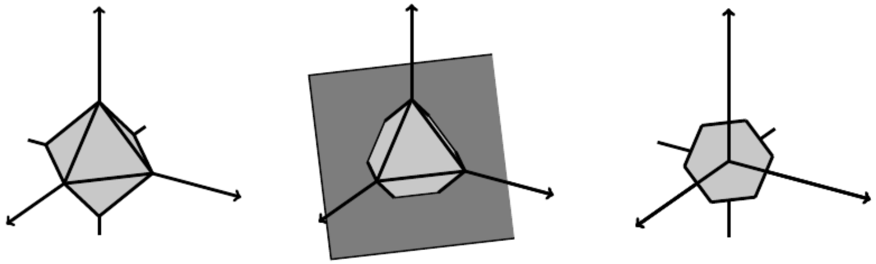


Figure 3.2: Left: the unit ball of the norm $\|.\|_1$. Center: the unit ball of the norm $\|.\|_1$ and the plane $\sum_i x_i = 0$. Right: the polyhedron $\mathcal{P}^*$

Although it can be observed directely from the definition of $\mathcal{P}$, it is maybe easier to see from the polar that stochastic matrices leave $\mathcal{P}$ invariant. Indeed, let A be a stochastic matrix

$$\begin{pmatrix}\frac{1}{2} & -\frac{1}{2} & 0 & \dots & 0\end{pmatrix} A = \frac{1}{2}(a_1^\top - a_2^\top),$$

with $a_i^\top$ the ith row of A. Similarly, for any other vertex v of the polar of $\mathcal{P}$ $\left\{z \mid z^\top \mathbf{1} = 0,\ \|z\|_1 \leq 1\right\}$,

$$v^\top A = \frac{1}{2}(a_i^\top - a_j^\top)$$

for some $i, j \in \{1, \dots, n\}$. Since $\forall k,\ a_k^\top \mathbf{1} = 1$,

$$\frac{1}{2}(a_i^\top - a_j^\top)\mathbf{1} = 0 \tag{3.13}$$

and

$$\left\|\frac{1}{2}(a_i^\top - a_j^\top)\right\|_1 \leq 1. \tag{3.14}$$

Hence

$$v^\top A \in \{z \mid z^\top \mathbf{1} = 0,\ \|z\|_1 \le 1\}$$

and the polyhedron $\mathcal{P}^*$ is invariant with respect to $A^\top$. By (3.7), the stochastic matrix A leaves $\mathcal{P}$ invariant.

We can observe that a matrix A leaves $\mathcal{P}$ invariant if and only if Conditions (3.13) and (3.14) are satisfied. This is not only the case of stochastic matrices. Examples of non-stochastic matrices that satisfy these conditions, and therefore leave $\mathcal{P}$ invariant, include

$$\begin{pmatrix} -1 & 0 \\ -\frac{1}{2} & -\frac{1}{2} \end{pmatrix},$$

which is the opposite of a stochastic matrix,

$$\begin{pmatrix} 2 & 0 \\ 1 & 1 \end{pmatrix},$$

which is the sum of a stochastic matrix and a rank-one term of the form $\mathbf{1}y^\top$ and also

$$A = \frac{1}{2}\begin{pmatrix} 1 & -1 & 0 & 0 \\ -1 & 1 & 0 & 0 \\ 0 & 0 & 1 & -1 \\ 0 & 0 & -1 & 1 \end{pmatrix},$$

which is the convex combination of a stochastic matrix and the opposite of a stochastic matrix. The results that we will develop in Chapters 4 and 5 hold for all matrices that leave $\mathcal{P}$ invariant or equivalently all matrices that satisfy Conditions (3.13) and (3.14).

3.2.1 Combinatorial Structure of $\mathcal{P}$

Thanks to Lemma 3.1, we know that the facets of the polyhedron $\mathcal{P}$ are the sets

$$\mathcal{P} \cap \left\{x \mid \frac{1}{2}(x_i - x_j) = 1\right\}$$

for $i \neq j$. The next lemma describes the faces of arbitrary dimension.

Lemma 3.4. *The set F is a proper face of $\mathcal{P}$ if and only if it is equal to*

$$F = \left\{x \in \mathcal{P} \mid \forall i \in S_1, j \in S_2,\ \frac{1}{2}(x_i - x_j) = 1\right\}, \tag{3.15}$$

with S_1 and S_2 disjoints nonempty subsets of $\{1, \ldots, n\}$. The dimension of F is given by

$$d = n - |S_1 \cup S_2| + 1.$$

Proof. The set

$$F = \left\{ x \in \mathcal{P} \mid \forall i \in S_1, j \in S_2,\ \frac{1}{2}(x_i - x_j) = 1 \right\}$$

is the intersection of the facets

$$\mathcal{P} \cap \left\{ x \mid \frac{1}{2}(x_i - x_j) = 1 \right\}$$

for which $i \in S_1$ and $j \in S_2$. By Lemma 3.2, we conclude that F is a face. On the other hand, if a set is a proper face, then it is the intersection of facets of $\mathcal{P}$ and can be written as (3.15).

The elements x_i with $i \notin (S_1 \cup S_2)$ can take any value between $\min_j x_j$ and $\max_j x_j = \min_j x_j + 2$. These, and the direction **1** correspond to the

$$n - |S_1 \cup S_2| + 1$$

dimensions of F. ■

Corollary 3.2. *The number of faces of $\mathcal{P}$ is $3^n - 2^{n+1} + 2$.*

Proof. We count the number of ordered pairs of disjoints nonempty subsets (S_1, S_2) of $\{1, \ldots, n\}$. The number of ordered pairs of disjoints (possibly empty) subsets (S_1, S_2) of $\{1, \ldots, n\}$ is 3^n. The number of these pairs in which $S_1 = \emptyset$ is 2^n, as is the number of pairs in which $S_2 = \emptyset$. There is 1 pair in which both S_1 and S_2 are empty. The total number of proper faces is thus $3^n - 2^{n+1} + 1$, and the number of faces (including $\mathcal{P}$) is $3^n - 2^{n+1} + 2$. ■

Corollary 3.3. *The number of faces of dimension $1 \leq d \leq n-1$ of $\mathcal{P}$ is*

$$f_d = \binom{n}{d-1}(2^{n-d+1} - 2).$$

There are no faces of dimension 0. There is one face of dimension n.

Proof. The number of pairs of dimension $1 \leq d \leq n-1$ is the number of ordered pairs of disjoints nonempty subsets (S_1, S_2) of $\{1, \ldots, n\}$ that satisfy

$$|S_1 \cup S_2| = n - d + 1.$$

The number of ways to choose the elements that belong to $S_1 \cup S_2$ is

$$\binom{n}{d-1} \tag{3.16}$$

and the number of ways to choose which of these elements belong to S_1 is

$$(2^{n-d+1} - 2). \tag{3.17}$$

The combination of (3.16) and (3.17) yields

$$f_d = \binom{n}{d-1}(2^{n-d+1} - 2).$$

There is no face of dimension 0 because for any $x \in \mathcal{P}$, $x + \mathbf{1} \in \mathcal{P}$. There is one face of dimension n by definition of a face. ■

Example 3.3. *When $n = 2$, $\mathcal{P} = \left\{x \mid \frac{1}{2}|x_1 - x_2| \leq 1\right\}$ as represented in Figure 3.3. The faces are*

- $\{x \mid \frac{1}{2}(x_1 - x_2) = 1\}$, *which is both a closed and an open face (because it is equal to its relative interior),*
- $\{x \mid \frac{1}{2}(x_2 - x_1) = 1\}$, *which is both a closed and an open face,*
- $\mathcal{P}$, *for which the corresponding open face is int($\mathcal{P}$).*

There are thus two faces of dimension 1, as predicted by Corollary 3.3:

$$f_1 = \binom{2}{1-1}(2^{2-1+1} - 2) = 2.$$

There are three faces in total as predicted by the Corollary 3.2: $3^2 - 2^3 + 2 = 3$.

Example 3.4. *The polyhedron $\mathcal{P}_3$, that we define as $\mathcal{P}$ for $n = 3$ is represented in Figure 3.4. Its faces of dimension 2 are the sets*

$$\left\{x \in \mathcal{P} \mid \frac{1}{2}(x_i - x_j) = 1\right\},$$

with $i \neq j$. The corresponding open faces are

$$\left\{x \mid \frac{1}{2}(x_i - x_j) = 1,\ x_j < x_k < x_i\right\},$$

with $i \neq j \neq k$ and $i \neq k$.

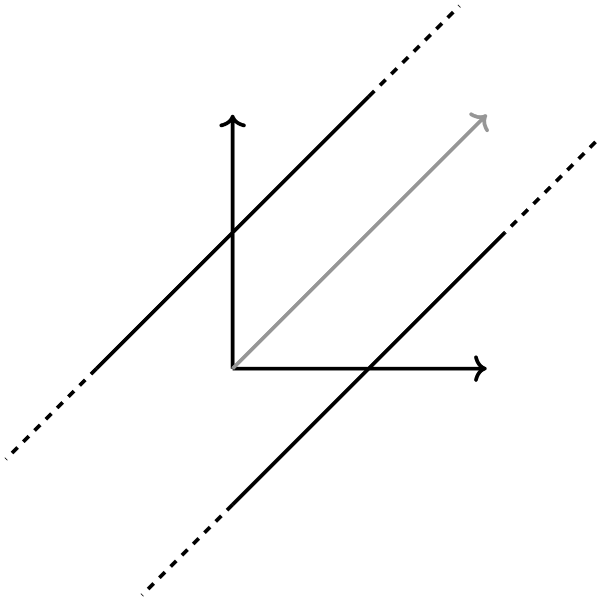


Figure 3.3: The polyhedron $\mathcal{P}$ in dimension 2.

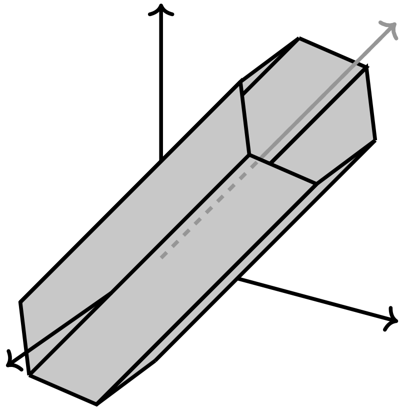


Figure 3.4: The polyhedron $\mathcal{P}$ in dimension 3.

The faces of dimension 1 are the sets

$$\left\{ x \mid \frac{1}{2}(x_i - x_j) = 1,\ x_i = x_k \right\}$$

and

$$\left\{ x \mid \frac{1}{2}(x_i - x_j) = 1,\ x_j = x_k \right\},$$

with i, j and k all different. These faces are both closed and open.

In turn, $\mathcal{P}_3$ has

- *6 two-dimensional faces,*
- *6 one-dimensional faces,*
- *$\mathcal{P}_3$.*

Corollary 3.3 gives $f_2 = \binom{3}{2-1}(2^{3-2+1}-2) = 6$ and $f_1 = \binom{3}{1-1}(2^{3-1+1}-2)$. There are thus 13 faces as predicted by the Corollary 3.2: $3^3 - 2^4 + 2 = 13$.

From the definition of $\mathcal{P}$, or maybe more clearly from Equation (3.9), we can see that the polyhedron $\mathcal{P}$ is symmetric around the origin:

$$\mathcal{P} = -\mathcal{P}.$$

$\mathcal{P}$ is the unit ball of the seminorm

$$\|x\|_{\mathcal{P}} \triangleq \frac{1}{2}(\max_i x_i - \min_i x_i).$$

3.2.2 Other Invariant Polyhedra for Stochastic Matrices

We end this section by mentioning that $\mathcal{P}$ is not the only polyhedron that is invariant for stochastic matrices. In fact a stochastic matrix A satisfies

$$\max_i (Ax)_i \leq \max_i x_i$$

and

$$\min_i (Ax)_i \geq \min_i x_i.$$

Therefore it leaves the unit ball of the infinity norm

$$\{x \mid \|x\|_\infty \leq 1\} = \left\{x \mid \max_i x_i \leq 1,\ \min_i x_i \leq 1\right\}.$$

The fact that a stochastic matrix leaves the unit ball of the infinity norm invariant is another way to see that a stochastic matrix leaves $\mathcal{P}$ invariant, because $\mathcal{P} = \mathcal{B}_\infty + \text{cone}\{\mathbf{1}, -\mathbf{1}\}$. Stochastic matrices also leave the nonnegative orthant invariant, because the nonnegative orthant can be seen as

$$\{x \mid \forall i,\ x_i \geq 0\} = \{x \mid \min_i x_i \geq 0\}.$$

In the next chapters, we will work with $\mathcal{P}$ as an invariant polyhedron for stochastic matrices. The reasons to use this polyhedron instead of another are the following.

- In applications such as consensus, we are interested in convergence to the space span$\{\mathbf{1}\}$ which is the characteristic cone of $\mathcal{P}$. We will see that iterative contraction of $\mathcal{P}$ will precisely lead to the convergence to this characteristic cone.
- $\mathcal{P}$ has less faces than the unit ball of the infinity norm. This number of faces will directly influence the complexity of our algorithms and the quality of our bounds.
- More matrices leave $\mathcal{P}$ invariant than $\mathcal{B}_\infty$, so that the result we will obtain will apply to a larger class of matrices.
- Our techniques do not work well with the nonnegative orthant to analyze consensus. In Chapter 4, we will see that, if, for a given set, there is an infinite product that does not converge to consensus, then there is a finite product P and an open proper face F of $\mathcal{P}$ such that

  $$PF \subseteq F. \tag{3.18}$$

  This implies that the closure of F contains a point x such that $Px = \lambda x$ with $|\lambda| = 1$ and that $\ldots PPPx$ never converges to consensus. The same reasoning does not work for the nonnegative orthant. Indeed, all faces of the nonnegative orthant contain the origin in their closure, and the origin is a consensus. The fact that a face F and a product P satisfy Equation (3.18) does not necessarily imply the existence of an infinite product that does not converge to consensus. When different products P_1 and P_2 satisfy Equation (3.18), deciding whether it is possible to construct a sequence that starts in F and that does not converge to zero/consensus is actually a hard problem. It is indeed equivalent to deciding whether the joint spectral radius of a projection of the set $\{P_1, P_2\}$ is strictly smaller than 1, a problem that is not known to be decidable in general [Jun09].

  In Chapter 6, however, we will see that our techniques applied to the nonnegative orthant are well suited to analyze a different problem, namely the primitivity of matrix sets.

Chapter 4

Complexity of Deciding Consensus

We define a *contractive set* as a set S of matrices that leave $\mathcal{P}$ invariant and for which all sufficiently long products contract[1] $\mathcal{P}$. This definition is a generalization of the notion of *consensus set* defined in [BO14], in the sense that a set of *stochastic matrices* is a contractive set if and only if it is a consensus set. We develop an algorithm that determines whether a set is contractive. We show that our algorithm has a complexity that is singly exponential in n the dimension of the matrices. This algorithm is an improvement both in complexity and generality over the doubly exponential decision procedure developed for consensus sets in [BO14].

This chapter is based on [CHJ14, CHJ15a] and is organized as follows.

- In Section 4.1, we introduce the notion of contractive set. We define the decision problem of determining whether a set is contractive.
- In Section 4.2, we develop an algorithm that decides whether a set of matrices is contractive.
- In Section 4.3, we discuss the complexity of this algorithm.
- In Section 4.4, we analyze the particular case of sets of two undirected stochastic matrices. We prove that deciding if a set is contractive can be done in polynomial time in this case.

[1] We recall that a matrix A contracts a polyhedron $\mathcal{Q}$ if $A\mathcal{Q} \in \mathrm{int}(\mathcal{Q})$.

4.1 Contractive Sets

We study finite sets of P-preserving matrices (matrices that leave the polyhedron $\mathcal{P}$ invariant), and we wonder whether all sufficiently long products of matrices from a set contract $\mathcal{P}$.

Assumption 4.1. *All matrices of S are P-preserving:*

$$\forall A \in S,\ A\mathcal{P} \subseteq \mathcal{P}.$$

As mentioned in Chapter 3, Assumption 4.1 is weaker than requiring the matrices to be stochastic. The assumption of stochasticity is common for linear discrete-time consensus systems and it is made, for example, in [BO14].

Definition 4.1 (Contractive set)**.** *We say that a set S of $n \times n$ matrices that satisfies Assumption 4.1 is a* contractive set *if there exist $\ell \in \mathbb{R}$ such that all products of ℓ or more matrices from S contract $\mathcal{P}$.*

This definition is consistent with the informal definition of a contractive set of stochastic matrices given in the introduction. Indeed, we will see in Proposition 4.3 that, for a matrix set S that satisfies Assumption 4.1 and $\forall A \in S,\ A\mathbf{1} = \mathbf{1}$, all left-infinite products converge to a rank-one matrix of the form $\mathbf{1}y^\top$ if and only if S is contractive. Additionally, Proposition 4.2 shows in the slightly more general case of a matrix set S that only satisfies Assumption 4.1, that all left-infinite products contract the space into the set span$\{\mathbf{1}\}$ if and only if S is contractive.

The next proposition shows that the geometric condition of a stochastic matrix A contracting $\mathcal{P}$ is equivalent to the algebraic condition of A being scrambling. In [Paz71, Section A.4 of Chapter II], a set S of stochastic matrices is said to satisfy the H_2 condition if all sufficiently long products of matrices from S are scrambling. The following proposition thus establishes the equivalence between the H_2 condition and contractivity, in the case of sets of stochastic matrices.

Proposition 4.1. *A stochastic matrix A contracts $\mathcal{P}$ if and only if it is scrambling.*

Proof. For a given row i of A, let us call S_i the set of columns k such that $a_{ik} > 0$. Recall that A contracts $\mathcal{P}$ if and only if

$$\forall i,j,\ \forall x \in \partial\mathcal{P},\ (Ax)_i - (Ax)_j < \max_k x_k - \min_k x_k.$$

We can compute, for arbitrary $i, j \in \{1, \ldots, n\}$ and $x \in \partial\mathcal{P}$:

$$(Ax)_i - (Ax)_j = \sum_{k \in S_i} a_{ik} x_k - \sum_{k \in S_j} a_{jk} x_k. \tag{4.1}$$

On the one hand, if A is not scrambling, then there are rows i and j such that

$$S_i \cap S_j = \emptyset$$

and one can see in (4.1) that if we take $x_k = 1$ when $k \in S_i$, $x_k = -1$ when $k \in S_j$ and $x_k = 0$ when $k \notin (S_i \cup S_j)$, we obtain

$$(Ax)_i - (Ax)_j = 2 = \max_k x_k - \min_k x_k.$$

Hence A does not contract $\mathcal{P}$.

On the other hand, if A is scrambling, then

$$\begin{aligned}(Ax)_i - (Ax)_j &= \sum_{k \in S_i} a_{ik} x_k - \sum_{k \in S_j} a_{jk} x_k \\ &< \max_k x_k \sum_{k \in S_i} a_{ik} - \min_k x_k \sum_{k \in S_j} a_{jk} \\ &= \max_k x_k - \min_k x_k,\end{aligned}$$

the strict inequality being because there is $p \in S_j \cap S_i$ for which not both $x_p = \max_k x_k$ and $x_p = \min_k x_k$ can hold (we recall that $x \in \partial\mathcal{P}$, hence $max_k x_k \neq \min_k x_k$). The matrix A therefore contracts $\mathcal{P}$, which concludes the proof. ■

Another consequence of this proposition is that a set of stochastic matrices S is contractive if and only if it is a consensus set. A consensus set is defined in [BO14] as a set of stochastic matrices such that the associated consensus system 1.3 converges for any initial condition and any sequence of matrices taken from the set. The authors of [BO14] proved that this condition is equivalent to the H_2 condition of [Paz71]. They proved that deciding whether a set of stochastic matrices is a consensus set is NP-hard but decidable. The decision procedure that they constructed to prove decidability has a doubly exponential time complexity.

Decision problem 4.1. *Given a set S of $n \times n$ matrices. Is S contractive?*

In the case of sets stochastic matrices, the results of [BO14] directly apply to Problem 4.1. Hence, we already know that the problem is decidable but

NP-hard in this case. In Section 4.2, we will develop a new decision algorithm and in Section 4.3, we will prove that its time complexity is singly exponential in n the dimension of the matrices. It will therefore be an improvement over the decision procedure of [BO14].

A positive answer to Problem 4.1 has many consequences. If the answer is positive, all left-infinite products send $\mathcal{P}$ into its characteristic cone span$\{\mathbf{1}\}$ (a formal statement will be provided in Proposition 4.2). If additionally, all matrices from S satisfy $A\mathbf{1} = \mathbf{1}$, then all left-infinite products actually have a limit and this limit is a rank-one matrix (as we will see in Proposition 4.3).

4.1.1 Properties of Contractive Sets

This section is dedicated to the properties of contractive sets. We show the relation between a positive answer to Problem 4.1 and convergence to a rank-one matrix.

In Chapter 3, we defined the seminorm $\|x\|_{\mathcal{P}} = \frac{1}{2}(\max_i x_i - \min_i x_i)$. We now define its induced matrix seminorm.

Definition 4.2. *The vector seminorm $\|.\|_{\mathcal{P}}$ induces the following matrix seminorm*

$$\|A\|_{\mathcal{P}} = \sup_{\|x\|_{\mathcal{P}}=1} \|Ax\|_{\mathcal{P}}$$

With this definition, Assumption 4.1 is equivalent to

$$\forall A \in S,\ \|A\|_{\mathcal{P}} \leq 1.$$

We also note that

$$\|A\|_{\mathcal{P}} = 0$$

if and only if

$$A = \mathbf{1}y^\top$$

for some y.

For a finite set $S = \{A_1, \dots, A_m\}$, we denote by $P_{\sigma(t)}$ the product

$$A_{\sigma(t)} \dots A_{\sigma(1)} A_{\sigma(0)},$$

with $\sigma : \mathbb{N} \mapsto \{1, \dots, m\} : t \mapsto \sigma(t)$ is an infinite sequence of indices. Let Σ denote the set of such sequences and let Σ_t denote the set of sequences limited to length t. The following proposition shows that if S is a contractive set, then for any sequence σ, the rows of $P_{\sigma(t)}$ become arbitrary similar as t

increases: $\lim_{t\to\infty}\|P_{\sigma(t)}\|_{\mathcal{P}} = 0$. We also prove in the same proposition that this convergence is exponential, as we will need this fact in the proof of a later proposition.

Despite the fact that the rows of $P_{\sigma(t)}$ become arbitrarily close to each other, the limit $\lim_{t\to\infty} P_{\sigma(t)}$ does not always exist, and we will prove convergence of the sequence $(P_{\sigma(t)})_t$ under stronger assumptions in Proposition 4.3.

Proposition 4.2. *Let S be a set of matrices that satisfies Assumption 4.1. Let the set S be contractive (or, equivalently, the answer to Problem 4.1 is positive). Then for any sequence $\sigma \in \Sigma$,*

$$\lim_{t\to\infty}\|P_{\sigma(t)}\|_{\mathcal{P}} = 0.$$

Additionally, there exist $C \in \mathbb{R}^+$ and $r < 1$ such that for any sequence σ:

$$\forall t,\ \|P_{\sigma(t)}\|_{\mathcal{P}} \leq Cr^t.$$

Proof. Let ℓ be, as in Definition 4.1, such that all products of ℓ or more matrices from S contract $\mathcal{P}$. Let us define

$$r_1 = \max_{\sigma\in\Sigma_\ell}\|P_{\sigma(\ell)}\|_{\mathcal{P}}.$$

We can write a $\max$ in the definition of r_1 because it is an optimization problem over a finite set Σ_ℓ. We have that $r_1 < 1$ because S is contractive. We have

$$\begin{aligned}\|P_{\sigma(t)}\|_{\mathcal{P}} &\leq \|A_{\sigma(t)}\dots A_{\sigma(t-\ell+1)}\|_{\mathcal{P}}\|P_{\sigma(t-\ell)}\|_{\mathcal{P}}\\ &\leq r_1\|P_{\sigma(t-\ell)}\|_{\mathcal{P}}\\ &\leq r_1^{\lfloor t/\ell\rfloor}\\ &\leq \frac{1}{r_1}\left(r_1^{\frac{1}{\ell}}\right)^t\end{aligned}$$

so that

$$\lim_{t\to\infty}\|P_{\sigma(t)}\|_{\mathcal{P}} = 0$$

and $C = \frac{1}{r_1}$ and $r = r_1^{\frac{1}{\ell}}$ yields the exponential convergence. ∎

Let us take, for example, the set made of a single matrix

$$\left\{A = \begin{pmatrix} 2 & 0\\ \frac{3}{2} & \frac{1}{2}\end{pmatrix}\right\}.$$

Since the set is a singleton, the only infinite product that can be generated is the sequence of powers of A:

$$A^t = \begin{pmatrix} 2^t & 0 \\ 1-(\frac{1}{2})^t & (\frac{1}{2})^t \end{pmatrix}. \tag{4.2}$$

As t increases, the rows of A^t become arbitrarily close to each others, in the sense that

$$\|A^t\|_{\mathcal{P}} = \frac{1}{2^t}$$

converges to zero. However the sequence of powers of A (Equation (4.2)) does not converge to a particular matrix. In the next proposition, we prove the convergence of all left-infinite products, under the additional condition that the matrices satisfy $A\mathbf{1} = \mathbf{1}$. We note that stochastic matrices satisfy $A\mathbf{1} = \mathbf{1}$ but that there are other matrices that satisfy $A\mathbf{1} = \mathbf{1}$ and Assumption 4.1. An example is given by

$$\begin{pmatrix} 2 & -1 \\ 1 & 0 \end{pmatrix}.$$

Proposition 4.3. *Let S be a set of matrices that satisfies Assumption 4.1 and such that $A\mathbf{1} = \mathbf{1}$ for any $A \in S$. If the set S is contractive, then for any sequence $\sigma \in \Sigma$, the corresponding left-infinite product exists and converges to a rank-one matrix of the form*

$$\lim_{t\to\infty} P_{\sigma(t)} = \mathbf{1}y^\top$$

for some y.

Proof. By Proposition 4.2, if the limit $\lim_{t\to\infty} P_{\sigma(t)}$ exists, it can only be equal to $\mathbf{1}y^\top$ for some y. We now prove that the limit exists. We start by bounding the difference between two successive iterates. Using $Q \triangleq I - \frac{\mathbf{1}\mathbf{1}^\top}{n}$ and the assumption that $A\mathbf{1} = \mathbf{1}$, $\forall A \in S$, we obtain

$$\begin{aligned} P_{\sigma(t+1)} - P_{\sigma(t)} &= (A_{\sigma(t+1)} - I)P_{\sigma(t)} \\ &= (A_{\sigma(t+1)} - I)QP_{\sigma(t)}. \end{aligned} \tag{4.3}$$

We observe that for any vector $x \in \mathbb{R}^n$,

$$\|x\|_{\mathcal{P}} \leq \|x\|_\infty$$

and

$$\|x\|_{\mathcal{P}} = \|Qx\|_{\mathcal{P}} \geq \frac{1}{2}\|Qx\|_\infty.$$

So that by using the definition of the induced norm $\|.\|_\infty$ and seminorm $\|.\|_\mathcal{P}$, we obtain for any matrix $B \in \mathbb{R}^{n\times n}$:

$$\|B\|_\mathcal{P} = \|QB\|_\mathcal{P} \geq \frac{1}{2}\|QB\|_\infty. \tag{4.4}$$

Using (4.3), (4.4) and Proposition 4.2 we obtain, for some appropriate constants C_1 and C_2,

$$\begin{aligned}
\|P_{\sigma(t+1)} - P_{\sigma(t)}\|_\infty &\leq \|A_{\sigma(t)} - I\|_\infty \|QP_{\sigma(t)}\|_\infty \\
&\leq (1 + \max_{A_i \in S} \|A_i\|_\infty) 2\|P_{\sigma(t)}\|_\mathcal{P} \\
&\leq C_1 \|P_{\sigma(t)}\|_\mathcal{P} \\
&\leq C_2 r^t
\end{aligned}$$

Therefore, for any $q > s$

$$\begin{aligned}
\|P_{\sigma(q)} - P_{\sigma(s)}\|_\infty &\leq \sum_{t=s}^{q-1} \|P_{\sigma(t+1)} - P_{\sigma(t)}\|_\infty \\
&\leq C_2 \sum_{t=s}^{q-1} r^t \\
&\leq C_2 r^s \sum_{t=s}^{\infty} r^{t-s} \\
&\leq C_3 r^s
\end{aligned}$$

where C_3 does not depend on s. Therefore the trajectory is a Cauchy sequence and converges. By Proposition 4.2 it can only converge to span$\{\mathbf{1}\}$. ■

4.2 Algorithmic Decision of Consensus

We now develop a singly exponential algorithm to decide Problem 4.1. For that we define the notion of *graph of double-faces* whose nodes are the pairs of opposite proper faces of $\mathcal{P}$, that we call *double-faces*. We then represent the products of matrices from a set as walks on the graph and we show that Problem 4.1 has a positive answer if and only if the graph has no cycle.

First, we present a lemma that plays a key role in the proof of the finiteness result. It shows how all the points in a face generate similar trajectories. It is similar to a claim in the proof of Theorem 4.1 in [LW95]; we state it here as an independent lemma because our hypotheses are slightly different.

Lemma 4.1 (Lagarias and Wang [LW95]). *Let S be a finite set of matrices having a common invariant polyhedron $\mathcal{Q}$. Then, for any $A \in S$ and any open face F of $\mathcal{Q}$, there exists exactly one open face G (possibly int$(\mathcal{Q})$) such that*

$$AF \subseteq G.$$

Proof. Since $\mathcal{Q}$ is equal to the disjoint union of its open faces (see Section 3.1) and $AF \subseteq \mathcal{Q}$, then AF intersects with at least one open face of $\mathcal{Q}$.

We now prove by contradiction that the image AF intersects at most one open face. Suppose that there were points $x_1, x_2 \in F$ such that Ax_1 and Ax_2 were in different open faces. These open faces differ in at least one facet constraint (Definition 3.5), with one having $b^\top x = c$ and the other $b^\top x < c$ (Lemma 3.1 and Corollary 3.1). Without loss of generality, suppose that $b^\top Ax_1 = c$ and $b^\top Ax_2 < c$. Since F is relatively open and convex, there exists $\varepsilon > 0$ such that

$$(1-\lambda)x_1 + \lambda x_2 \in F \text{ for } -\varepsilon \leq \lambda \leq 1+\varepsilon.$$

In particular $y = (1+\varepsilon)x_1 + -\varepsilon x_2 \in F$ and $b^\top Ay > c$, which implies $Ay \notin \mathcal{Q}$, contradiction with $A(\mathcal{Q}) \subseteq \mathcal{Q}$.

■

Since $\mathcal{P}$ is symmetric around the origin, its proper faces form pairs that are opposite to each other. We define the notion of double-face to represent these pairs of opposite faces. These double-faces will be the nodes of our graph of double-faces and dealing with double-faces instead of pairs allows to have less nodes in the graph and to reduce the complexity of our algorithm.

Definition 4.3 (Double-face). *A* double-face *is a set equal to $F \cup -F$ for some proper face F. A double-face is called open if the face F is open, and closed otherwise.*

Lemma 4.1 naturally holds for open double-faces as well: if $A\mathcal{Q} \subseteq \mathcal{Q}$, then for any open double-face F of $\mathcal{Q}$, there is an open double-face G such that $AF \subseteq G$.

4.2.1 Graph Representation of Products

We present a method to represent products as walks on a graph. The graph of double-faces is constructed from a set of matrices S and a polyhedron $\mathcal{Q}$ that is symmetric around the origin and that is invariant with respect to the matrices of S. The nodes depend only on the polyhedron $\mathcal{Q}$ while the edges depend on the matrices. The nodes represent double-faces of a polyhedron $\mathcal{Q}$ and the edges represent the possibility to jump from one double-face to another using a matrix of S. We will show that this graph captures enough information to decide Problem 4.1. We define the graph of double-faces for an arbitrary centrally symmetric polyhedron and then we will use the graph of double-faces of $\mathcal{P}$ to decide Problem 4.1.

Definition 4.4 (Graph of double-faces)**.** *Given a finite set S of matrices and $\mathcal{Q}$ an invariant polyhedron that is symmetric around the origin ($\mathcal{Q} = -\mathcal{Q}$), we call the graph of double-faces $\mathcal{G}$ the graph having*

- *one vertex for each double-face of $\mathcal{Q}$, one node representing int$(\mathcal{Q})$ that we call "node 1" by convention.*
- *one edge from node i to node j if they correspond to double-faces F_i and F_j and there is $A \in S$ such that $AF_i \subseteq F_j$. In particular there is one edge from node i to node 1 if node i corresponds to a double-face F_i and there is $A \in S$ such that $AF_i \subseteq$ int$(\mathcal{Q})$ and one edge going from node 1 to itself.*

Example 4.1. *We construct the graph of double-faces of the set*

$$\left\{ A = \begin{pmatrix} 0 & \frac{1}{2} \\ -1 & -\frac{1}{2} \end{pmatrix}, B = \begin{pmatrix} -\frac{1}{4} & \frac{3}{4} \\ -\frac{3}{4} & \frac{1}{4} \end{pmatrix} \right\}$$

and the polyhedron

$$\mathcal{Q} = \{x \mid \|x\|_1 \leq 1\}.$$

The polyhedron is invariant for matrices A and B as depicted on Figure 4.1.

To make the construction easier to follow, we start with one node for each face (instead of one for each double-face). The graph has therefore nine nodes: one for int$(\mathcal{Q})$, one for each vertex (the corners) and one for each facet (the sides of the square). The image by A of vertex $F_1 = \begin{pmatrix} 1 & 0 \end{pmatrix}^\top$ is $\begin{pmatrix} \frac{1}{2} & -\frac{1}{2} \end{pmatrix}^\top$, which is in the face

$$F_2 = \{x \mid x_1 - x_2 = 1,\ x_1 + x_2 < 1,\ -x_1 - x_2 < 1\}.$$

There is therefore an edge from the node representing F_1 to the node representing F_2, as depicted on the left of Figure 4.2. By doing the same for each face, we find the entire graph for matrix A and polyhedron $\mathcal{Q}$.

We then add the edges corresponding to matrix B (Figure 4.3, left). The last step is to merge the nodes representing opposite faces and removing the edges that appear twice. We obtain the final graph of double-faces (Figure 4.3, right).

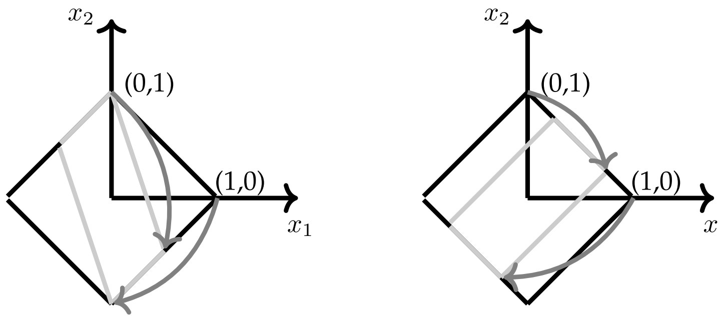


Figure 4.1: Left: The polyhedron $\mathcal{Q}$ (black), $A\mathcal{Q}$ (light grey). The dark grey arrows represent $A(1 \quad 0)^\top = (0 \quad -1)^\top$ and $A(0 \quad 1)^\top = \left(\frac{1}{2} \quad -\frac{1}{2}\right)^\top$. Right: the same for matrix B.

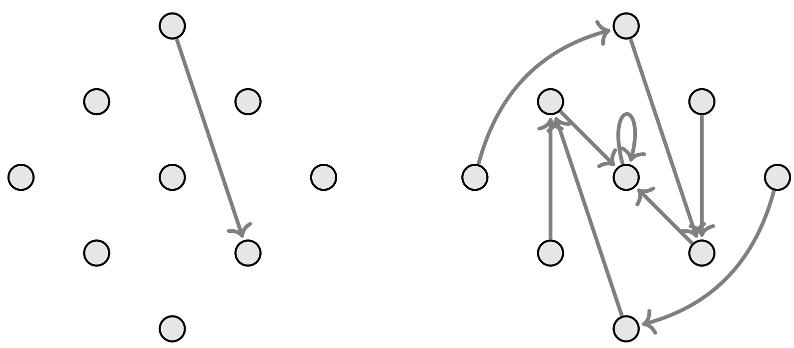


Figure 4.2: Left: the nodes of the graph with the edge from F_1 to F_2. The node in the middle represents int($\mathcal{Q}$), the nodes in the corners represent the vertices of the polyhedron and the other nodes represent the facets of the polyhedron. The edge is the one from F_1 to F_2. Right: the graph with all edges for matrix A (one edge from each node).

Theorem 4.1. *Let S be a finite set of matrices satisfying Assumption 4.1. S is contractive if and only if the self-loop of node 1 is the only cycle in the graph of double-faces.*

Proof. Only if If there is a cycle in the graph of double-faces other than the self-loop, then there is an open double-face F and a product P such that $PF \subseteq F$ and $P^tF \subseteq F$ for any t. Therefore, the S is not contractive. Hence S is a contractive set only if the self-loop of node 1 is the only cycle in the graph of double-faces.

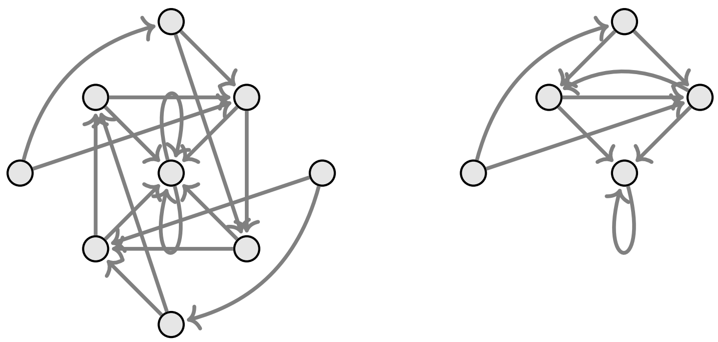


Figure 4.3: Left : the graph for matrices A and B. Right : the graph of double-faces as defined in Definition 4.4.

If If there is no cycle other than the self-loop, then all sufficiently long walks end in node 1, because each node has an outdegree of at least one. This means that for any sufficiently long product P, the image of each double-face is in int($\mathcal{P}$) and hence any sufficiently long product contracts $\mathcal{P}$. ■

4.3 Computational aspects

Now that we have necessary and sufficient conditions (Theorem 4.1) for Problem 4.1, we estimate the algorithmic complexity of evaluating these conditions.

To construct the graph of double-faces, we need two basic operations: to compute in which face a point is, and to find a point in a given face. In Lemma 3.4, we have described the (open) faces of $\mathcal{P}$. From this description, it is computationally easy (in $\mathcal{O}(n)$) to determine in which open face a point $x \in \mathcal{P}$ is. Finding a point in a face can also be done with a complexity of $\mathcal{O}(n)$.

We are now able to prove our complexity result.

Theorem 4.2. *Problem 4.1 can be decided in $\mathcal{O}(3^n mn^2)$ operations.*

Proof. Construction of the graph of double-faces: The graph has $N = \frac{1}{2}(3^n - 2^{n+1} + 1)$ nodes. Each node has at most m outgoing edges (at most Nm in total), corresponding to the m matrices in S.

To compute the edge starting from node i (representing face F) and corresponding to transition matrix A, we need to find the face G such that

$$AF \subseteq G.$$

By Lemma 4.1, we know that G is the face containing Ax where x is any point in F. Finding G can be done in $\mathcal{O}(n^2)$ operations : take a point x in F (in $\mathcal{O}(n)$), compute Ax (in $\mathcal{O}(n^2)$), and find the face in which Ax is (in $\mathcal{O}(n)$). Therefore the complexity of constructing the graph of double-faces is $\mathcal{O}(3^n mn^2)$.

Decision problems on the graph: Once the graph is constructed, Problem 4.1, which is equivalent to the existence of cycles in the graph (see Theorem 4.1), can be decided using a topological sorting algorithm which has a complexity of $\mathcal{O}(|E| + |V|) = \mathcal{O}(3^n m)$ with $|E|$ and $|V|$ being respectively the numbers of edges and vertices [Kah62]. The complexity of checking the existence of cycles is thus lower than the complexity of constructing the graph. ■

Since [BO14, CHJ14], the problem of deciding whether a set is a consensus set was known to be decidable in $\mathcal{O}(m^{3^n} n^\omega)$ operations, where $\mathcal{O}(n^\omega)$ is the complexity of multiplying two $n \times n$ matrices. We have now an algorithm with a complexity that is singly exponential in n. We note however that the space complexity of our algorithm is very large. Indeed we need to store the graph of double-faces, which has a number of nodes that is exponential in n. More precisely, the number of nodes is $\Theta(3^n)$ and the number of edges is $\Theta(3^n m)$ and so is the total space complexity of our method.

4.3.1 Discussion

We have obtained a singly exponential algorithm to solve the problem of deciding whether a set is contractive. This is an improvement over the decision procedure of [BO14]. Indeed, the authors of [BO14] show that their procedure has a doubly exponential time complexity and a short analysis shows that it has a singly exponential space complexity. Our algorithm implies that the problem belongs to EXPTIME, while it was known that it belongs to EXPSPACE. It is also known that the problem is NP-hard.

Recall the classical complexity hierarchy:

$$\text{P} \subseteq \text{NP} \subseteq \text{PSPACE} \subseteq \text{EXPTIME} \subseteq \text{EXPSPACE} \subseteq \text{2-EXPTIME},$$ [2]

[2] with P the class of problems that can be solved in polynomial time, NP the class of problems for which instances with positive answer have a polynomial time verifiable certificate, PSPACE the class of problems that can be solved with polynomial space, EXPTIME the class of problems that can be solved in singly exponential time, EXPSPACE the class of problems that can be solved with singly exponential space and 2-EXPTIME the class of problems that can be solved in doubly exponential time.

Whether the problem belongs to NP, or to PSPACE are open questions, and so are the questions of whether the problem is PSPACE-hard and EXPTIME-hard. These considerations might seem theoretical but they have direct algorithmic implications. For example, if the problem is EXPTIME-hard, then there is no algorithm with a polynomial space complexity, unless PSPACE = EXPTIME (it is usually believed that these classes are different).

4.4 Sets of two stochastic undirected matrices

In this last section, we study the complexity of Problem 4.1 in the case of *undirected* matrices. A nonnegative matrix is called *undirected* [BO14] if

$$a_{ij} > 0 \Leftrightarrow a_{ji} > 0.$$

Undirected matrices are also sometimes called *type-symmetric*. In the context of consensus systems, they correspond reciprocal communication between agents: an agent i influences the value of an agent j if and only if agent j influences the value of agent i. It is known that reciprocity plays an important role in the convergence of consensus systems [HT13]. It has been proven that Problem 4.1 is NP-hard for sets of three undirected matrices and for sets of two stochastic matrices in general [BO14]. The authors have left open the case of sets of two undirected matrices. We prove that, in this case, it can be solved in polynomial time.

It is worth noticing that the product of two stochastic matrices is a stochastic matrix, and that stochastic matrices satisfy the relations:

$$\max_i (Ax)_i \leq \max_i x_i$$

$$\min_i (Ax)_i \geq \min_i x_i.$$

Sets of stochastic matrices satisfy therefore Assumption 4.1.

The next lemma presents a simple yet crucial observation about undirected stochastic matrices.

Lemma 4.2. *Let A be an undirected stochastic matrix. Then A^2 has a positive diagonal.*

The following lemma shows the effect of a transition matrix with a positive diagonal.

Lemma 4.3. *Let A, B and C be stochastic matrices and let B have a positive diagonal. If ABC is not SIA, then AC is not SIA either.*

Proof. The matrices A, B and C are nonnegative (because they are stochastic). Since B has a positive diagonal, an element $(AC)_{ij}$ of AC can be positive only if the element $(ABC)_{ij}$ at the same position in ABC is positive as well. Hence, if ABC is not SIA, then AC is not SIA either. ■

Theorem 4.3. *Let $S = \{A_1, A_2\}$ be a set of two stochastic undirected matrices. The set S is contractive if and only if A_1, A_2 and A_1A_2 are SIA.*

Proof. If A_1 is not SIA, then the sequence $(A_1^t)_t$ does not converge to a rank-one matrix and the set S is not contractive. By the same argument, we conclude that A_2 and A_1A_2 are SIA if S is contractive.

Now suppose that S is not contractive. Essentially, we will prove that there is an infinite product that is not SIA and that if the product contains an infinite number of A_1 and A_2, we can remove subproducts A_1A_1 and A_2A_2 (with the use of Lemma 4.3) in order to obtain the infinite product $\ldots A_2A_1A_2A_1$ (or $\ldots A_1A_2A_1A_2$) and prove that it is not SIA.

By Theorem 4.1, there is a cycle other than the self-loop at node 1 in the graph of double-faces of S. We take a node on the cycle, representing an open double-face F. The cycle starting from F then represents a product $P_{\sigma(k)}$ such that

$$P_{\sigma(k)}F \subseteq F.$$

Therefore there is no t such that P^t contracts $\mathcal{P}$. By Proposition 4.1 (that establishes the equivalence between the contraction of $\mathcal{P}$ and the matrix being scrambling) there is thus no t such that P^t is scrambling. Since any SIA matrix has a power that is scrambling, there is also no t such that P^t is SIA. Taking the square of the product provides a product of even length $2k$ having the following property: there is no t such that $(P^2_{\sigma(k)})^t$ is SIA. Let us now take the shortest sequence σ^* of even length k^* such that

$$\forall\, t,\ (P_{\sigma^*(k^*)})^t \text{ is not SIA.} \tag{4.5}$$

If this sequence σ^* has a length of 2, then $P_{\sigma^*(k^*)}$ is equal to A_1A_2, to A_2A_1, to A_1^2 or to A_2^2 and the proof is done. Otherwise, the sequence σ^* has a length of at least 4. Suppose that σ^* contains the subsequence i, i with $i = 1$ or $i = 2$:

$$\forall\, t,\ (P_{\sigma^*(k^*)})^t \text{ is not SIA.}$$

which is

$$\forall\, t,\ (A_{\sigma^*(k^*)} \dots A_{\sigma^*(l)} A_i A_i P_{\sigma^*(l-3)})^t \text{ is not SIA.}$$

By Lemma 4.3 we can state that

$$\forall\, t,\ (A_{\sigma^*(k^*)} \dots A_{\sigma^*(l)} P_{\sigma^*(l-3)})^t \text{ is not SIA}$$

and σ^* was not the *shortest* sequence of even length that satisfies Property (4.5). We have a contradiction and we can conclude that σ^* does not contain the subsequences $1, 1$ nor $2, 2$. But then σ^* is equal to $1, 2$ or $2, 1$ (would σ^* be longer, for example $\sigma^* = 1, 2, 1, 2$ then the sequence $1, 2$ would be shorter, would have an even length and would also satisfy Property (4.5)).

■

By Corollary 2.2, checking whether a stochastic matrix is SIA can be done in polynomial time and Problem 4.1 can therefore be decided in polynomial time for sets of two stochastic undirected matrices.

Corollary 4.1. *For sets of two stochastic undirected matrices, Problem 4.1 can be decided with a complexity of $\mathcal{O}(\omega(n)\log(n))$, where $\mathcal{O}(\omega(n))$ is the complexity of the multiplication of two $n \times n$ matrices.*

4.5 Conclusion

The goal of this chapter was to investigate the complexity of deciding whether a set is contractive (Problem 4.1). We have obtained a geometric characterization allowing for a singly exponential algorithm to solve this problem. As discussed in Subsection 4.3.1, this does not entirely answer the question of the complexity of the problem. Indeed, we now know that the problem belongs to the complexity class EXPTIME and we knew already [BO14] that it is NP-hard. But we do not know whether the problem is in PSPACE, nor do we know whether it is in NP. Similarly, PSPACE-hardness and EXPTIME-hardness are open questions.

We have also improved the state-of-the-art complexity in the particular case of sets of two undirected matrices. We proved the existence of a polynomial-time algorithm for this case.

Consensus systems with stochastic matrices have an invariant polyhedron, which makes them naturally suited for the analysis that we have developed.

We would like to mention however that this reasoning can apply for any discrete time linear switched system that admits a common invariant polyhedron. The singly exponential complexity would still hold in those cases. The exact complexity may be different. Indeed, one of the building blocks of the method is to determine in which face a point is. We can do it here in $\mathcal{O}(n)$ operations because of the representation of the polyhedron given by Lemma 3.4. This compact representation is possible for this particular polyhedron but not necessarily for all of them. In Chapter 6, we will apply these techniques to another polyhedron (namely the nonnegative orthant). We will show that constructing its graph of faces provides a new method to decide whether a set of nonnegative matrices is primitive (in the sense that all sufficiently long products are positive).

Chapter 5

Tight Bound for Consensus

For a set of matrices that share a common nonincreasing polyhedral norm, all infinite products converge to zero if and only if all infinite *periodic* products with period smaller than a certain value converge to zero. Moreover, bounds on that value are available [LW95].

In this chapter, we provide a stronger bound that holds for polyhedral norms and also for all *seminorms*. In the latter case, the matrix products do not necessarily converge to zero, but to a common invariant subspace. We prove that our bound is tight for all polyhedral seminorms (and thus tight for all norms). We study the particular case of matrices that leave $\mathcal{P}$ invariant and we obtain that there is an infinite product that does not contract $\mathcal{P}$ if and only if there is an infinite periodic product with period at most

$$\binom{n}{\lfloor n/3 \rfloor}(2^{n-\lfloor n/3 \rfloor - 1} - 1) \tag{5.1}$$

that does not contract $\mathcal{P}$.

This chapter presents the main results of [CHJ16] and is organized as follows.

- In Section 5.1, we introduce and formalize the problem.
- In Section 5.2, we develop a method to compute a bound for general centrally symmetric polyhedra (which can be seen as unit balls of seminorms).
- In Section 5.3, we apply the results of the previous section to $\mathcal{P}$ to obtain the bound (5.2).

5.1 Introduction

We consider sets of matrices that leave invariant a polyhedron $\mathcal{Q}$ that is symmetric around the origin: $\mathcal{Q} = -\mathcal{Q}$. As discussed in Subsection 3.1.5, this is equivalent to assuming that the seminorm associated with the polyhedron $\|.\|_{\mathcal{Q}}$ is nonincreasing for the matrices of the set. As in Subsection 3.1.5, we will simply call *centrally symmetric* the polyhedra that are symmetric around the origin.

For any centrally symmetric polyhedron, we would like to study a quantity that we call *finiteness bound*.

Definition 5.1 (Finiteness bound)**.** *Let $\mathcal{Q}$ be a centrally symmetric polyhedron and let $S = \{A_1, \ldots, A_m\}$ be a set of matrices that leave $\mathcal{Q}$ invariant and such that not all infinite products of matrices of S contract $\mathcal{Q}$. We define ℓ as the length of the shortest product P such that no power of P contract $\mathcal{Q}$.*

For any centrally symmetric polyhedron $\mathcal{Q}$, we call a finiteness bound *p any upper bound on ℓ that holds for all sets that leave $\mathcal{Q}$ invariant.*

For a centrally symmetric polyhedron $\mathcal{Q}$, a finiteness bound p and a set $S = \{A_1, \ldots, A_m\}$ of matrices that leave $\mathcal{Q}$ invariant, the existence of an infinite product that does not contract $\mathcal{Q}$ implies the existence of an infinite periodic product with period at most p that does not contract $\mathcal{Q}$.

When there is no left-infinite product of matrices from a finite set S that does not contract $\mathcal{Q}$, for any infinite product $\ldots A_{\sigma(2)} A_{\sigma(1)}$,

$$\forall x \in \mathcal{Q}, \lim_{t \to \infty} \|A_{\sigma(t)} \ldots A_{\sigma(2)} A_{\sigma(1)} x\|_{\mathcal{Q}} = 0,$$

where $\mathcal{Q}$ is the seminorm associated to $\mathcal{Q}$.[1] Therefore, in the particular case of a *bounded* polyhedron $\mathcal{Q}$, the nonexistence of an infinite product that does not contract $\mathcal{Q}$ is equivalent to the convergence to zero of all infinite products. Hence, if p is a finiteness bound for a bounded polyhedron $\mathcal{Q}$ and S is a finite set of matrices that leave $\mathcal{Q}$ invariant, all infinite products converge to zero if and only if all infinite products with period at most p converge to zero.

The number ℓ in Definition 5.1 is equal to the length of the shortest cycle in the graph of double-faces (Definition 4.4) corresponding to S and $\mathcal{Q}$. A finiteness bound is thus a bound on these lengths, that holds for all sets that leave $\mathcal{Q}$ invariant.

[1] This is a consequence of the compactness of S, of the finite number of faces of $\mathcal{Q}$ and the linearity of the matrix product and the formal proof of this statement is very similar to the proof of Proposition 4.2.

The goal of this chapter is to find the best finiteness bounds.

Question 5.1. *Given a centrally symmetric polyhedron $\mathcal{Q}$, what is p^* the smallest possible finiteness bound?*

Question 5.1 has found a partial answer in [LW95, Theorem 4.1]. The authors have shown that when the matrices of a finite matrix set have a common nonincreasing polyhedral norm, there is an infinite product that does not converge to zero if and only if there is an infinite periodic product with period at most $k = \frac{f}{2}$ that does not contract $\mathcal{Q}$, where f is the number of faces of the unit ball of the norm. We will extend this result to seminorms and we will see that smaller finiteness bounds exist.

This question has also been studied in the case of stochastic matrices. It has been established [Paz71] that, for any set of stochastic matrices S, if there is ℓ such that all products of length at least ℓ of matrices of S are scrambling, then all products of length at least $B = \frac{1}{2}(3^n - 2^{n+1} + 1)$ are scrambling. Subsequently, Blondel and Olshevsky [BO14] showed that a set is contractive if and only if all products of length $B = \frac{1}{2}(3^n - 2^{n+1} + 1)$ are scrambling. We will show that a set S is contractive if and only if all products of length at most

$$p^* = \binom{n}{\lfloor n/3 \rfloor}(2^{n-\lfloor n/3 \rfloor - 1} - 1) \approx \frac{3}{2\sqrt{\pi n}}B \tag{5.2}$$

are SIA.

The bound of Paz and Blondel and Olshevsky, $B = \frac{1}{2}(3^n - 2^{n+1} + 1)$, is in fact equal to half the number of proper faces of $\mathcal{P}$. The number of faces of $\mathcal{P}$ has been computed in Corollary 3.2 and is equal to $3^n - 2^{n+1} + 2$, one of them being the non-proper face $\mathcal{P}$. In fact, we have shown in [CHJ14] that the result of Blondel and Olshevsky could be obtained by using the invariant polyhedron $\mathcal{P}$. We then realized that this bound was not tight and in [CHJ16] we obtained the tight bound of Equation (5.2). One motivation to find the best finiteness bound was to obtain a better algorithm to decide if a set is contractive. Indeed, a polynomial bound would have yielded a polynomial space algorithm, which would thus have been better that the algorithm developed in Chapter 4. This was unfortunately not the case, as the bound that we obtained is exponential in n.

5.2 The General Case

In this section, we answer Question 5.1 in general. We start by recalling some definitions of *poset*, *lattice* and *antichain*. We follow the terminology of [Zie95].

Definition 5.2 (Poset and lattice). *A* partially ordered set *or* poset *is a set P with a binary relation $\preceq$ that is transitive, antisymmetric and reflexive. We also note $x \prec y$ for the relation $x \preceq y$ and $x \neq y$.*

A poset is called a lattice *if any pair of elements has a unique infimum and a unique supremum.*

Definition 5.3 (Graded poset). *A poset $(P, \preceq)$ is called* graded *if it can be equipped with a rank function*

$$r : P \mapsto \mathbb{N}$$

that is compatible with the ordering:

$$x \preceq y \Rightarrow r(x) \leq r(y)$$

and such that any two comparable elements $y \prec x$ either have ranks that differ by 1 or there is a element z between them $y \prec z \prec x$:

$$\forall\ x, y \in P,\ \textit{such that}\ y \prec x,\ \ r(x) = r(y) + 1\ \textit{or}\ \exists z,\ y \prec z \prec x.$$

The set of all elements of a given rank is called a rank level.

Definition 5.4 (lattice of double-faces). *Given a centrally symmetric polyhedron $\mathcal{Q}$ (i.e., a polyhedron $\mathcal{Q} = -\mathcal{Q}$), we call* lattice of double-faces *the poset $(P, \subseteq)$ where $\subseteq$ is the inclusion relation and P is a set whose members are*

- *double-faces of $\mathcal{Q}$ ($r =$ dimension of the face)*
- *$\mathcal{Q}$ ($r = n$)*
- *$\emptyset$ ($r = d_{\min} - 1$, where $d_{\min}$ is the lowest dimension of faces of $\mathcal{Q}$).*

It can be verified that this poset is a lattice (Definition 5.2) and that it is graded; a rank function is given between brackets.

The notion of *antichain* will play a key role in this chapter.

Definition 5.5 (Antichain). *Let $(P, \preceq)$ be a poset. An* antichain *is a subset $S \subseteq P$ whose elements are not comparable:*

$$\forall x, y \in S,\ \ x \npreceq y.$$

For instance, a set of double-faces that are not included in one another form an antichain in the lattice of double-faces.

Example 5.1. *The Polyhedron $\mathcal{P}$, as well as some of its double-faces, is represented in Figure 5.1. Its lattice of double-faces and its largest antichain are represented in Figure 5.2.*

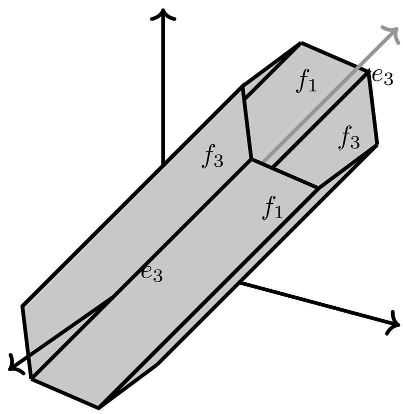


Figure 5.1: The polyhedron $\mathcal{P}$ for $n = 3$. The gray arrow indicates the direction $a\mathbf{1}$. The polyhedron has 6 double-faces: three that correspond to the 6 facets and 3 that correspond to the 6 edges. The sets f_1, f_3 and e_3 are double-faces.

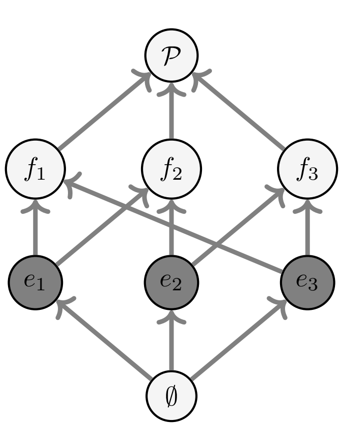


Figure 5.2: The lattice of double-faces of the polyhedron $\mathcal{P}$ for $n = 3$. The elements f_1, f_2 and f_3 represent the three pairs of opposite facets while e_1, e_2 and e_3 represent the three pairs of opposite edges. In dark gray, a largest antichain in this lattice.

Definition 5.6 (Width of a Poset)**.** *We call the width $W(P)$ of a poset P the number of elements of the largest antichain of P. We also write $W(\mathcal{Q})$ for the width of the lattice of double-faces of a given centrally symmetric polyhedron $\mathcal{Q}$.*

The next theorem provides a finiteness bound. It is an improvement of [LW95, Theorem 4.1] as we extend it to seminorms and we provide a stronger bound.

Theorem 5.1. *Let S be a finite set of matrices and let $\mathcal{Q}$ be a polyhedron that is symmetric around the origin and invariant for S. If there is a left-infinite product of matrices from S that does not contract $\mathcal{Q}$, there is one that is periodic with a period p not larger than*

$$p^* = W(\mathcal{Q})$$

that does not contract $\mathcal{Q}$.

Proof. We first observe that p is finite. Suppose there exists an infinite noncontracting product $\dots A_{\sigma(2)}A_{\sigma(1)}$ and therefore a point x_0 such that

$$\forall i,\ A_{\sigma(i)} \dots A_{\sigma(1)}x_0 \notin \text{int}(\mathcal{Q}).$$

Since the number of faces is finite, there is an open double-face O and indices $i < j$ such that

$$A_{\sigma(i)} \dots A_{\sigma(1)}x_0 \in O \text{ and } A_{\sigma(j)} \dots A_{\sigma(1)}x_0 \in O.$$

By Lemma 4.1, we have

$$A_{\sigma(j)} \dots A_{\sigma(i+1)}O \subseteq O.$$

Therefore, the infinite power of $A_{\sigma(j)} \dots A_{\sigma(i+1)}$ is an infinite *periodic* noncontracting product, proving that the theorem is true for some finite period $p = j - i$ smaller than the number of double-faces.

We now prove the full theorem. Let P be such that $\dots PPP$ is an infinite noncontracting product with the smallest period p and

$$P = A_{\sigma(p)} \dots A_{\sigma(1)}.$$

Let O_1 be an open double-face such that

$$\forall t \geq 0,\ (P)^t O_1 \not\subseteq \text{int}(\mathcal{Q})$$

(such a face exists due to Lemma 4.1 and the fact that $\dots PPP$ does not contract $\mathcal{Q}$), let O_2 be the double-face containing $A_{\sigma(1)}O_1$ (by Lemma 4.1, there is exactly one such double-face), O_3 containing $A_{\sigma(2)}A_{\sigma(1)}O_1$ up to O_p containing $A_{\sigma(p-1)}\dots A_{\sigma(1)}O_1$. Let also $F_1 = \mathrm{cl}(O_1), \dots, F_p = \mathrm{cl}(O_p)$.

We now prove that $\{F_1, \dots, F_p\}$ in an antichain in the lattice of double-faces. Suppose, to obtain a contradiction, that for some i, j with $i > j$, $F_i \subseteq F_j$. Then,

$$A_{\sigma(i-1)}\dots A_{\sigma(j)}F_j \subseteq F_i \subseteq F_j,$$

and thus

$$\forall t \geq 0,\ (A_{\sigma(i-1)}\dots A_{\sigma(j)})^t F_j \subseteq F_j.$$

This contradicts the assumption that $\dots PPP$ is the infinite periodic noncontracting product *with the smallest period*. Similarly, if for some i, j with $i < j$, $F_i \subseteq F_j$, then

$$\forall t \geq 0,\ (A_{\sigma(i-1)}\dots A_{\sigma(1)}A_{\sigma(p)}\dots A_{\sigma(j)})^t F_j \subseteq F_j,$$

and again we have a contradiction. Hence $\{F_1, \dots, F_p\}$ is an antichain in the lattice of double-faces and P has a length smaller than or equal to $W(\mathcal{Q})$ ■

The bound p^* of Theorem 5.1 cannot be decreased in general: it is tight for any polyhedron in any dimension n, as we prove in the next theorem.

Theorem 5.2. *Let $\mathcal{Q}$ be a polyhedron in dimension n that is centrally symmetric around the origin and let $p^* = W(\mathcal{Q})$ as in Theorem 5.1. There is a set S of $n \times n$ matrices that leave $\mathcal{Q}$ invariant and such that*

- *all left-infinite periodic products with periods smaller than p^* contract $\mathcal{Q}$,*
- *there is a left-infinite product that does not contract $\mathcal{Q}$.*

Proof. We construct a set of matrices such that the infinite noncontracting product that has the smallest period has a period equal to $p^* = W(\mathcal{Q})$. Let $X = \{F_1, \dots, F_{p^*}\}$ be the largest antichain in the lattice of double-faces and let $O_1, \dots, O_{p^*}$ be the corresponding open double-faces.

By definition, each double-face F_i is the union of two opposite proper faces $G_i, -G_i$ and the proper face G_i is the intersection of $\mathcal{Q}$ with a hyperplane

$$G_i = \mathcal{Q} \cap \{x : b_i^\top x = c_i\}$$

such that $\mathcal{Q}$ is in one halfspace defined by the hyperplane:

$$\mathcal{Q} \subseteq \{x : b_i^\top x \leq c_i\}.$$

We also have $c_i \neq 0$. Indeed, if $c_i = 0$, then $\mathcal{Q} \subseteq \{x : b_i^\top x \leq 0\}$ and because $\mathcal{Q} = -\mathcal{Q}$, and $\mathcal{Q} \subseteq \{x : -b_i^\top x \leq 0\}$ and this implies $G_i = \mathcal{Q} \cap \{x : b_i^\top x = 0\} = \mathcal{G}$ and G_i is not a proper face. Therefore, $c_i \neq 0$ and we can scale b_i and c_i to have $\forall i,\ c_i = 1$. Finally, $F_i = G_i \cup -G_i = \mathcal{Q} \cap \{x : b_i^\top x = \pm 1\}$.

By taking any v_i in the open double-face $O_{(i \bmod p^*)+1}$ and defining

$$A_i = v_i b_i^\top \text{ and } S = \{A_1, \ldots, A_{p^*}\},$$

we have

$$\begin{aligned}
\forall i,\ A_i F_i &= A_i(\mathcal{Q} \cap \{x : b_i^\top x = \pm 1\}) \\
&\subseteq A_i\{x : b_i^\top x = \pm 1\} \\
&= \{A_i x : b_i^\top x = \pm 1\} \\
&= \{v_i b_i^\top x : b_i^\top x = \pm 1\} \\
&= \{\pm v_i\} \\
&\subseteq O_{(i \bmod p^*)+1}.
\end{aligned} \tag{5.3}$$

We have as well

$$\begin{aligned}
\forall i,\ A_i(\mathcal{Q} \backslash F_i) &= A_i(\mathcal{Q} \cap \{x : -1 < b_i^\top x < 1\}) \\
&\subseteq \{v_i b_i^\top x : -1 < b_i^\top x < 1\} \\
&= \{\lambda v_i : -1 < \lambda < 1\} \\
&\subseteq \{\lambda y : -1 < \lambda < 1,\ y \in \mathcal{Q}\} \\
&= \text{int}(\mathcal{Q}).
\end{aligned} \tag{5.4}$$

By (5.3) and (5.4), for any $j \neq (i \bmod p^*) + 1$,

$$\begin{aligned}
A_j A_i \mathcal{Q} &\subseteq A_j \left(\text{int}(\mathcal{Q}) \cup O_{(i \bmod p^*)+1}\right) \\
&= A_j \text{int}(\mathcal{Q}) \cup A_j O_{(i \bmod p^*)+1} \subseteq \text{int}(\mathcal{Q}).
\end{aligned}$$

Therefore,

$$\ldots A_{(h+2 \bmod p^*)+1} A_{(h+1 \bmod p^*)+1} A_{(h \bmod p^*)+1} A_h$$

is the only infinite noncontracting product starting with A_h. For any h, this product has a period of p^* (because the matrices $A_1, \ldots, A_{p^*}$ are all different). We conclude that all infinite periodic products with periods smaller than $m = p^*$ contract $\mathcal{Q}$ and the theorem is proven. ■

The number of matrices of the set S that we construct in Theorem 5.2 is equal to half the number of facets of the corresponding polyhedron. Therefore this does not guarantee the tightness in the particular case where the number of matrices is limited. For example, the existence of a smaller bound p in the case of sets of two matrices remains open.

Giving an explicit value to the size of the largest antichain may prove difficult in some cases. However, since a set of double-faces of same dimension always constitutes an antichain, the largest antichain has at least $\max_i f_i$ elements, and we have the following lower bound

$$p^* = W(\mathcal{Q}) \geq \max_i f_i, \tag{5.5}$$

where f_i is the number of faces of dimension i. The inequality (5.5) might not be very useful, because $p*$ is in essence an upper bound. However, if the equality holds, the exact value of p^* can be known. This is the case when the lattice of double-faces of Q has the *Sperner property*:

Definition 5.7 (Sperner Property [Eng97])**.** *A graded poset is said to have the Sperner property if its largest antichain is equal to its largest rank level (Definition 5.3).*

In the next section, we will study the structure of the lattice of double-faces of $\mathcal{P}$. We will see that the lattice has the Sperner property and we will be able to compute an explicit value for our bound p^* by computing the number of elements of its largest rank level.

5.3 The case of the Polyhedron $\mathcal{P}$

We investigate the case of sets that leave polyhedron $\mathcal{P}$ invariant. In Theorem 5.3, we prove that the lattice of double-faces of $\mathcal{P}$ has the Sperner property. In Theorem 5.4 we compute the value of p^* for $\mathcal{P}$. Finally, in Theorem 5.5, we prove that the bound p^* is tight for sets of stochastic matrices. We end this section by discussing the relation between these results, the results of [BO14], and SIA matrices.

Definition 5.8 (Upper and Lower Shadow [Eng97])**.** *Let $(P, \preceq)$ be a graded poset and let $R \subseteq P$ be such that $\exists k,\ \forall x \in R,\ rank(x) = k$. We call the* upper shadow

$$\nabla(R) = \{x \in P : \exists y \in R,\ y \preceq x,\ rank(x) = k+1\}\,.$$

Similarly, we define the lower shadow

$$\Delta(R) = \{x \in P : \exists y \in R,\ x \preceq y,\ rank(x) = k-1\}.$$

In Proposition 3.4 of Chapter 3, we have shown that the proper faces of $\mathcal{P}$ are the sets

$$F = \left\{ x \in \mathcal{P} \mid \forall i \in S_1, j \in S_2,\ \frac{1}{2}(x_i - x_j) = 1 \right\}, \tag{5.6}$$

for S_1 and S_2 some disjoint nonempty subsets of $\{1, \ldots, n\}$. Additionally, if d is the dimension of the face F, then $|S_1 \cup S_2| = n - d + 1$. The corresponding double-face is

$$F = \left\{ x \in \mathcal{P} \mid \forall i \in S_1, j \in S_2,\ \frac{1}{2}|x_i - x_j| = 1 \right\},$$

Therefore, the lower shadow of each single double-face F of $\mathcal{P}$ of dimension $2 \leq d \leq n-1$ contains

$$|\Delta(\{F\})| = 2(d-1)$$

elements (the double-faces obtained by adding an element to either S_1 or S_2). The upper shadow of a double-face of dimension $1 \leq d \leq n-2$ is the set of double-faces obtained by removing an element from either S_1 or S_2, while keeping them both nonempty. In the case that both $|S_1| \geq 2$ and $|S_2| \geq 2$, we obtain

$$|\nabla(\{F\})| = n - d + 1,$$

while in the case that $|S_1| = 1$ or $|S_2| = 1$ we have

$$|\nabla(\{F\})| = n - d.$$

The case $|S_1| = 1$ or $|S_2| = 1$ corresponds to double-faces of dimension $n-1$ whose upper shadow is $\mathcal{P}$, the face of dimension n:

$$|\nabla(\{F\})| = 1 = n - (n-1).$$

In turn, the upper shadow of a double-face of dimension $1 \leq d \leq n-1$ satisfies

$$|\nabla(\{F\})| \geq n - d.$$

Theorem 5.3. *In any dimension n, the lattice of double-faces of $\mathcal{P}$ has the Sperner property. One largest antichain is the set of double-faces of dimension $d^* = \lfloor n/3 \rfloor + 1$.*

Proof. For any set of double-faces of $\mathcal{P}$ of the same dimension $d \geq 1$ (that is, any subset of a rank level, with $d \geq 1$, in the lattice of double-faces), we define

$$E_+ = \{(F_1, F_2) : F_1 \in R,\ F_2 \in \nabla(\{F_1\})\}.$$

E_+ is thus the set of pairs of double-faces in respectively R and $\nabla(R)$ that are comparable. Since the upper shadow of each element of R has at least $n - d$ elements, we have

$$|E_+| \geq |R|(n - d).$$

Since the lower shadow of each element of ∇R contains exactly $2d$ elements, not all of which belonging to R, we have

$$|E_+| \leq |\nabla(R)|2d.$$

Combining the two inequalities, we obtain $|\nabla(R)| \geq |R|\frac{n-d}{2d}$ and

$$\forall\, 1 \leq d \leq \frac{n}{3},\ \forall R,\ \ |\nabla(R)| \geq |R|. \tag{5.7}$$

By a similar reasoning, we obtain $|\Delta(R)| \geq |R|\frac{2(d-1)}{n-d+2}$ and

$$\forall d \geq \frac{n+4}{3},\ \forall R,\ \ |\Delta(R)| \geq |R|. \tag{5.8}$$

Let now X be a largest antichain, let d^- be the smallest dimension of an element in $X$and let R^- be the intersection of the antichain with the level d^-. We have $d^- \geq 1$ because the only element with rank level $d^- = 0$ is the empty face $\emptyset$ and this element does not belong to any largest antichain. If $d^- \leq \frac{n}{3}$, Equation (5.7) tells us that the antichain

$$(X \backslash R^-) \cup \nabla(R^-)$$

has at least as many elements as X. We can repeat this process until the antichain contains only faces of dimension strictly larger than $\frac{n}{3}$. Similarly we use (5.8) to obtain an antichain with at least as many elements of rank strictly smaller than $\frac{n+4}{3}$. Since

$$\frac{n}{3} < d < \frac{n+4}{3}$$

has a unique integer solution $d^* = \lfloor n/3 \rfloor + 1$, the final antichain contains only faces of dimension d^*. ∎

5.3.1 A New Finiteness Bound for Contractive Sets

By Theorem 5.3, the largest antichain in the lattice of double-faces is the set of all double-faces of dimension $d^* = \lfloor n/3 \rfloor + 1$. From Equation (5.6), one can see that the number of faces of dimension d of $\mathcal{P}$ is equal to the number of pairs of disjoint nonempty subsets S_1, S_2 such that

$$|S_1 \cup S_2| = n - d + 1.$$

Therefore, the number of faces of dimension $1 \leq d \leq n-1$ of $\mathcal{P}$ is

$$f_d = \binom{n}{d-1}(2^{n-d+1} - 2).$$

The number of double-faces of dimension d is then

$$\binom{n}{d-1}(2^{n-d} - 1),$$

and the width of the lattice of double-faces of $\mathcal{P}$, which is equal to the number of double-faces of dimension $\lfloor n/3 \rfloor + 1$, is

$$p^* = \binom{n}{\lfloor n/3 \rfloor}(2^{n-\lfloor n/3 \rfloor - 1} - 1). \tag{5.9}$$

Combining this value of p^* with Theorem 5.1 and [CHJ15a, Proposition 1.a] yields the next theorem.

Theorem 5.4. *A finiteness bound for $\mathcal{P}$ is given by*

$$p^* = \binom{n}{\lfloor n/3 \rfloor}(2^{n-\lfloor n/3 \rfloor - 1} - 1) = \mathcal{O}\left(\frac{3^n}{\sqrt{n}}\right).$$

A finiteness result similar to Theorem 5.4 was known [BO14] with $B = \frac{1}{2}(3^n - 2^{n+1} + 1)$ instead of p^*. This value of B is in fact the number of double-faces of $\mathcal{P}$. Indeed, Corollary 3.2 indicates that $\mathcal{P}$ has $3^n - 2^{n+1} + 2$ faces, so that the number of proper faces is $3^n - 2^{n+1} + 1$ and the number of double-faces is B. The new bound p^*, on the other hand, is the size of the largest antichain in the lattice of double-faces. It is approximately equal to $\frac{3}{2\sqrt{\pi n}}B$, as illustrated in Figure 5.3.

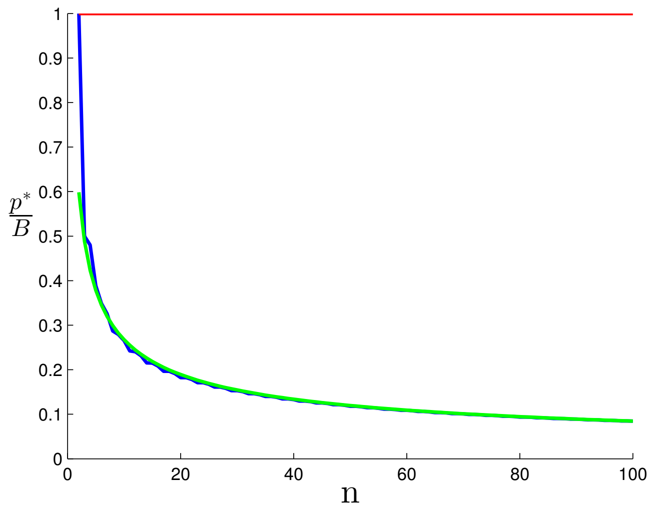


Figure 5.3: In blue: ratio between our bound p^* and the previous bound B. This ratio is below one (red) and close to $\frac{3}{2\sqrt{\pi n}}$ (green).

5.3.2 Stochastic Matrices

In this subsection, we prove next that Theorem 5.4 is tight for stochastic matrices. This is not a consequence of Theorem 5.2. Indeed, Theorem 5.2 applied to the polyhedron $\mathcal{P}$ guarantees that for any dimension n, there is a set of matrices such that Theorem 5.1 is tight for $\mathcal{P}$. However, the matrices in this set are not necessarily stochastic. We construct here a set of stochastic matrices for which Theorem 5.4 is tight.

This set contains p^* matrices. Therefore, Theorem 5.5 proves the tightness of p^* in general but not in the particular case where the number of matrices in the set is limited.

Theorem 5.2 is also not a consequence of this theorem since it applies to all centrally symmetric polyhedra and not only $\mathcal{P}$.

Theorem 5.5. *For any $n \geq 2$, there is a set of $n \times n$ stochastic matrices such that:*

- *There is a product of length $p^* = \binom{n}{\lfloor n/3 \rfloor}(2^{n-\lfloor n/3 \rfloor -1} - 1)$ whose powers do not converge to a rank-one matrix;*
- *For any product P of length $\leq p^* - 1$, the sequence of powers converges to a rank-one matrix.*

Proof. We will construct *stochastic* matrices $\{A_1, \dots, A_{p^*}\}$ that have the two properties:

$$\forall i,\ A_i F_i \subseteq O_{(i \bmod p^*)+1} \tag{5.10}$$

$$\forall i,\ A_i(\mathcal{P} \backslash (F_i \cup -F_i)) \subseteq \text{int}(\mathcal{P}). \tag{5.11}$$

Then the same argument as in the proof of Theorem 5.2 will allow allow us to conclude. Recall that each face F of $\mathcal{P}$ can be written as

$$F = \left\{ x \in \mathcal{P} \mid \forall i \in S_1, j \in S_2,\ \frac{1}{2}(x_i - x_j) = 1 \right\}$$

Or alternatively as

$$F = \{x \in \partial\mathcal{P} : \forall i \in S_1,\ x_i = \max_j x_j,\ \forall i \in S_2,\ x_i = \min_j x_j\} \tag{5.12}$$

for certain disjoint nonempty sets $S_1, S_2 \subset \{1, \dots, n\}$.

Let F_i, F_j be two faces satisfying $F_j = F_{(i \mod p^*)+1}$, let S_{1i} be the set S_1 of Equation 5.12 for the face F_i and let S_{1j}, S_{2i} and S_{2j} be defined similarly to S_{1i}. We will construct a matrix A_i such that any element of $A_i x$ in position $\in S_{1j}$ is equal to the average of the elements of x in positions in S_{1i}. This ensures that when $x \in F_i$, any element of $A_i x$ in position $\in S_{1j}$ is equal to $\max_k x_k$ (because it is equal to the average of some elements of x that are equal to $\max_k x_k$). Similarly, A_i will be constructed such that any element of $A_i x$ in position $\in S_{2j}$ is equal to the average of the elements of x in positions in S_{2i}. Finally, A_i will be such that any element of $A_i x$ in position $\in (\{1, \dots, n\} \backslash (S_{1j} \cup S_{2j}))$ is equal to the average of the elements of x in positions in $\{1, \dots, n\} \backslash (S_{1i} \cup S_{2i})$.

We now construct this matrix in the particular case where the elements of S_{1i} and S_{1j} are the first elements of $\{1, \dots, n\}$ and the elements S_{2i} and S_{2j} are the last elements of $\{1, \dots, n\}$. Let F_i be a face such that $S_{1i} = \{1, \dots, a_i\}$ and $S_{2i} = \{n - c_i + 1, \dots, n\}$ for some a_i and c_i and similarly let $F_j = F_{(i \mod p^*)+1}$ be such that $S_{1j} = \{1, \dots, a_j\}$ and $S_{2j} = \{n - c_j + 1, \dots, n\}$ for some a_j and c_j. Let $b_i = n - a_i - c_i$ and $b_j = n - a_j - c_j$. One matrix satisfying properties (5.10) and (5.11) is

$$A_i = \begin{pmatrix} +_{a_j \times a_i} & 0 & 0 \\ +_{b_j \times a_i} & +_{b_j \times b_i} & +_{b_j \times c_i} \\ 0 & 0 & +_{c_j \times c_i} \end{pmatrix}$$

where $+$ represents a positive element chosen such that the sum of the elements on each row sum to one. Let us see why property (5.10) is satisfied. Let $x \in F_i$, we have that the first a_j elements of $A_i x$ are averages of the first a_i

elements of x and therefore they are equal to $\max_k x_k$. Similarly, the last c_j elements of $A_i x$ are weighted averages of the last c_i elements of x and therefore, they are equal to $\min_k x_k$. The remaining elements are weighted averages of all elements of x and therefore they are strictly smaller than $\max_k x_k$ and strictly larger than $\min_k x_k$. These three facts imply $A_i x \in O_j$ and since it is the case for any $x \in F_i$, property (5.10) is satisfied. Property (5.11) is proved in a similar manner.

Without the assumption on the specific form of the faces F_i and F_j, the matrix A_i is the same up to some permutations of the rows and of the columns.

■

5.3.3 Relation with SIA Matrices

Theorem 5.4 establishes that

$$p^* = \binom{n}{\lfloor n/3 \rfloor}(2^{n-\lfloor n/3 \rfloor -1} - 1)$$

is a finiteness bound for $\mathcal{P}$. Therefore, for a set of *stochastic* matrices, all left-infinite products converge to a rank-one matrix if and only if *all periodic left-infinite products, with period smaller than or equal to p^*, converge to a rank-one matrix*. This is in fact equivalent to the condition that all products of length smaller than or equal to p^* are SIA[2]. In the case of stochastic matrices, Theorem 5.4 thus becomes.

Theorem 5.6. *Let S be a finite set of stochastic matrices. Any left-infinite product of matrices from S converges to a rank-one matrix if and only if any product of length smaller than or equal to p^* (as defined in Equation* (5.2)*) is SIA.*

5.4 Conclusion

We have studied sets of matrices that admit a nonincreasing polyhedral seminorm, and we wondered whether all infinite products of these matrices map the state space onto points whose seminorm is equal to zero (the consensus problem is a particular case of this setting). We have improved the available finiteness bound by leveraging the combinatorial structure of (an abstraction of) the dynamical system described by these matrices. We have

[2]We recall that a matrix is SIA if and only if it is stochastic and the sequence of its powers converge to a rank-one matrix.

shown that the largest antichain in the lattice of double-faces provides a tight finiteness bound. In the particular case of the polyhedron $\mathcal{P}$, we were able to compute the size of the largest antichain to obtain a finiteness bound of $p^* = W(\mathcal{P}) = \binom{n}{\lfloor n/3 \rfloor}(2^{n-\lfloor n/3 \rfloor -1} - 1)$, which is an improvement of about $\frac{3}{2\sqrt{\pi n}}$ over the previously known bound.

A question that we have left open is the influence of the number of matrices of the set on the finiteness bound. The examples that we constructed to prove the tightness of our bounds contain many matrices. It is not clear what the best finiteness bound would be, for example, for sets containing only two matrices. Tackling this question could be done by analyzing the structure of the graph of double-faces. Indeed, a finiteness bound is a bound on the length of the shortest cycle in this graph, and long cycles might be much harder to construct with only two matrices.

Chapter 6

Application to Other Polyhedra

We apply techniques developed in Chapters 4 and 5 to other polyhedra.

In Section 6.1, we use the nonnegative orthant $\mathbb{R}^n_+ = \{x \in \mathbb{R}^n \mid \forall i,\ x_i \geq 0\}$, to study sets of nonnegative matrices all sufficiently long products of which are positive. We call these sets *all-products-primitive matrix sets*.

- We recover a result by Cohen and Sellers [CS82] stating that if a set is all-products-primitive, then all products of length $2^n - 2$ are *positive*.
- We show that the face lattice of $\mathbb{R}^n_+$ has the Sperner property, from which we deduce that a set is all-products-primitive iff all products of length $\binom{n}{n/2}$ are *primitive*.
- We show that deciding all-products-primitivity can be done with a singly exponential time-complexity.
- We show that deciding all-products-primitivity is NP-hard.

In Section 6.2, we construct a centrally symmetric polytope whose lattice of double-faces does not have the Sperner property, showing that not all lattice of double-faces have this property. This polytope is the simplest example that we were able to construct. Its construction uses old ideas from Danzer and Eckhoff [Eck06].

In the previous chapter, we have seen (Theorem 5.1) that $p^* = W(\mathcal{Q})$ is a tight finiteness bound. In general $W(\mathcal{Q})$ may be hard to compute but, in the particular case of the polyhedron $\mathcal{P}$, $W(\mathcal{Q})$ is equal to the largest rank level of the lattice of double-faces of $\mathcal{P}$. The new example that we constuct shows that p^* is not always equal to the largest rank level of the lattice of double-faces.

6.1 Primitivity of Matrix Sets

We consider the problem of determining whether all products of a set of nonnegative matrices are positive.

Definition 6.1 (All-Products-Primitivity of a Matrix Set). *We say that a set of nonnegative matrices is all-products-primitive if there is a length ℓ such that all products of length ℓ are positive.*

All-products-primitivity is a generalization of the primitivity of a single matrix – i.e., the existence of a power of that matrix that is positive – in the sense that a set that contains a single matrix is all-products-primitive if and only if this matrix is primitive. All-products-primitivity is sometimes called *primitivity* [CS82] but we call it here all-products-primitive to avoid any confusion with other notions of primitivity, such as the existence of a positive product [BJO15].

Conditions for all-products-primitivity has been known for a long time, and these conditions can be verified algorithmically.

Theorem 6.1 (Exponent of primitive sets [CS82]). *A set of nonnegative matrices S is all-products-primitive if and only if all products of length $2^n - 2$ are positive. The length $2^n - 2$ cannot be decreased.*

The analysis of the polyhedron $\mathbb{R}_+^n$ provides an alternative proof for the "only if" part of this result. Nonnegative matrices leave the nonnegative orthant $\mathbb{R}_+^n$ invariant. Additionally, a nonnegative matrix is positive if and only if it contracts $\mathbb{R}_+^n \backslash \{0\}$:

$$A\mathbb{R}_+^n \backslash \{0\} \in \text{int}(\mathbb{R}_+^n).$$

The open faces of the polyhedron $\mathbb{R}_+^n$ are the sets that can be written as

$$F = \{x \mid \forall i \in S,\ x_i > 0 \text{ and } \forall i \notin S,\ x_i = 0\} \tag{6.1}$$

for S some subset of $\{1, \dots, n\}$. The number of open faces is thus 2^n (including $\text{int}(\mathbb{R}_+^n)$).

If the set is all-products-positive, then the image of $\mathbb{R}_+^n \backslash \{0\}$ by sufficiently long products of matrices from S is $\text{int}(\mathbb{R}_+^n)$. By Lemma 4.1, the image of each open face of $\mathbb{R}_+^n \backslash \{0\}$ is in another open face of $\mathbb{R}_+^n \backslash \{0\}$ and we can construct the graph of faces of $\mathbb{R}_+^n \backslash \{0\}$[1]. This graph has $2^n - 1$ nodes and does not

[1] In this case it will be a graph of faces and not a graph of double-faces as in Chapter 4 because the polyhedron is not centrally symmetric. The idea is exactly the same

contain any cycle other than the loop from int($\mathbb{R}^n_+$) to itself. Therefore, the longest possible path from a node to int($\mathbb{R}^n_+$) has a length of $2^n - 2$. This means that for any product P of length $2^n - 2$ and any vector $x \in \mathbb{R}^n_+ \backslash \{0\}$, $Px \in$ int($\mathbb{R}^n_+$) and therefore P is positive, which concludes our alternative proof of Theorem 6.1.

The face lattice of $\mathbb{R}^n_+$ is isomorphic to the *Boolean lattice*, which is the poset of all subsets of an n-element set, ordered by inclusion. This isomorphism can be seen from Equation (6.1), where there is one face F for each subset S of $\{1, \dots, n\}$ and a (closed) face $F_1 = \{x \mid \forall i \in S_1,\ x_i \geq 0 \text{ and } \forall i \notin S_1,\ x_i = 0\}$ is included in a closed face $F_2 = \{x \mid \forall i \in S_2,\ x_i \geq 0 \text{ and } \forall i \notin S_2,\ x_i = 0\}$ if and only if $S_1 \subseteq S_2$. In fact, the boolean lattice was the very first poset for which the Sperner property has been established [Spe28].

Theorem 6.2 (Sperner property of the Boolean lattice [Spe28])**.** *Let n be a positive integer and $\mathcal{F}$ be a family of subsets of $\{1, \dots, n\}$ such that no member of $\mathcal{F}$ is included in another member of $\mathcal{F}$. Then*

- $\mathcal{F} \leq \binom{n}{\lfloor n/2 \rfloor}$.
- *Equality holds iff $\mathcal{F}$ is the set of subsets containing exactly $\lfloor n/2 \rfloor$ elements.*

Thanks to this theorem, we know that the largest antichain of the face lattice of $\mathbb{R}^n_+$ contains $\binom{n}{\lfloor n/2 \rfloor}$ elements, from which we obtain.

Theorem 6.3. *A set of nonnegative matrices S is all-products-primitive if and only if all products of length $\leq \binom{n}{n/2}$ are primitive. The length $\binom{n}{n/2}$ cannot be decreased.*

Proof. Only if: If a product (of any length) is not primitive, then its powers are not positive and the set is not all-products-primitive.
If: If the set is not all-products-primitive, there is a cycle in the graph of faces. This cycle corresponds to a product of matrices from S that is not *primitive.* If this cycles contains two nodes that correspond to faces that are included in one another, it is possible to construct a shorter cycle (as was done in Chapter 4). In turn, there is a cycle that contains only nodes corresponding to faces that are not included in one another. These faces form an antichain in the face lattice of $\mathbb{R}^n_+$, and this antichain has a size $\leq \binom{n}{n/2}$. Finally, the cycle of length $\leq \binom{n}{n/2}$ corresponds to a nonprimitive product of length $\leq \binom{n}{n/2}$. ■

By constructing the graph of faces of $\mathbb{R}^n_+$ for a given set S and by checking whether it contains cycles, the all-products-primitivity can be decided in singly exponential time.

The graph of faces of $\mathbb{R}^n_+$ also provides a way to decide whether a set is all-products-primitive.

Theorem 6.4. *Given a set S of nonnegative matrices, deciding whether S is all-products-primitive can be decided algorithmically with a singly exponential time complexity.*

Proof. This can be done by constructing the graph of faces of $\mathbb{R}^n_+$ for a given set S and by checking whether it contains cycles. The number of nodes of this graph is singly exponential in the size of the matrices and checking whether it contains a cycle is linear in the number of edges (which is at most quadratic in the number of nodes). ■

This decision procedure has an exponential space complexity, that comes from the storage of the graph of faces.

We conclude this section by proving that deciding all-products-complexity is NP-hard.

Theorem 6.5. *Given a finite set S of nonnegative matrices, deciding whether S is all-products-primitive is NP-hard.*

Proof. The reduction is from the Consensus-Set problem[2] in the particular case that all matrices share a common row, that has a 1 on the diagonal and 0 everywhere else on this row. In the reduction for the NP-hardness proof of Consensus-Set [BO14], it can be seen that all the constructed matrices have this common row and thus deciding Consensus-Set in this restricted case is also NP-hard.

We can assume, without loss of generality, that the common row is the first one. Let thus S be a set of stochastic matrices with a first row equal to $\begin{pmatrix}1 & 0 & \dots & 0\end{pmatrix}$ and let S' be a set in which the first rows of the matrices of S have been replaced by $\begin{pmatrix}1 & 1 & \dots & 1\end{pmatrix}$.

Claim: S is a consensus set if and only if S' is all-products-primitive.

If: Suppose that S is a consensus set. Then there is ℓ such that all products of length ℓ of matrices from S have a positive column [BO14, Corollary 2.10]. This positive column is the first column. Indeed, no other column can be positive in products of matrices that their first rows equal to $\begin{pmatrix}1 & 0 & \dots & 0\end{pmatrix}$. Therefore, all products of length ℓ of matrices from S' also have a positive first column, and all products of length $\ell + 1$ of matrices from S' are positive,

[2]We recall that consensus sets are contractive sets in the case of stochastic matrices.

because the product of a nonnegative matrix that has a positive first column by a nonnegative matrix that has a positive first row is positive.

Only if: Suppose that S' is all-products-primitive. Then there is ℓ such that all products of length ℓ of matrices from S' have a positive first column (because they are positive). We show that for a given product $P' \triangleq A'_{\sigma(t)} \dots A'_{\sigma(2)} A'_{\sigma(1)}$ of matrices from S' that has positive first column, the same product on the matrices of S: $P \triangleq A_{\sigma(t)} \dots A_{\sigma(2)} A_{\sigma(1)}$ also has a positive first column.

By definition of the matrix product and using the nonnegativity of the matrices, we have that a positive element P'_{i1} implies the existence of a sequence $j_t, j_{t-1}, j_{t-2} \dots j_0$ with $j_t = i$ and $j_0 = 1$ such that

$$\forall 1 \leq k \leq t,\ (A'_{\sigma(k)})_{j_k j_{k-1}} > 0.$$

We want to show that there exists such a sequence that does not use any element that is positive in $A'_{\sigma(k)}$ but not in $A_{\sigma(k)}$. These elements are the elements $(A'_{\sigma(k)})_{j_k j_{k-1}}$ with $j_k = 1$ and $j_{k-1} \neq 1$. If k^* is the largest k such that $j_k = 1$, then we can take

$$j_{k^*-1} = j_{k^*-2} = \dots = j_0 = 1$$

and we have a sequence of positive elements $(A'_{\sigma(k)})_{j_k j_{k-1}}$ that does not use any element that is positive in $A'_{\sigma(k)}$ but not in $A_{\sigma(k)}$. Hence P_{i1} is also positive and we can conclude that P also has a positive column.

In turn, every product of length ℓ of matrices from S has a positive first column and S is a consensus set. ■

6.2 A Face Lattice without the Sperner Property

We construct a polytope that is centrally symmetric around the origin and for which the lattice of double-faces does not have the Sperner property. Before presenting the example, we need some definitions.

Definition 6.2 (Simplicial and simple polytopes). *An n-dimensional polytope is* simplicial *if its facets contain exactly n vertices. An n-dimensional polytope is* simple *if its vertices are adjacent to exactly n edges.*

An octahedron is an example of a simplicial polytope because each of its faces is a triangle. A cube is an example of a simple polytope.

Definition 6.3 (F-vector). *The* f-vector *of a polytope of dimension n is the n-dimensional vector whose i^{th} element is equal to the number of faces of dimension $i-1$.*

For example, the f-vector of a cube is equal to $\begin{pmatrix} 8 & 12 & 6 \end{pmatrix}$. We give here an example of a centrally symmetric polytope of dimension 4 whose face lattice does not have the Sperner property. We have represented the equivalent construction in dimension 3 in Figures 6.1, 6.2, 6.3 and 6.4. However, our example in dimension 3 has the Sperner property and we have to construct it in dimension 4 to not have the Sperner property.

Example 6.1. *We use the* connected sum $\#$ *presented in [Zie95, Eck06]. The first step (illustrated in dimension 3 in Figure 6.1) is to cut off a vertex from a simple polytope P' (we will use a hypercube). The second step (Figure 6.2) is to apply a suitable projective transformation T to the rest of P'. The third and last step (Figure 6.4) is to glue it on a facet of a simplicial polytope P (we will use a hyperoctahedron). The transformation T can be chosen such that the result $P\#P'$ is a convex polytope whose faces are those of P except the facet on which the transformed sliced P' was glued, and those of P' except the cut off vertex. In particular, all open faces of P', except the vertex that has been cut off, remain present (albeit modified) and do not intersect with P. Similarly, all faces of P, except the face on which the modified P' has been glued, remain present and do not intersect with the modified P'. In turn, the f-vector of $P\#P'$ is equal to*

$$\begin{aligned} & f + f' - \begin{pmatrix} 1 & 0 & \dots & 0 & 1 \end{pmatrix} \\ = & \begin{pmatrix} f_0 + f'_0 - 1 & f_1 + f'_1 & \dots & f_{n-1} + f'_{n-1} - 1 \end{pmatrix}, \end{aligned}$$

where f is the f-vector of P and f' that of P'.

We use this construction twice to glue two 4-dimensional *hypercubes*

$$\mathcal{C} = \{x \in \mathbb{R}^4 : \|x\|_\infty \leq 1\}$$

(that will be the P' of the connected sum described above) on opposite faces of a hyperoctahedron

$$\mathcal{O} = \{x \in \mathbb{R}^4 : \|x\|_1 \leq 1\}$$

(that will correspond to P above). The hypercubes have an f-vector of

$$\begin{pmatrix} 16 & 32 & 24 & 8 \end{pmatrix},$$

while the hyperoctahedra have an f-vector of

$$\begin{pmatrix} 8 & 24 & 32 & 16 \end{pmatrix}.$$

The connected sum is a convex polytope with an f-vector equal to

$$\begin{aligned} & 2f + f' - \begin{pmatrix} 2 & 0 & 0 & 2 \end{pmatrix} \\ = & 2\begin{pmatrix} 16 & 32 & 24 & 8 \end{pmatrix} + \begin{pmatrix} 8 & 24 & 32 & 16 \end{pmatrix} - \begin{pmatrix} 2 & 0 & 0 & 2 \end{pmatrix} \\ = & \begin{pmatrix} 38 & 88 & 80 & 30 \end{pmatrix}. \end{aligned}$$

The largest rank level has thus a size of 88 *(and it is the level of the faces of dimension 1). To construct a large antichain in the face lattice, we can take the 64 faces of dimension 1 from the hypercubes and the 32 faces of dimension 2 of the hyperoctahedron. This set of faces is an antichain in the face lattice because:*

- *The faces of dimension 2 of $\mathcal{O}$ cannot be subsets of the faces of dimension 1 of the transformed hypercube, because of the dimensions.*
- *The faces of dimension 1 of the transformed sliced hypercubes are not subsets of $\mathcal{O}$ and can therefore not be subsets of the faces of dimension 2 of $\mathcal{O}$. These faces of dimension 1 cannot be subsets of $\mathcal{O}$ because*
  - *Each face of dimension 1 of the hypercubes contains two vertices.*
  - *At least one of them corresponds to an original vertex of the hypercube (i.e., it is not a vertex created by the cut).*
  - *After transformation, this vertex does not belong to $\mathcal{O}$. Hence, the face of dimension 1 cannot be a subset of $\mathcal{O}$.*
- *The faces of dimension 2 of $\mathcal{O}$ are different and therefore cannot be included in one another.*
- *Similarly, the faces of the sliced hypercube cannot be included in one another.*

We conclude that the set that we constructed is an antichain in the face lattice. We obtain an antichain of $96 > 88$ *elements, which proves that the polytope does not have the Sperner property. The constructed polytope can be chosen to be centrally symmetric. To illustrate, Figures 6.1, 6.2, 6.3 and 6.4 present the same construction in dimension 3. The difference is that the antichain that is obtained is not large enough and the polytope is Sperner.*

6.2.1 Existence of a Suitable Projective Transformation

This subsection deals with the technical details of the transformation that makes Example 1 possible. We prove that we indeed obtain a convex polytope with the structure that we claimed. We start by recalling some basics on

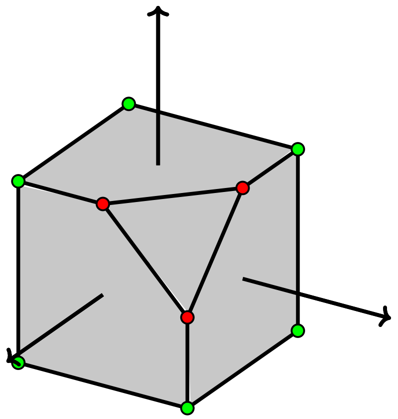


Figure 6.1: sliced cube. In green a set of faces making up an antichain in the face lattice and with none of these faces being entirely in the new face $\mathcal{C} \cap \{x_1 + x_2 + x_3 \leq 2\}$.

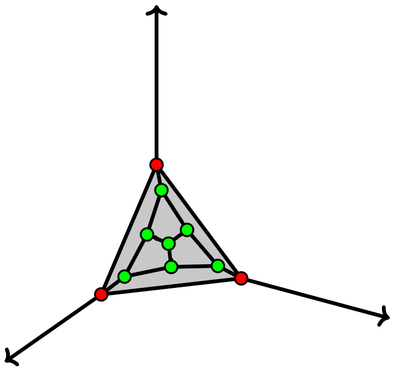


Figure 6.2: A transformation of the sliced cube, preserving its combinatorial structure.

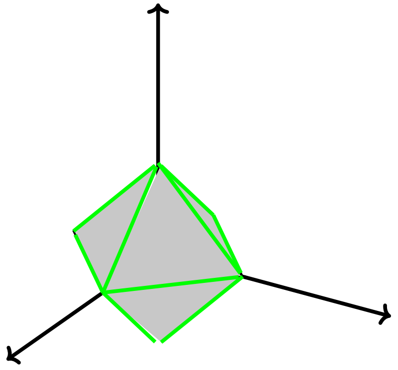


Figure 6.3: Octahedron. The set of green faces is an antichain in the face lattice.

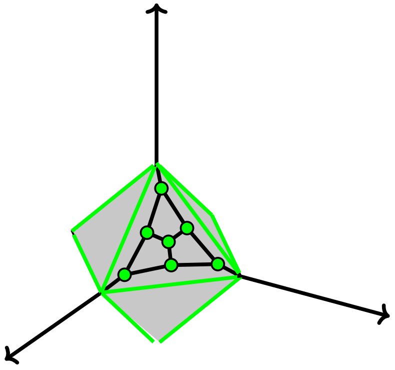


Figure 6.4: Resulting polytope. The set of green faces is an antichain in the face lattice.

projective transformations, a class of transformations that preserve collinearity [Grü03, Page 4].

Definition 6.4 (Projective transformation [Grü03])**.** *A projective transforma-*

tion *is a transformation of the form*

$$T(x) = \frac{Ax + b}{c^\top x + \delta},$$

with $A \in \mathbb{R}^{n\times n}$*,* $b,\ c \in \mathbb{R}^n$*,* $\delta \in \mathbb{R}$ *and at least one of* c *and* δ *being different from zero.*

Definition 6.5 (Nonsingular projective transformation [Grü03])**.** *A projective transformation* T *is called* nonsingular *if*

$$\begin{pmatrix} A & b \\ c^\top & \delta \end{pmatrix}$$

is invertible.

Definition 6.6 (Permissible projective transformation [Grü03])**.** *A projective transformation* T *is called* permissible for a set K *if*

$$\{x \mid c^\top x + \delta = 0\} \cap K = \emptyset.$$

Theorem 6.6 (Theorem 3.1.4 in [Grü03])**.** *Any permissible projective transformation of a polytope is a polytope.*

Definition 6.7 (Combinatorial equivalence [Grü03])**.** *Two polytopes are said* combinatorially equivalent *if their face lattices are the same.*

Theorem 6.7 (Theorem 3.2.3 in [Grü03])**.** *Let* $\mathcal{Q}$ *be a polytope. If* T *is a nonsingular projective transformation, permissible for* $\mathcal{Q}$*, then* $\mathcal{Q}$ *and* $T(\mathcal{Q})$ *are combinatorially equivalent.*

We now present formally the 4-dimensional construction. The hyperoctahedron $\mathcal{O}$ can be described by the equations

$$\forall b \in \{-1, 1\}^4,\ b^\top x \le 1.$$

The hypercubes $\mathcal{C}$ can be described by the equations

$$\forall i \in \{1, 2, 3, 4\},\ |x_i| \le 1, \tag{6.2}$$

while the sliced hypercube $\mathcal{S}$ is described by the equations

$$\begin{cases} \forall i \in \{1, 2, 3, 4\},\ |x_i| \le 1 \\ \mathbf{1}^\top x \le 3. \end{cases} \tag{6.3}$$

We apply to the sliced cube the projective transformation

$$T(x) = \frac{(I + \frac{2}{3}\mathbf{1}\mathbf{1}^\top)x - 3.\mathbf{1}}{3.\mathbf{1}^\top x - 10} \tag{6.4}$$

and we want to prove that

$$\mathcal{R} \triangleq T(\mathcal{S}) \cup \mathcal{O} \cup -T(\mathcal{S})$$

is a polytope and that its f-vector is

$$2f + f' - \begin{pmatrix} 2 & 0 & 0 & 2 \end{pmatrix} = \begin{pmatrix} 38 & 88 & 80 & 30 \end{pmatrix},$$

where f and f' are the f-vectors of the hypercube $\mathcal{C}$ and of the hyperoctahedron $\mathcal{O}$ respectively.

Lemma 6.1. *Let T be the projective transformation defined in* (6.4) *and $\mathcal{S}$ the 4-dimensional sliced hypercube defined in* (6.3)*. $T(\mathcal{S})$ and $-T(\mathcal{S})$ are polytopes.*

Proof. We prove that transformation T is permissible (Definition 6.6) for $\mathcal{S}$:

$$\forall x \in \mathcal{S},\ 3.\mathbf{1}^\top x - 10 \neq 0,$$

because $x \in \mathcal{S}$ satisfies

$$\mathbf{1}^\top x \leq 3.$$

By Theorem 6.6, the permissibility of the transformation implies that $T(\mathcal{C})$ and $-T(\mathcal{C})$ are polytopes. ■

Lemma 6.2. *Let T be the projective transformation defined in* (6.4) *and $\mathcal{S}$ the 4-dimensional sliced hypercube defined in* (6.3)*. The set $T(\mathcal{S})$ is a full-dimensional polytope with the same combinatorial structure as $\mathcal{S}$.*

Proof. We have already proved in Lemma 6.1 that T is permissible for $\mathcal{S}$. We prove the nonsingularity of T and Theorem 6.7 allows us to conclude. The transformation T is singular if the 5×5 matrix

$$\begin{pmatrix} I + \frac{2}{3}\mathbf{1}\mathbf{1}^\top & -3.\mathbf{1} \\ 3.\mathbf{1}^\top & -10 \end{pmatrix}$$

is full rank. The block (-10) has obviously a rank of 1 and the Schur complement

$$I + \frac{2}{3}\mathbf{1}\mathbf{1}^\top - \frac{9}{10}\mathbf{1}\mathbf{1}^\top = I - \frac{7}{30}\mathbf{1}\mathbf{1}^\top$$

is diagonally dominant and therefore positive definite. In turn, the matrix is full rank. ■

Lemma 6.3 (Vertices and facets of the 4-dimensional transformed sliced hypercube)**.** *The vertices of $T(\mathcal{S})$ are*

$$\begin{aligned} &\begin{pmatrix}1 & 0 & 0 & 0\end{pmatrix} \\ \frac{1}{6}&\begin{pmatrix}4 & 1 & 1 & 1\end{pmatrix} \\ \frac{1}{5}&\begin{pmatrix}2 & 2 & 1 & 1\end{pmatrix} \\ \frac{1}{24}&\begin{pmatrix}8 & 8 & 8 & 5\end{pmatrix} \\ \frac{10}{33}&\begin{pmatrix}1 & 1 & 1 & 1\end{pmatrix} \end{aligned} \tag{6.5}$$

and the points obtained by permuting elements of these vectors. Its facet inequalities are (every $1 \le i \le 4$ generates different facet inequalities):

$$\begin{aligned} (\boldsymbol{1} - e^i)^\top x &\le 1 \\ \left(\frac{8.\boldsymbol{1} + e^i}{10}\right)^\top x &\le 1 \\ \boldsymbol{1}^\top x &\ge 1. \end{aligned} \tag{6.6}$$

Proof. The vertices of the 4-dimensional sliced hypercube $\mathcal{S}$ are the vertices of the hypercube $\mathcal{C}$ except the vertex $\begin{pmatrix}1 & 1 & 1 & 1\end{pmatrix}$ and adding the four vertices created by the additional constraint $\mathbf{1}^\top x \le 3$. The vertices of $T(\mathcal{S})$ are thus obtained by computing the projective image of these vertices.

$$\begin{aligned} T\left(\begin{pmatrix}0 & 1 & 1 & 1\end{pmatrix}\right) &= \begin{pmatrix}1 & 0 & 0 & 0\end{pmatrix} \\ T\left(\begin{pmatrix}-1 & 1 & 1 & 1\end{pmatrix}\right) &= \frac{1}{6}\begin{pmatrix}4 & 1 & 1 & 1\end{pmatrix} \\ T\left(\begin{pmatrix}-1 & -1 & 1 & 1\end{pmatrix}\right) &= \frac{1}{5}\begin{pmatrix}2 & 2 & 1 & 1\end{pmatrix} \\ T\left(\begin{pmatrix}-1 & -1 & -1 & 1\end{pmatrix}\right) &= \frac{1}{24}\begin{pmatrix}8 & 8 & 8 & 5\end{pmatrix} \\ T\left(\begin{pmatrix}-1 & -1 & -1 & -1\end{pmatrix}\right) &= \frac{10}{33}\begin{pmatrix}1 & 1 & 1 & 1\end{pmatrix}. \end{aligned} \tag{6.7}$$

The facet inequalities of $\mathcal{S}$ are the facet inequalities of the hypercube $\mathcal{C}$ and the facet inequality created by the cut $\mathbf{1}^\top x \le 3$. It can be verified that

$$\begin{aligned} e^{i\top} x \le 1 &\Leftrightarrow (\mathbf{1} - e^i)^\top T(x) \le 1 \\ -1 \le e^{i\top} x &\Leftrightarrow \left(\frac{8.\mathbf{1} + e^i}{10}\right)^\top T(x) \le 1 \\ \mathbf{1}^\top x \le 3 &\Leftrightarrow \mathbf{1}^\top T(x) \ge 1. \end{aligned}$$

■

Lemma 6.4. *The set $\mathcal{R} = T(\mathcal{S}) \cup \mathcal{O} \cup -T(\mathcal{S})$ is defined by the facet constraints of $T(\mathcal{S})$, $\mathcal{O}$ and $-T(\mathcal{S})$ except the facets by with these polytopes are glued together. That is, $\mathcal{R}$ is the set of points satisfying*

$$(\boldsymbol{1} - e^i)^\top x \leq 1 \tag{6.8}$$

$$\left(\frac{8.\boldsymbol{1} + e^i}{10}\right)^\top x \leq 1 \tag{6.9}$$

$$-(\boldsymbol{1} - e^i)^\top x \leq 1 \tag{6.10}$$

$$-\left(\frac{8.\boldsymbol{1} + e^i}{10}\right)^\top x \leq 1 \tag{6.11}$$

$$\forall\, b \in \{-1, 1\}^4 \backslash \{\boldsymbol{1}, -\boldsymbol{1}\},\ b^\top x \leq 1. \tag{6.12}$$

Proof. Let x be a point that satisfies (6.8) – (6.12). At least one of the following conditions is satisfied.

- If $\mathbf{1}^\top x \geq 1$, then $x \in T(\mathcal{S})$ because (6.8), (6.9) and $\mathbf{1}^\top x \geq 1$ are the facet constraints of $T(\mathcal{S})$, as we have seen in Equation (6.6).
- If $\mathbf{1}^\top x \leq 1$ and $-\mathbf{1}^\top x \leq 1$, then $x \in \mathcal{O}$ because these contraints and (6.12) are all the facet constraints of $\mathcal{O}$.
- If $-\mathbf{1}^\top x \leq 1$, then $x \in -T(\mathcal{S})$ because (6.10), (6.11) and $-\mathbf{1}^\top x \leq 1$ are the facet constraints of $-T(\mathcal{S})$.

In turn, $x \in (T(\mathcal{S}) \cup \mathcal{O} \cup -T(\mathcal{S})) = \mathcal{R}$.

On the other hand, the vertices of $\mathcal{O}$: $\pm e^i$ satisfy (6.8) – (6.12) and by convexity of $\mathcal{O}$, all points of $\mathcal{O}$ satisfy these constraints. Similarly, the vertices of $T(\mathcal{S})$ and $-T(\mathcal{S})$ (given in (6.7)) satisfy (6.8) – (6.12) as well and therefore all points of $T(\mathcal{S})$ and $-T(\mathcal{S})$ satisfy these constraints. In turn, all points of $\mathcal{R}$ satisfy (6.8) – (6.12). ■

We have now proven that $\mathcal{R}$ is a polytope (it is a polyhedron by Lemma 6.4 and it is also clearly bounded). It thus has the f-vector claimed in Example 6.1 because the facets are as claimed and the lower dimensional faces are intersections of facets.

Corollary 6.1. *The set $\mathcal{R} = T(\mathcal{S}) \cup \mathcal{O} \cup -T(\mathcal{S})$ is a polytope. Its f-vector is*

$$\begin{pmatrix} 38 & 88 & 80 & 30 \end{pmatrix}.$$

We have thus constructed the polytope $\mathcal{R}$ as claimed in Example 6.1. It has an antichain of 96 elements (see Example 6.1) and a largest rank level of 88. We thus have an example of a centrally symmetric polytope whose face lattice does not have the Sperner property.

Part II

Almost Contractive Sets

Is there one left-infinite product that converges to a rank-one matrix?

Chapter 7

Synchronizing Automata

In this short chapter, we present an introduction to automata and synchronization. Automata are simple models of computation involving a set of states and a set of possible actions that define transitions between the states. Synchronization is a sequence of actions that bring the automaton into a given state, independently of the initial state. An important question is the number of steps that are needed to achieve synchronization. The Černý conjecture asserts that if an automaton with n states has a synchronizing sequence, then it has one of length at most $(n-1)^2$. This conjecture has been open for more than 50 years and is still an active research topic [Szy18, GJT15, Vol08].

Automata can be represented by matrix sets and synchronization by products that have a positive column. In the next chapter, we will use results from automata theory in our study of almost contractive sets. The results of this chapter are all known results from the literature.

- In Section 7.1, we introduce automata.
- In Section 7.2, we present the Černý conjecture.
- In Section 7.3, we present an algorithm that decides whether an automaton is synchronizing.
- In Section 7.4, we present some negative complexity result regarding the computation of the reset threshold of automata.

7.1 Automata and Synchronization

In this section, we define automata and some of their representations.

Definition 7.1 (Automaton)**.** *A* deterministic finite automaton (DFA) *or simply* automaton *is a tuple $\mathcal{A} = \langle Q, \Sigma, \delta \rangle$, where Q is a set of states, Σ is an input alphabet and $\delta : Q \times \Sigma \to Q$ is a transition function defining an action of the letters in Σ on Q.*

Givin a state and a letter, the automaton will be in a new state, given by the transition function. Given an initial state and a sequence of input letters, called a *word*, the automaton will be in a certain state, given by the composition of the transition function. An automaton is often defined with an *initial state* and a set of ACCEPT *states*. A word is said to be accepted if starting from the initial state and applying the word sends the automaton to one of the ACCEPT states. The set of words that an automaton $\mathcal{A}$ accepts is called the *language recognized by* $\mathcal{A}$. An automaton without initial and ACCEPT states, as defined in Definition 7.1 is sometimes called a *semi-automaton*. In this thesis, however, we will use definitions and notations of [Vol08] and we will call it simply an automaton.

An automaton can be represented by a graph, called the state diagram, in which the nodes represent the set of states, the edges represent the transition function and the label of the edges are the different actions.

Example 7.1. *An automaton is represented in Figure 7.1. If the automaton is in state 1 and the letter a is applied, the automaton will be in state 2. If the automaton is in state 1 and the word aab is applied, it will be in state 3.*

It is also common to represent an automaton by a set of matrices, in which the set of rows represents the set of states and the matrices represent the different letters or actions. The matrices will have one 1 in each row, and the matrix will have a 1 in position i, j if the corresponding action sends the state i to the state j:

$$i \times a \to j.$$

The state of the automaton is then represented as a row vector $v^\top$ with a 1 in the column corresponding to the state and 0 everywhere else. The new state after applying action a is then computed with the product $v^\top A$, where A is the matrix representing action a.

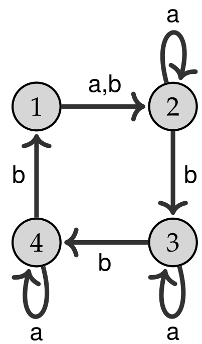


Figure 7.1: A simple automaton with four states and two letters.

Example 7.2. *The automaton of Example 7.1 can be represented by the set of matrices*

$$\left\{ A = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}, B = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \end{pmatrix} \right\}.$$

The state 1 is represented by

$$\begin{pmatrix} 1 & 0 & 0 & 0 \end{pmatrix}$$

and the transition when letter a is applied by

$$\begin{pmatrix} 1 & 0 & 0 & 0 \end{pmatrix} A = \begin{pmatrix} 0 & 1 & 0 & 0 \end{pmatrix}.$$

The word aab will correspond to the product

$$AAB = \begin{pmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \end{pmatrix}$$

and in this representation as well, we can see that the word aab sends state 1 onto state 3, because the element in position $1,3$ in AAB is equal to 1, and thus

$$\begin{pmatrix} 1 & 0 & 0 & 0 \end{pmatrix} AAB = \begin{pmatrix} 0 & 0 & 1 & 0 \end{pmatrix}.$$

In this thesis, we call *automaton matrices* the matrices that can possibly appear in this representation.

Definition 7.2 (Automaton matrix)**.** *A matrix is called* automaton matrix *if it has exactly one 1 on each row and 0 everywhere else.*

It is easy to verify that automaton matrices are stochastic and the product of two automaton matrices is an automaton matrix.

7.2 Reset Threshold and Černý Conjecture

This section is an introduction to synchronization and the Černý conjecture. An automaton is synchronizing if there is a sequence of actions that sends all states on the same state. Such a sequence is called a *synchronizing word*. To a synchronizing word corresponds a product of the matrices that has a positive column, and to the state on which all states are sent corresponds the column that is positive. For a synchronizing automaton, the *reset threshold* is defined as the length of the shortest synchronizing word.

Conjecture 7.1 (Černý conjecture [Čer64]). *Let $\mathcal{A}$ be a synchronizing automaton with n states. $\mathcal{A}$ has a reset threshold of at most $(n-1)^2$.*

The best known bound on the length of the reset threshold is $\frac{114n^3}{685} + \mathcal{O}(n^2)$ [Szy18] which is a slight improvement over the following theorem, due to Pin and Frankl.

Theorem 7.1 (Bound on the Reset Threshold [Pin83, Fra82]). *Let $\mathcal{A}$ be a synchronizing automaton with n states. $\mathcal{A}$ has a synchronizing word of length at most $\frac{n^3-n}{6}$.*

It is known that, if the Černý conjecture is true, then it is tight for all n. Indeed, for any n there is an synchronizing automaton whose reset threshold is $(n-1)^2$. We now describe these automata.

Definition 7.3 (Černý family of automata [Čer64]). *The* Černý family of automata *is defined as*

$$\mathcal{C}_n = \left\{ A = \begin{pmatrix} 0 & 1 & & \\ 0 & 1 & & \\ & & \ddots & \\ & & & 1 \end{pmatrix}, B = \begin{pmatrix} & 1 & & \\ & & \ddots & \\ & & & 1 \\ 1 & & & \end{pmatrix} \right\},$$

where the omitted elements are zeros. In particular, C_4 is the automaton of Example 7.1 and Figure 7.1.

The Černý family of automata will provide a lower bound on the Černý function, that we define now.

Definition 7.4 (Černý function). *The* Černý function $C(n)$ *is defined as the largest reset threshold among all automata with* n *states.*

The Černý family is the only known infinite series of automata with a reset threshold of $(n-1)^2$.

Theorem 7.2 (Lower Bound on $C(n)$ [Čer64]). *The Černý automaton* $\mathcal{C}_n$ *defined in Definition 7.3 has a reset threshold of* $(n-1)^2$.

Combining Theorems 7.2 and 7.1 yields

$$(n-1)^2 \leq C(n) \leq \frac{n^3-n}{6},$$

while Theorem 7.2 allows to state the Černý conjecture as

$$C(n) = (n-1)^2.$$

7.3 Deciding Synchronization

We have seen that synchronization is the existence of a word that brings all states to a common state. In fact, a necessary and sufficient condition for synchronization is that for any *pair* of states, there is a word that brings the two states of the pair to a common state. This was established in Černý's original article [Čer64, Theorem 2]. Eppstein [Epp90, Theorem 4] has proposed a polynomial time algorithm to decide synchronization that is based on this result. It uses a structure that we call the graph of pairs.

Definition 7.5 (Graph of pairs). *For a given automaton* S*, we call the* graph of pairs $\mathcal{F}(S)$ *the graph defined as follows.*

- *One node for each unordered pair of states with repetition (*$\frac{n(n+1)}{2}$ *nodes),*
- *an edge from pair* i_1, i_2 *to pair* j_1, j_2 *if there is a matrix* $A_k \in M$ *such that*

$$(A_k)_{i_1 j_1}, (A_k)_{i_2 j_2} > 0 \text{ (or } (A_k)_{i_1 j_2}, (A_k)_{i_2 j_1} > 0\text{)}, \tag{7.1}$$

  where i_1, i_2, j_1, j_2 *are not necessarily different.*

The graph of pairs is usually defined with one edge from pair i_1, i_2 to pair j_1, j_2 *for each* matrix $A_k \in M$ that satisfies (7.1) and called the automaton of pairs. We define it here as a graph without multiple edges. Removing these multiple edges does not change the reachability condition that we will define in Theorem 7.3. However, these multiple edges would increase the complexity of the decision procedure described in Section 8.6.1 of Chapter 8 and this is why we define the graph of pairs without these multiple edges.

Example 7.3. *The graph of pairs of the automaton of Figure 7.2 is represented on Figure 7.3.*

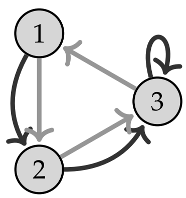


Figure 7.2: A simple automaton with three states and two letters (black and grey edges).

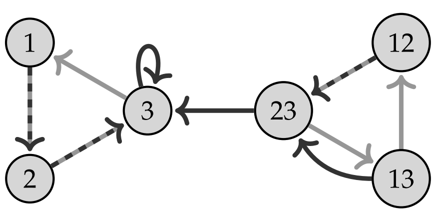


Figure 7.3: Graph of pairs for automaton of Figure 7.2.

The synchronization of an automaton is equivalent to a simple reachability property in its graph of pairs.

Theorem 7.3 (Criteria for synchronization [Epp90, Theorem 4])**.** *The automaton S is synchronizing if and only if, in its graph of pairs (as defined above), from each node representing a pair, there is a path to a node representing a single state.*

The graph of pairs has $\frac{n(n+1)}{2}$ nodes, where n is the number of states. Checking the criteria of Theorem 7.3 with a breadth-first search algorithm allows to verify whether an automaton is synchronizing with a complexity of $\mathcal{O}(n^2m)$ where m the number of letters.

7.4 Negative Complexity Results

We present negative complexity results regarding the computation and approximation of the reset threshold of an automaton. The first states that computing the reset threshold is an NP-hard problem.

Theorem 7.4 (Eppstein [Epp90, Theorem 8]). *The problem of deciding whether a given automaton has a reset threshold smaller than or equal to ℓ is NP-hard.*

In fact, stronger hardness results have been obtained. Even approximating the reset threshold is an NP-hard problem. For the problem of minimizing a function $f(x)$, a ρ-approximation algorithm is an algorithm that computes a solution x such that for any instance of the problem

$$\text{OPT} \leq f(x) \leq \rho\text{OPT},$$

where OPT is the optimal solution. The following result shows that, unless P=NP, there is no $n^{1-\epsilon}$-approximation algorithm that computes the reset threshold. Equivalently, unless P=NP, there is no algorithm that for any synchronizing automaton $\mathcal{A}$ with n states finds a synchronization word of length smaller than or equal to $n^{1-\epsilon}$ times the reset threshold of $\mathcal{A}$.

Theorem 7.5 ([GS15, Theorem 5.1]). *For every $\epsilon > 0$, it is NP-hard to approximate the reset threshold of automata of size $n \times n$ within a multiplicative factor $n^{1-\epsilon}$.*

Before Theorem 7.5, a reduction from the *SET-COVER problem* to the computation of the reset threshold was obtained by Gerbush and Heeringa in [GH11]. We will generalize it in the next chapter to prove the hardness of approximating the length of the shortest SIA product of a set.

Definition 7.6 (SET-COVER problem). *Given a set of elements $U = \{1, 2, \dots, n\}$ and a collection $\mathcal{F}$ of subsets of U whose union is equal to U, i.e., $U = \cup_{T \in \mathcal{F}} T$. The SET-COVER problem is to compute a sub-collection $\mathcal{F}' \subseteq \mathcal{F}$ of the smallest possible size, whose union still equals to U.*

Theorem 7.6 ([DS14, Corollary 1.5]). *For every $0 < \alpha < 1$, it is NP-hard to approximate SET-COVER to within $(1 - \alpha)\log_2(n)$.*

Unless P=NP, this result is tight in the sense that there is a polynomial time $\log_2(n)$-approximation algorithm for the SET-COVER problem.

We can see that the approximation ratio of Theorem 7.6 is smaller than that of Theorem 7.5, so that the reduction from the SET-COVER problems yields a weaker result than Theorem 7.5.

7.5 Conclusion

We have presented classical results from automata theory that we will use in the next chapter. We will see that these results generalize easily to sets of stochastic matrices. This will allow us to answer some questions such as bounding the length of the smallest SIA product of a set of stochastic matrices or deciding whether a set of stochastic matrices is SIA.

Chapter 8

Almost Contractive Sets of Matrices

We study the shortest SIA[1] products of sets of matrices. We observe that the shortest SIA product of a set of matrices is usually very short and we provide a first upper bound on the length of the shortest SIA product (if one exists) of any set of stochastic matrices.

When particularized to automata, the problem becomes that of finding periodic synchronizing words, and we develop the consequences of our results in relation with the celebrated Černý conjecture in automata theory.

We also study the related notions of positive-column, Sarymsakov, and scrambling matrices.

This chapter presents the results of [CGHJ17] and [CHJ15b] and is organized as follows.

- In Section 8.1, we recall the definition of almost contractive sets and we motivate the study of shortest SIA and positive-column products of these sets.
- In Section 8.2, we define several *indices*: to a set of stochastic matrices S, we associate a quantity called the *SIA index of S* that is equal to the length of the shortest SIA product of matrices from S. We define similar indices characterizing the shortest positive-column, Sarymsakov, or scrambling product of a set. We also define quantities equal to the largest SIA index, largest positive-column index etc. among all sets of almost contractive sets.

[1] We recall that a matrix A is SIA if it is stochastic and $\lim_{t\to\infty} A^t$ exists and is a rank-one matrix.

- Section 8.3 contains the main theoretical contributions of this chapter. We study the relation between the indices. We also show that, for any dimension n the largest SIA index among sets of $n \times n$ matrices can always be obtained with a set of *automaton* matrices. We prove the same result for the positive-column index, which also establishes a link with the Černý conjecture, because the positive-column index of a set of automaton matrices is equal to the reset threshold of this set.

- In Section 8.4, we give lower and upper bounds on the largest positive-column index.

- Section 8.5 is devoted to the properties of the SIA index. Lower and upper bounds on the largest SIA index are provided, and we present an experiment that computes the largest SIA index for small values of n.

- In Section 8.6 we discuss the procedure to decide whether a given set of matrices is almost contractive and the hardness of computing and approximating the SIA index and the other indices.

8.1 Introduction

In this chapter, we study almost contractive sets of matrices. We recall that a stochastic matrix is SIA if the limit

$$\lim_{t\to+\infty} P^t$$

exists and all of its rows are equal, i.e., the limit is a rank-one matrix.

Definition 8.1 (Almost contractive set)**.** *We say that a set of stochastic matrices S is* almost contractive *if there is a product of matrices from S with repetitions allowed that is an SIA matrix.*

Almost contractive sets are the sets that have a left-infinite product that converge to a matrix with all its rows equal, as indicated en the following proposition.

Proposition 8.1. *A set $S = \{A_1, \dots, A_m\}$ of stochastic matrices is almost contractive if and only if there is a infinite sequence of indices $\sigma \in \Sigma$ such that*

$$\lim_{t\to\infty} A_{\sigma(t)} \dots A_{\sigma(1)} A_{\sigma(0)} = \mathbf{1} y^\top$$

for some $y \in \mathbb{R}^n$ satisfying $y^\top \mathbf{1} = 1$.

Proof. The "only if" part is evident, as repeating infinitely the SIA product yields the desired infinite product (and the condition that $y^\top \mathbf{1} = 1$ simply follows from the stochasticity of the matrices in the product).
If: since the limit $\lim_{t\to\infty} A_{\sigma(t)} \dots A_{\sigma(1)} A_{\sigma(0)} = \mathbf{1} y^\top$ has a positive column, there is a finite t such that $P \triangleq A_{\sigma(t)} \dots A_{\sigma(1)} A_{\sigma(0)}$ has a positive column and this product P is SIA (by Proposition 2.3). ■

For an almost contractive set S, we denote by $\mathrm{sia}(S)$, $\mathrm{sar}(S)$, $\mathrm{scr}(S)$ and $\mathrm{pc}(S)$ the lengths of the shortest SIA, Sarymsakov (Definition 2.3), scrambling and positive-column products of matrices from S[2], and we call these quantities the *SIA index*, the *Sarymsakov index*, the *scrambling index* and the *positive-column index* of the set S. The positive-column index has been studied extensively in the context of automata theory [Vol08]. Indeed, for sets of automaton matrices, the positive-column index is equal to the reset threshold of the automaton represented by this set. The scrambling index has been studied in the case of a single matrix [AK09].

[2]We will see that these products always exist.

These indices are similar and related to the classical and well studied notion of *exponent* – the length of the shortest product that is entrywise positive, if one exists (see [BR91, Section 3.5] for a survey of the single matrix case, and [GGJ18, PV12] for more recent work on matrix sets).

In the context of a system whose switching sequence can be controlled, an SIA product corresponds to a switching sequence that can be repeated to make the system converge to a rank-one matrix. Thus, shorter SIA products correspond to simpler controllers. Furthermore, the length of the SIA product has an influence on the converging rate. Indeed, if $P = A_\ell \dots A_2 A_1$ is an SIA product, the sequence $\dots A_\ell \dots A_2 A_1 A_\ell \dots A_2 A_1$ converges to a rank-one matrix at an average rate of $\lambda_2^{1/\ell}$, where λ_2 is the second largest eigenvalue of P.

Finally, the study of these indices brings new insights to synchronizing automata and the *Černý conjecture* [Čer64] that we presented in the previous chapter and that states that for any set of automaton $n \times n$ matrices S either there is a product of length at most $(n-1)^2$ of matrices from S having a positive column or there is no such product at all. The conjecture has been open for half a century and the best bound obtained so far is cubic in n [Vol08]. As we will soon see, a good upper bound on the SIA index of $n \times n$ matrices would improve the state of the art on the Černý conjecture. In particular, any subquadratic bound would bring a breakthrough. Additionally, we will see that the Černý conjecture can be generalized to sets of stochastic matrices, in the sense that it holds for sets of stochastic matrices if and only if it holds in its classical formulation, for automata.

8.2 Indices and Bounds

In the present section we associate a natural combinatorial parameter with each of the classes of SIA, Sarymsakov, scrambling and positive-column matrices. We study the relation between these indices and we provide bounds on these indices that depend only on the size of the matrices.

Definition 8.2 ($\mathcal{X}$ index)**.** *Let $\mathcal{X}$ be a class of matrices. The* $\mathcal{X}$ index *of a set of matrices S is the smallest ℓ such that there is a product of length ℓ of matrices from S belonging to $\mathcal{X}$.*

If we say that set belongs to a class $\mathcal{X}$ if there is a product of the set that belongs to $\mathcal{X}$, then the $\mathcal{X}$ index can be seen as the length of the shortest witness proving that a set does belong to the class $\mathcal{X}$. If no product of S belongs to $\mathcal{X}$, we will agree that the $\mathcal{X}$ index is undefined. In this section, we will focus on SIA, Sarymsakov, scrambling and positive-column indices of almost contractive sets and we will denote the corresponding index of a set S by $\mathrm{sia}(S)$, $\mathrm{sar}(S)$, $\mathrm{scr}(S)$, $\mathrm{pc}(S)$. We note that our terminology is in agreement with the case of a single matrix, where the scrambling index has received a lot of attention [AK09]. We would also like to remark the positive-column index has been extensively studied for sets of automaton matrices, since it is equal to the reset threshold of the corresponding automaton.

An almost contractive set has well-defined Sarymsakov, scrambling and positive-column indices, as shown in the next proposition.

Proposition 8.2. *For any set of stochastic matrices S the four statements are equivalent.*

*(i) There exists an **SIA** product of matrices from S*

*(ii) there exists a **Sarymsakov** product of matrices from S*

*(iii) there exists a **scrambling** product of matrices from S*

*(iv) there exists a **positive-column** product of matrices from S.*

Proof. By Proposition 2.3, for any SIA product P, there is a power p such that P^p has a positive column. Hence $(i) \Rightarrow (iv)$. Additionally, the inclusions of Proposition 2.4

$$S_{\text{PC}} \subset S_{\text{SCR}} \subset S_{\text{SAR}} \subset S_{\text{SIA}}.$$

imply

$$(iv) \Rightarrow (iii) \Rightarrow (ii) \Rightarrow (i).$$

∎

An almost contractive set is therefore a set of stochastic matrices that has an SIA product, a Sarymsakov, a scrambling and a positive-column product.

One of the basic questions arising in regard with these indices is how large the index of a set of $n \times n$ stochastic matrices can be? Such questions have received a lot of attention for the exponent [BR91, Section 3.5] that can be seen as the $\mathcal{X}$ index, where $\mathcal{X}$ is the class of positive matrices. Likewise, the reset threshold of automata – that can be seen as the positive-column index in the case of automaton matrices – has been studied extensively [Vol08].

We will study the quantities $\mathrm{sia}(n)$, $\mathrm{sar}(n)$, $\mathrm{scr}(n)$ and $\mathrm{pc}(n)$, that we define as the largest SIA, Sarymsakov, scrambling and positive-column indices among all sets of $n \times n$ stochastic matrices. We also study $\mathrm{sia}_A(n)$ that we define as the largest SIA index among all sets of $n \times n$ *automaton matrices*.

8.3 Relation between Indices

In this section, we study the relation the quantities $\mathrm{sia}(n)$, $\mathrm{sar}(n)$, $\mathrm{scr}(n)$, $\mathrm{pc}(n)$, as well as $\mathrm{sia}_A(n)$ and $C(n)$. We recall that $C(n)$ is defined as the largest reset threshold among all automaton with n states, or equivalently, as the largest positive-column index among sets of automaton matrices. We show that

$$\mathrm{pc}(n) = C(n), \tag{8.1}$$

and

$$\mathrm{sia}(n) = \mathrm{sia}_A(n) \tag{8.2}$$

meaning that the largest SIA and positive-column indices can be reached by sets of *automaton* matrices. We then establish the following relations between the indices.

$$\mathrm{sia}(n) \leq \mathrm{sar}(n) = \mathrm{scr}(n) = \mathrm{pc}(n) \leq (n-1)\,\mathrm{sia}(n).$$

We start by showing that, in the case of automaton matrices, the notions of Sarymsakov, scrambling and positive-column coincide.

Proposition 8.3. *Let P be an automaton matrix. The three properties are equivalent:*

(i) P is Sarymsakov

(ii) P is scrambling

(iii) P is positive-column.

Proof. The inclusions $S_{\mathrm{PC}} \subset S_{\mathrm{SCR}} \subset S_{\mathrm{SAR}} \subset S_{\mathrm{SIA}}$ of Proposition 2.4 imply (iii) $\Rightarrow$ (ii) and (ii) $\Rightarrow$ (i), so that (i) $\Rightarrow$ (iii) remains to be proved. Let P be Sarymsakov. Recall that a stochastic matrix is Sarymsakov if for any two disjoint nonempty subsets S and S' the consequent function F satisfies either

$$F(S) \cap F(S') \neq \emptyset \tag{8.3}$$

or

$$|F(S) \cup F(S')| > |S \cup S'|. \tag{8.4}$$

Let $i, j \in \{1, \ldots, n\}$. Condition (8.4) cannot be satisfied for sets $S = \{i\}$ and $S' = \{j\}$ because

$$|F_P(\{i\}) \cup F_P(\{j\})| > |\{i\} \cup \{j\}| = 2$$

would imply

$$|F_P(\{i\})| > 1 \text{ or } |F_P(\{j\})| > 1,$$

which means that row i or row j has more than one positive element, which is impossible by the definition of an automaton matrix. Therefore, Condition (8.3) is satisfied for any pair of singletons $S = \{i\}, S' = \{j\}$, meaning that for any rows i, j, the positive element in row i is in the same column as the positive element in row j, so that the matrix P is in fact positive-column. ■

Corollary 8.1. *For any set of automaton matrices S,*

$$\mathrm{sar}(S) = \mathrm{scr}(S) = \mathrm{pc}(S).$$

In order to prove Equation (8.1), the equality between $\mathrm{pc}(n)$ and the largest reset threshold of n states automata, we define an associated automaton to any set of stochastic matrices, a construction that has been developped in [BJO15] for the study of primitive matrix sets.

Definition 8.3 (Pattern domination)**.** *Let A and B be nonnegative matrices. We write $A \succeq B$ and say that matrix A* dominates *matrix B if*

$$\exists a \in \mathbb{R} \text{ such that } aA \geq B,$$

where $\geq$ is an elementwise inequality. This corresponds to the matrix A having a positive element at each position at which the matrix B has a positive element.

Definition 8.4 (Automaton associated with a set)**.** *Let $S = \{A_1, \ldots, A_m\}$ be a set of stochastic matrices. We call the* automaton associated with the set *S, the automaton S' containing all automaton matrices that are dominated by some matrix of S, that is*

$$S' \triangleq \{A' \mid A' \in S_{AUTO} \text{ and } \exists A \in S \text{ s. t. } A \succeq A' \},$$

with S_{AUTO} the set of automaton matrices.

Note that the associated automaton can contain a very large number of matrices but we will only use it in proofs and we will not construct it explicitly in any algorithm.

Example 8.1. *The automaton associated to set*

$$S = \left\{ A_1 = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0.8 & 0.2 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}, A_2 = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \end{pmatrix} \right\}$$

is

$$S' = \left\{ A_1' = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix},\ A_2' = A_2,\ A_3' = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \right\}$$

We now define a way to represent products or words on a graph. This construction is again inspired by a similar construction in [BJO15].

Definition 8.5 (Graph associated with a word)**.** *Given a set $S = \{A_1, \ldots, A_m\}$ of stochastic matrices and a word $w = w_\ell \ldots w_1$ on the alphabet $\{1, \ldots, m\}$, we call* graph associated with the word w *the graph whose adjacency matrix is*

$$\begin{pmatrix} 0 & A_{w_\ell} & & & \\ & \ddots & \ddots & & \\ & & \ddots & A_{w_2} & \\ & & & \ddots & A_{w_1} \\ & & & & 0 \end{pmatrix}.$$

Example 8.2. *For the set S of Example 8.1, the graph associated with the word 11221 is depicted in Figure 8.1.*

In Lemma 8.1, we will look for an *in-tree* in the graph associated with a word because the in-tree corresponds to a synchronizing word of the associated automaton.

Definition 8.6 (In-tree and spanning in-tree)**.** *We call an* in-tree *a directed graph in which, for a vertex r called the* root *and any other vertex v, there is exactly one directed path from r to v. For a digraph $\mathcal{G}$, we call a* spanning in-tree *an in-tree that has the same set of nodes as $\mathcal{G}$ and whose set of edges is a subset of that of $\mathcal{G}$.*

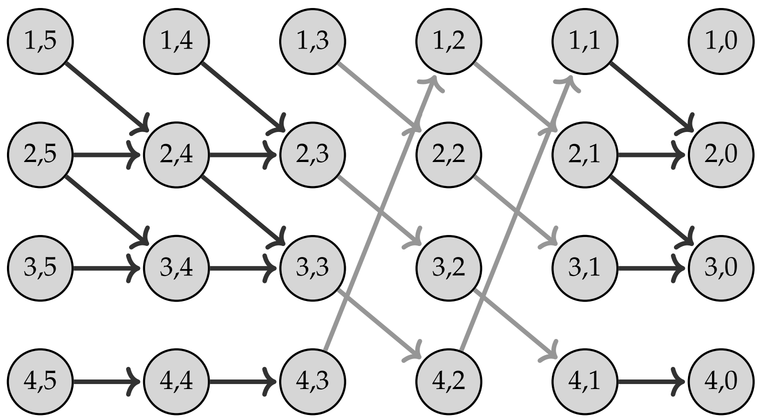


Figure 8.1: graph associated with the word 11221 for S defined in Example 8.1. The edges corresponding to matrix A_1 are in black, those corresponding to matrix A_2 are in grey.

The next lemma relates positive-column products of a set S of stochastic matrices and synchronizing products of its associated automaton S. This lemma means in particular that a set S is almost contractive if and only if its associated automaton is synchronizing and additionally that they have column-positive products of the same lengths.

Lemma 8.1. *Let $S = \{A_1, \dots, A_m\}$ be a set of stochastic matrices, and let $S' = \{A'_1, \dots, A'_{m'}\}$ be the associated automaton of S, as defined in Definition 8.4. A word $w = w_\ell \dots w_1$ over the alphabet $\{1, \dots, m\}$ is positive-column for S if and only if there is a word $w' = w'_\ell \dots w'_1$ over the alphabet $\{1, \dots, m'\}$ that is synchronizing for S' and such that*

$$\forall i \in \{1, \dots, \ell\},\ A_{w_i} \succeq A'_{w'_i}.$$

Proof. If: The product $A_w = A_{w_\ell} \dots A_{w_1}$ dominates $A'_{w'} = A'_{w'_\ell} \dots A'_{w'_1}$ because each A_{w_i} dominates $A'_{w'_i}$ and domination is preserved under multiplication. In particular, if $A'_{w'}$ has a positive column, the same column is positive in A_w.

Only if: We call the node (i, j) with $i \in \{1, \dots, n\},\ j \in \{0, \dots, \ell\}$ the node corresponding to the i^{th} row of the $(\ell - j)^{\text{th}}$ block-row of the matrix of Definition 8.5. This numbering is represented in Figure 8.1. Suppose that A_w has a positive k^{th} column. Therefore, in the graph $\mathcal{G}(w)$ associated with the word w, from each node (i, ℓ), there exists a path to node $(k, 0)$. The graph of these paths has a spanning in-tree rooted in k (an example is given in Figure 8.2). In $\mathcal{G}(w)$, for each node, there is at most one outgoing edge that belongs to the spanning in-tree. Therefore, some edges of $\mathcal{G}(w)$ can be removed such that the graph still has the same spanning in-tree and each node has exactly one outgoing edge. We perform the corresponding operations on the matrices that form the product $A_{w_\ell} \dots A_{w_1}$, that is, we set to zero positive elements that do not correspond to edges of the spanning in-tree and such that on each row of each matrix, exactly one element remains positive. Then, we set to 1 all remaining positive elements. We obtain a new product $A'_{w'_\ell} \dots A'_{w'_1}$ for which

- the k^{th} column is positive
- $\forall i,\ A_{w_i} \succeq A'_{w'_i}$
- $\forall i,\ A'_{w'_i}$ is an automaton matrix by construction,

from which we conclude that each $A'_{w'_i}$ belongs to S' the automaton associated with the set S and that w' is synchronizing for S'. ∎

Example 8.3. *The graph associated with the word 11221 is represented in Figure 8.2. The in-tree is in black. We see that removing the dashed edges allows keeping the in-tree and having exactly one outgoing edge from each node. Without these dashed*

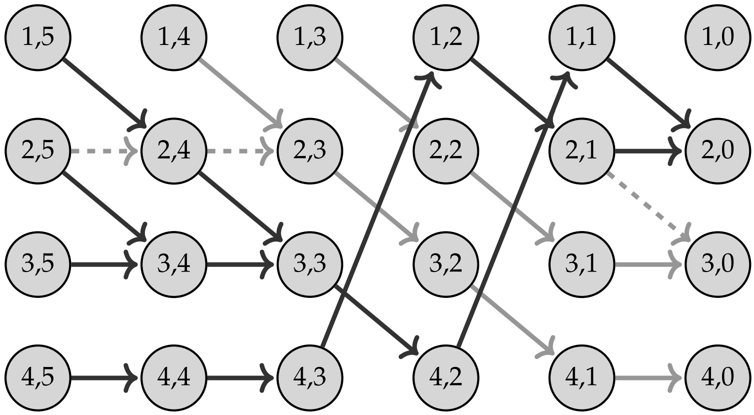


Figure 8.2: Graph associated with the word 11221. In black: the in-tree. Without the dashed edges, the graph is that of the word 11223 of automaton S'.

edges, the graph becomes that associated with the word 11223 of automaton $M = \{A'_1, A'_2, A'_3\}$.

The next two results highlight the importance of sets of automaton matrices in the study of $\mathrm{sia}(n)$ and $\mathrm{pc}(n)$, the largest SIA and positive-column indices.

Theorem 8.1. *For any dimension $n \in \mathbb{N}$, the largest positive-column index among all sets of $n \times n$* stochastic *matrices is equal to the largest positive-column index among all sets of $n \times n$* automaton *matrices:*

$$\mathrm{pc}(n) = C(n).$$

Proof. Since the set of sets of automaton matrices is a subset of the set of sets of stochastic matrices and since $C(n)$ and $\mathrm{pc}(n)$ can be seen as optimization problems on these sets, we have that $\mathrm{pc}(n) \geq C(n)$.

Now, let S be a set of $n \times n$ stochastic matrices with $\mathrm{pc}(S) = \mathrm{pc}(n)$ (such a set exists by definition of $\mathrm{pc}(n)$). By Lemma 8.1, its associated automaton S' satisfies

$$\mathrm{pc}(S') = \mathrm{pc}(S) = \mathrm{pc}(n).$$

Moreover, we have, by definition of $C(n)$,

$$\mathrm{pc}(S') \leq C(n).$$

Hence

$$\mathrm{pc}(n) = \mathrm{pc}(S') \leq C(n)$$

and

$$\mathrm{pc}(n) = C(n).$$

■

Theorem 8.2. *For any dimension $n \in \mathbb{N}$, the largest SIA index among all sets of $n \times n$* stochastic *matrices is equal to the largest SIA index among all sets of $n \times n$* automaton *matrices:*

$$\mathrm{sia}(n) = \mathrm{sia}_A(n).$$

Proof. The proof is similar to that of Theorem 8.1. We have that $\mathrm{sia}(n) \geq \mathrm{sia}_A(n)$ because automaton matrices are stochastic.

Now, let S be a set of $n \times n$ stochastic matrices with $\mathrm{sia}(S) = \mathrm{sia}(n)$ (such a set exists by definition of $\mathrm{sia}(n)$). By Lemma 8.1, its associated automaton S' is synchronizing and therefore also almost contractive. We observe that any SIA product of S' has a length larger than or equal to $\mathrm{sia}(S)$. Indeed, since the matrices of S dominate those of S', we have that if $P = A'_1 \dots A'_\ell$ is an SIA product of length ℓ of matrices from S', then replacing each A'_i by a matrix A_i of S satisfying $A_i \succeq A'_i$ yields an SIA product of length ℓ of matrices from S. Hence,

$$\mathrm{sia}_A(n) \geq \mathrm{sia}(S') \geq \mathrm{sia}(S) = \mathrm{sia}(n)$$

and the proof is complete. ■

We will now prove the main result of this section. First, observe that for every almost contractive set S, we have the following inequalities as a direct consequence of Proposition 2.4:

$$\operatorname{sia}(S) \leq \operatorname{sar}(S) \leq \operatorname{scr}(S) \leq \operatorname{pc}(S). \tag{8.5}$$

Theorem 8.3. *Let* $\operatorname{sia}(n)$, $\operatorname{sar}(n)$, $\operatorname{scr}(n)$, $\operatorname{pc}(n)$ *be the largest values of the corresponding indices among all almost contractive sets of* $n \times n$ *matrices. For any dimension* $n \in \mathbb{N}$*, we have*

$$\operatorname{sia}(n) \leq \operatorname{sar}(n) = \operatorname{scr}(n) = \operatorname{pc}(n) \leq (n-1)\operatorname{sia}(n). \tag{8.6}$$

Proof. First, notice that we have the following inequalities as a direct consequence of (8.5):

$$\operatorname{sia}(n) \leq \operatorname{sar}(n) \leq \operatorname{scr}(n) \leq \operatorname{pc}(n).$$

It remains to prove that $\operatorname{pc}(n) \leq \operatorname{sar}(n)$ and $\operatorname{pc}(n) \leq (n-1)\operatorname{sia}(n)$. We first prove $\operatorname{pc}(n) \leq \operatorname{sar}(n)$. Let $\operatorname{sar}_A(n)$ be the largest Sarymsakov index among all $n \times n$ sets of automaton matrices.

$$\begin{aligned} \operatorname{pc}(n) &= C(n) \qquad (8.7)\\ &= \operatorname{sar}_A(n) \qquad (8.8)\\ &\leq \operatorname{sar}(n). \qquad (8.9) \end{aligned}$$

The first equality (8.7) is Theorem 8.1. The second (8.8) is a consequence of the fact that an automaton matrix is Sarymsakov if and only if it is positive-column (see Proposition 8.3). Inequality (8.9) holds because automaton matrices form by definition a subset of stochastic matrices.

We now prove $\operatorname{pc}(n) \leq (n-1)\operatorname{sia}(n)$. We have

$$\begin{aligned} \operatorname{pc}(n) &= C(n) \qquad (8.10)\\ &\leq (n-1)\operatorname{sia}_A(n) \qquad (8.11)\\ &= (n-1)\operatorname{sia}(n). \qquad (8.12) \end{aligned}$$

The first equality (8.10) is Theorem 8.1, the second (8.11) is due to Corollary 2.3, that states that the $n-1$st power of an SIA automaton matrix has a positive column and the third (8.12) is Theorem 8.2. ■

8.4 Bounds on the Positive-Column Index

In this short section, we discuss bounds on $\mathrm{pc}(n)$. Theorem 8.1 that establishes the equality

$$\mathrm{pc}(n) = C(n)$$

allows us to use known results from automata theory and to apply them to $\mathrm{pc}(n)$. The best known upper bound on the reset threshold of automata of size n is $\frac{n^3-n}{6}$ (Theorem 7.1) and it translates immediately to a bound on $\mathrm{pc}(n)$. Furthermore, any new bound on $C(n)$, would provide a new bound on $\mathrm{pc}(n)$.

This applies to lower bounds on $C(n)$ as well since these $C(n) = \mathrm{pc}(n)$ are equal. In particular, $(n-1)^2$ is the best known lower on $C(n)$ and this translates to $(n-1)^2 \leq \mathrm{pc}(n)$.

Finally, Theorem 8.1 also means that any new bound on $\mathrm{pc}(n)$ would provide a new bound on $C(n)$.

Theorem 8.4. *Let S be an almost contractive set. The positive-column index of S is at most $\frac{n^3-n}{6}$ and at most $(n-1)^2$ if Conjecture 7.1 holds.*

Proof. This a consequence of Corollary 8.1, Conjecture 7.1 and Theorem 7.1. ■

8.5 Bounds on the SIA index

In the previous section, we have given lower and upper bounds on $\mathrm{pc}(n)$. In Theorem 8.3, we have seen that $\mathrm{sar}(n) = \mathrm{scr}(n) = \mathrm{pc}(n)$ so these bounds also apply to $\mathrm{sar}(n)$ and $\mathrm{scr}(n)$.

We now study the SIA index and in particular $\mathrm{sia}(n)$, the largest SIA index among all sets of $n \times n$ stochastic matrices. To the best of our knowledge, the SIA index is a new notion. The study of $\mathrm{sia}(n)$ has potential consequences for the Černý conjecture. Indeed, we have seen in Theorem 8.3 that $C(n) \leq \mathrm{sia}(n)(n-1)$ and a good upper bound on $\mathrm{sia}(n)$ could therefore improve the state of the art of the conjecture.

We provide general upper and lower bounds on $\mathrm{sia}(n)$ that depend only on n. We conjecture that $\mathrm{sia}(n)$ is bounded by $2n$ and we describe the results of our computational experiments to estimate $\mathrm{sia}(n)$ for small values of n.

8.5.1 Upper bounds

Thanks to Theorems 8.3 and 8.1, any upper bound on $C(n)$ translates into a bound on $\mathrm{sia}(n)$. In particular, Theorem 7.1 provides the upper bound

$$\mathrm{sia}(n) \leq \frac{n^3 - n}{6}.$$

In fact, we believe that $\mathrm{sia}(n)$ is much smaller and we propose the following conjecture.

Conjecture 8.1. *The SIA index of a set of $n \times n$ stochastic matrices is bounded by $2n$.*

In the next subsections we will support this conjecture by providing results of computational experiments and analysis of sets that are extremal for the Černý conjecture. We believe that Conjecture 8.1 offers a new angle on the Černý conjecture and can bring new insights. First, $\mathrm{sia}(n)(n-1)$ is an upper bound on $C(n)$ (Theorem 8.3) and therefore, if Conjecture 8.1 is true, then any synchronizing automaton of size n has a reset threshold at most $2n(n-1)$, which is a significant improvement of the state of the art. Second, the SIA index tends to be surprisingly small for automata with large reset thresholds highlighting the structural properties of these particular cases: Rystsov's automata [Rys97, p. 279], Černý automata [Vol08], and other slowly synchronizing automata [AVG13] have small SIA-indices.

8.5.2 Numerical results

We now present the results of our computational experiments that support our Conjecture 8.1. Since the bound on SIA index for automaton matrices is equal to the bound on the SIA index for stochastic matrices (Proposition 8.2), we only investigate automaton matrices. We have computed on a computer cluster the SIA index of all automata made of two matrices up to $n = 7$, and up to $n = 9$ for *initially connected automata*, a notion that we will define soon. The results are summarized in Table 8.1. We have done the same for all triplets of automaton matrices up to $n = 5$ (Table 8.1), and we obtain exactly the same maximum SIA-indices.

The maximum SIA index grows approximately like $2n$, as shown in Figure 8.3. We were not able to find a pattern in the sequence $0, 1, 3, 5, 8, 10, 13, 15, 16$ of largest SIA indices nor could we relate it to a well-known integer sequence.

n	maximum SIA index, two matrices, all automata	maximum SIA index, two matrices, IC automata.	maximum SIA index, three matrices, IC automata
1	0	0	0
2	1	1	1
3	3	3	3
4	5	5	5
5	8	8	8
6	10	10	
7	13	13	
8		15	
9		16	

Table 8.1: Exhaustive tests for pairs and triplets of automaton matrices. The first column is the size of the matrices, the second column is the maximum SIA index of two-matrices automata, the third column is the maximum SIA index of two-matrices initially connected automata (see Definition 8.8), and the fourth column is the maximum SIA index of three-matrices automata.

Examples of sets of 8×8 and 9×9 matrices that have an SIA index of 15 and 16 are depicted on Figure 8.4.

Our methodology is the following. For the exhaustive tests, we have enumerated all sets of automaton matrices and for each set we have computed its SIA index. In order to compute the SIA index, we have enumerated all matrix products corresponding to Lyndon words (Definition 8.7 below) of increasing length, until an SIA product is found. Proposition 8.4 guarantees that the correct SIA index is computed.

To compute the SIA index of a set, it is not necessary to compute all products up to a given length. We observe that for any stochastic matrices (or products of matrices) A_1 and A_2:

$$A_1 A_2 \text{ is SIA } \Leftrightarrow A_2 A_1 \text{ is SIA} \tag{8.13}$$

and that

$$\forall p \geq 2,\ A_1^p \text{ is SIA } \Leftrightarrow A_1 \text{ is SIA} . \tag{8.14}$$

For a set $S = \{A_1, A_2, \ldots, A_k\}$, we can define an arbitrary ordering $A_1 \prec A_2 \prec \ldots \prec A_k$. Clearly, the products of matrices from S are in one-to-one correspondence with *words* – the sequences of the symbols $A_1, A_2, \ldots, A_k$. The

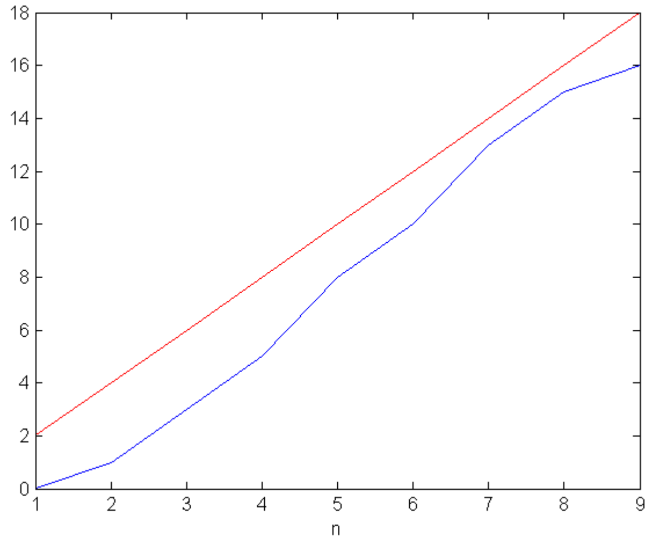


Figure 8.3: Maximum SIA index for pairs of matrices (blue) and the curve $y = 2n$ (red).

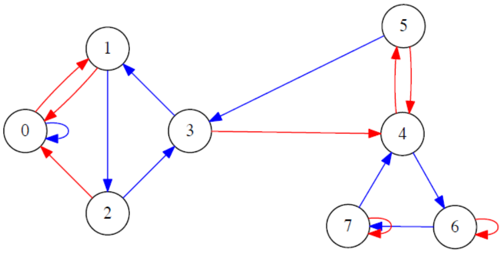
lexicographic order on the words is: $P \prec Q$ if either $Q = PU$ for some word U; or $P = UA_iV$ and $Q = UA_jW$ for $A_i \prec A_j$ and some words U, V, W.

Definition 8.7 (Lyndon word [Lyn54])**.** *A* cyclic shift *of a word P is a word of the form VU if $P = UV$. A non-empty word P is Lyndon if it is strictly smaller in the lexicographic order than all of its cyclic shifts.*

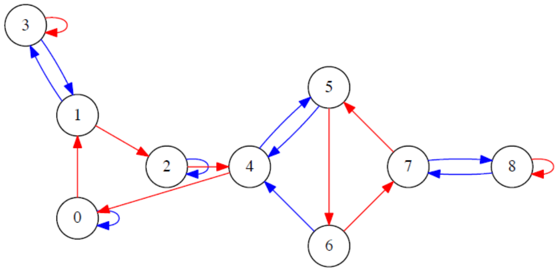
Proposition 8.4. *Let S be an almost contractive set of matrices such that* $\mathrm{sia}(S) = \ell$*. There is a Lyndon word of length ℓ such that the corresponding product of matrices from S is SIA.*

Proof. Properties (8.13) and (8.14) guarantee that there is always a shortest SIA product P that is aperiodic and that is not larger than any of its cyclic permutation. We can then invoke the classical result about the Lyndon words stating that an aperiodic word that is not larger than any of its cyclic shifts is actually Lyndon [Lot97, Proposition 5.1.2]. ■

We have noticed in our tests that all extremal examples are *initially connected* (in fact, even strongly connected).



(a) Pair of 8×8 matrices that has an SIA index of 15. This is the only such pair up to relabelling of the nodes.



(b) Pair of 9×9 matrices that has an SIA index of 16. There are 12 different (up to relabelling of the nodes) sets that have an SIA index of 16.

Figure 8.4: Examples of matrix sets that have extremal SIA index. The matrices are represented by their graphs, the blue edges represent the positive elements of one matrix and the red edges the other matrix.

Definition 8.8 (Initially connected). *A set of automaton matrices $\{P_1, \ldots, P_m\}$ is called* initially connected *or* IC *if in the graph associated with the matrix $P_1 + \cdots + P_m$ there exists a node q such that there is a path from q to any node of the graph. In particular, if the graph associated to $P_1 + \cdots + P_m$ is strongly connected, the set is initially connected.*

Therefore, we have decided to analyze larger values of n restricted to the case of initially connected automaton matrices. This has allowed us to perform the tests up to $n = 9$ instead of $n = 7$.

8.5.3 Lower bounds

Corollary 2.3 implies that automaton matrices with a large reset threshold have a large SIA index as well. Therefore, we focused on automata that are known to be tight for the Černý conjecture. The results are summarized in Table 8.2. A list of automata that are known to be tight for the Černý conjecture can be found in [Vol08].

Automaton	SIA index
Černý family (with $n \geq 3$)	n
3 states (3 different automata)	3
4 states (3 automata)	5
5 states	7
6 states (Kari automaton)	9

Table 8.2: SIA-indices of automata that are known to be tight for the Černý conjecture.

In the next proposition, we establish a lower bound on $\mathrm{sia}(n)$. We do so by proving that the Černý family of automata (see Definition 7.3) has an SIA index of n. Recall that

$$\mathcal{C}_n = \left\{ A = \begin{pmatrix} 0 & 1 & & \\ 0 & 1 & & \\ & & \ddots & \\ & & & 1 \end{pmatrix}, B = \begin{pmatrix} & 1 & & \\ & & \ddots & \\ & & & 1 \\ 1 & & & \end{pmatrix} \right\}. \tag{8.15}$$

We will see that its shortest SIA product is equal to AB^{n-1}. For small values of n, the values computed in Table 8.1 provide slightly better lower bounds.

Proposition 8.5. *The Černý set $\mathcal{C}_n$ of matrices of dimension n has an SIA index of* $\mathrm{sia}(\mathcal{C}_n) = n$.

Proof. Observe first that $(AB^{n-1})^{n-2}A$ (with A and B as in Equation 8.15) has a positive column because the corresponding word $(ab^{n-1})^{n-2}a$ is synchronizing [Čer64]. Thus $(AB^{n-1})^{n-1}$ also has a positive column and AB^{n-1} is SIA. Thus, $\mathrm{sia}(\mathcal{C}_n) \leq n$.

Furthermore, $\mathrm{sia}(\mathcal{C}_n) \geq \frac{\mathrm{pc}(\mathcal{C}_n)}{n-1}$ (Corollary 2.3) and $\mathrm{pc}(\mathcal{C}_n) = (n-1)^2$ (i.e., the reset threshold of the Černý automaton is $(n-1)^2$, see e.g. [Gus13]). Therefore, $\mathrm{sia}(\mathcal{C}_n) \geq (n-1)$.

It remains to show now that the case $\mathrm{sia}(\mathcal{C}_n) = (n-1)$ is impossible. Assume to the contrary that P is an SIA product of length $n-1$. Therefore, P^{n-1} has a positive column (by Corollary 2.3). By [Gus13, Proposition 4] every positive-column product of $\mathcal{C}_n$ has at least of $n^2 - 3n + 2$ occurrences of A and $n-1$ occurrences of B, thus, P has exactly one occurrence of A and $n-2$ occurrences of B. Applying Proposition 8.4 we further conclude that the word $A^{n-2}B$ corresponds to an SIA product as well, which is not the case.

■

Now we will analyze a set of matrices derived from the Wielandt series of matrices that have the largest possible exponent among $n \times n$ matrices [BR91, Chapter 3.5]. Matrix sets of this kind often appear in the study of combinatorial characteristics of matrix sets, e.g. generalizations of the exponents [SS03] or in the study of positive-column indices [AVG13]. We define the Wielandt set of automaton matrices as

$$\mathcal{W}_n = \left\{ A = \begin{pmatrix} & 1 & & & \\ & & 1 & & \\ & & & \ddots & \\ & & & & 1 \\ 0 & 1 & & & \end{pmatrix}, B = \begin{pmatrix} & 1 & & & \\ & & 1 & & \\ & & & \ddots & \\ & & & & 1 \\ 1 & 0 & & & \end{pmatrix} \right\},$$

where the omitted elements are zeros.

Proposition 8.6. *The Wielandt set of matrices $\mathcal{W}_n$ has an SIA index of* $\mathrm{sia}(\mathcal{W}_n) = n-1$.

Proof. Recall that the shortest positive-column product of $\mathcal{W}_n$ has length $n - 3n + 3$ [AVG13, Theorem 2]. Since $(n-1)\,\mathrm{sia}(\mathcal{W}_n)$ cannot be strictly smaller than $n - 3n + 3$ by Corollary 2.3, we immediately conclude that $\mathrm{sia}(\mathcal{W}_n) \geq (n-1)$. This bound is tight, since AB^{n-2} is the desired SIA product. ■

8.6 Complexity Results

In the present section we address algorithmic problems related to almost contractive sets and the SIA index.

8.6.1 Deciding whether a Set is Almost Contractive

Proposition 8.2 states that for a set S, the existences of a positive-column product, of a scrambling product, of a Sarymsakov product, and of an SIA product are equivalent. An algorithm to decide whether a set is almost contractive has been described in [PV12, Section 5]. Essentially, the decision procedure to determine whether a set S has a positive-column product amounts to applying Theorem 7.3 to the associated automaton of S. The graph of pairs of S' can be computed directly from S, without actually computing S' that can contain a very large number of matrices. The total complexity is $\mathcal{O}(mn^4)$ [PV12, Section 5].

8.6.2 Complexity of Approximating the Indices

In this section we prove some negative complexity results related to the approximation of the indices $\mathrm{sia}(S)$, $\mathrm{sar}(S)$, $\mathrm{scr}(S)$ and $\mathrm{pc}(S)$ for a given set of stochastic matrices S. The first result is a direct generalization of Theorem 7.5.

Theorem 8.5. *For every $\epsilon > 0$, it is NP-hard to approximate the Sarymsakov, scrambling and positive-column indices of sets of $n \times n$ stochastic matrices within a multiplicative factor $n^{1-\epsilon}$.*

Proof. For a set of automaton matrices S, $\mathrm{sar}(S) = \mathrm{scr}(S) = \mathrm{pc}(S)$ (Corollary 8.1). Therefore, the problems of approximating $\mathrm{sar}(S)$, $\mathrm{scr}(S)$ and $\mathrm{pc}(S)$ within a multiplicative factor $n^{1-\epsilon}$ are NP-hard when restricted to sets of automaton matrices (Theorem 7.5) and hence NP-hard in general. ■

Next, we prove an inapproximability result for the SIA index (Theorem 8.6). This result also holds in the case where the matrices have a positive diagonal. Matrices with positive diagonal elements often appear in consensus applications, where a positive diagonal element represents the weight of the own value of an agent in the computation of its new value. Restricting a problem can make it easier and many problems become easier for sets of stochastic matrices with positive diagonals. For example a set S is almost contractive if and only if its graph $\mathcal{G}(S)$, defined as the union of the graphs $\mathcal{G}(A_k)$ of which A_k are the adjacency matrices, has a spanning in-tree. Constructing this graph $\mathcal{G}(A_k)$ is easier than constructing the graph of pairs and thus deciding whether a set is almost contractive is easier in the case of matrices with positive diagonal. Almost contractive sets of matrices with a positive diagonal have a positive-column index smaller than or equal to $n-1$ (and a positive-column product is given by the spanning in-tree).

Theorem 8.6 also holds for the Sarymsakov index, the scrambling index and the positive-column index. This is new in the case of matrices with a positive diagonal. In the general case, however, Theorem 8.5 is stronger because it shows the inapproximability with larger ratios.

The reduction that we use is an adaptation of the reduction in [GH11] from COVER-SET to computing the reset threshold of an automaton. Here, we will reduce the COVER-SET problem to the problem of computing the SIA index of sets of stochastic matrices with a particular pattern. This pattern is the following:

$$\begin{pmatrix} 1 & & & \\ x & x & & \\ \vdots & & \ddots & \\ x & & & x \end{pmatrix}, \tag{8.16}$$

where the elements that are not indicated are 0 and x denotes an element that can be either positive or zero (at least one in each row will be positive because the matrices are stochastic). This pattern is preserved under multiplication. To prove the inapproximability of the SIA index, we will need the fact that a matrix with this pattern is SIA if and only if it is positive-column. It is a consequence of the following lemma.

Lemma 8.2. *Let S be a set of stochastic matrices with a zero pattern as described in Equation* (8.16) *and let $P = A_{w_1} \dots A_{w_\ell}$ be a product of matrices from S. Then P has a positive $(i,1)$ element if and only if one of the matrices A_{w_k} has a positive $(i,1)$ element:*

$$P_{i1} > 0 \Leftrightarrow \exists k,\ (A_{w_k})_{i1} > 0.$$

Proof. If: Let A_{w_k} be the smallest k such that $(A_{w_k})_{i1} > 0$. Then, by definition of k:

$$\forall j < k,\ (A_{w_j})_{i1} = 0$$

and thus

$$\forall j < k,\ (A_{w_j})_{ii} > 0 \tag{8.17}$$

because stochastic matrices have at least one positive element in each row. We also have

$$\forall j > k,\ (A_{w_j})_{11} > 0,$$

which, combined with $(A_{w_k})_{i1} > 0$ and (8.17) gives $P_{i1} > 0$.

Only if: Suppose $P_{i1} > 0$ and let us assume, to obtain a contradiction, that $\nexists k,\ (A_{w_k})_{i1} > 0$, that is $\forall k,\ (A_{w_k})_{i1} = 0$. We have then

$$(A_{w_k})_{ij} = \begin{cases} 1 & \text{if } i = j; \\ 0 & \text{otherwise} \end{cases}$$

and

$$(A_{w_k} A_{w_h})_{i1} = \sum_j (A_{w_k})_{ij} (A_{w_h})_{j1} = (A_{w_h})_{i1} = 0.$$

Iterating this reasoning conclude the "only if" part of this lemma. ■

Corollary 8.2. *Let P be a stochastic matrix with a pattern as in Equation 8.16. The following assertions are equivalent*

(i) P is SIA

(ii) P is Sarymsakov

(iii) P is scrambling

(iv) P is positive-column.

Proof. The implications $(iv) \Rightarrow (iii) \Rightarrow (ii) \Rightarrow (i)$ are consequences of $S_{\text{PC}} \subset S_{\text{SCR}} \subset S_{\text{SAR}} \subset S_{\text{SIA}}$ (Proposition 2.4). If P is SIA then there is p such that P^p has a positive column (Proposition 2.3). This positive (first) column is also positive in P by Lemma 8.2. ■

Theorem 8.6. *For every $\alpha > 0$, it is NP-hard to approximate the SIA, Sarymsakov, scrambling or positive-column index of sets of matrices of size $n \times n$ within a factor $(1 - \alpha)\log(n)$. This remains true if the problem is restricted to automaton matrices or to matrices with positive diagonal.*

Proof. We will make two reductions from the COVER-SET problem (Definition 7.6).

- In the first reduction, we construct a set S of automaton matrices from a COVER-SET instance and we prove that $\text{sia}(S)$ is equal to the size of the smallest cover of U. This first reduction will prove the NP-hardness of approximating the SIA index within a factor $(1 - \alpha)\log(n)$ in general and when the problem is restricted to automaton matrices.

- In the second reduction, we construct a set S' of matrices with a positive diagonal and we prove that $\mathrm{sia}(S')$ is equal to the size of the smallest cover of U. This will prove the NP-hardness of approximating the SIA index within a factor $(1-\alpha)\log(n)$ when the problem is restricted to matrices with a positive diagonal.

The hardness of approximating the other indices follows from Corollary 8.2 (matrices of both S and S' satisfy the hypotheses of this corollary).

We start from a SET-COVER instance $U = \{1, \ldots, n\}$ and $\mathcal{F}$ a collection of subsets of U. For every set $T \in \mathcal{F}$ we construct an $(n+1) \times (n+1)$ matrix A_T:

$$\begin{aligned} (A_T)_{11} &= 1 \\ (A_T)_{i1} &= 1 \text{ when } (i-1) \in T \\ (A_T)_{ii} &= 1 \text{ when } (i-1) \notin T \end{aligned}$$

and the other elements are 0. Clearly, each A_T is an automaton matrix (and it is therefore also stochastic). Let $S = \{A_T \mid T \in \mathcal{F}\}$. By Corollary 8.2 a product of matrices from S is SIA if and only if it has a positive (first) column. By Lemma 8.2 a product $P = A_{T_1} \ldots A_{T_\ell}$ of matrices from S has a positive column if and only if

$$U = \bigcup_{i=1,\ldots,\ell} T_i.$$

Therefore, $\mathrm{sia}(S)$ is equal to the size of the smallest cover of U and the first reduction is complete.

The second reduction is the same with the set $S' = \{A'_T \mid T \in \mathcal{F}\}$, where each A'_T is a stochastic matrix with a positive diagonal, defined as:

$$\begin{aligned} (A'_T)_{11} &= 1 \\ (A'_T)_{i1} &= 0.5 \text{ and } (A'_T)_{ii} = 0.5 \text{ when } (i-1) \in T \\ (A'_T)_{ii} &= 1 \text{ when } (i-1) \notin T. \end{aligned}$$

■

8.7 Conclusion

Summary of the Results

In this chapter we studied almost contractive sets of stochastic matrices. We have defined different indices corresponding to the classes of matrices that are often used in the analysis of the convergence of products of stochastic matrices. We have clarified the relation between these indices.

We have studied $\mathrm{sia}(n)$, $\mathrm{sar}(n)$, $\mathrm{scr}(n)$ and $\mathrm{pc}(n)$, the largest values that these indices can take among all sets of $n \times n$ stochastic matrices. We have shown that three of these quantities are equal and we therefore focused on the remaining quantities $\mathrm{sia}(n)$ and $\mathrm{pc}(n)$.

We have shown that $\mathrm{pc}(n)$ is in fact equal to the well studied quantity $C(n)$, the largest reset threshold among all n states automata. This allows to conclude that

$$(n-1)^2 \leq \mathrm{pc}(n) \leq \frac{n^3 - n}{6}.$$

We have obtained similar results for $\mathrm{sia}(n)$, as we have proved that

$$\mathrm{sia}(n) = \mathrm{sia}_A(n),$$

where $\mathrm{sia}_A(n)$ is the largest SIA index among all sets of *automaton matrices*. This has led to the following bounds:

$$n \leq \mathrm{sia}(n) \leq \frac{n^3 - n}{6}$$

and we have conjectured that

$$\mathrm{sia}(n) \leq 2n.$$

To support this conjecture, we have computed the SIA-index of

- all 2 letters automata up to 7 states (or said differently all sets of two automaton matrices),
- all 2 letters initially connected automata up to 9 states,
- all 3 letters initially connected automata up to 5 states,
- all automata that are known to be tight for the Černý conjecture.

Finally, we have studied the complexity of deciding whether a set of stochastic matrix has an SIA, Sarymsakov, scrambling or positive-column index smaller than or equal to a given ℓ.

Remaining Challenges

Several problems remain open. First, there are significant gaps between the lower and upper bounds on $\mathrm{sia}(n)$ and $\mathrm{pc}(n)$. Closing the gap between the upper and the lower bound on $\mathrm{pc}(n)$ seems to be a very hard problem since it is equivalent to solving the Černý conjecture, that has been open for more than fifty years. Proving Conjecture 8.1 also seems very hard for the same reason. However, we do not know whether the relation between Conjecture 8.1 and the Černý conjecture goes both ways. Indeed, we have shown that a proof of our conjecture would greatly improve the state of the art on the Černý conjecture. But we do not know if a counterexample to our conjecture would help building a counterexample to the Černý conjecture.

Another open problem is to find the best approximation ratio achievable by polynomial-time algorithms computing the SIA index of a given matrix set. Since our contribution is based on synchronizing automata theory, the methods used in [GS15] can potentially be used to establish the exact ratio.

We have proved that for each matrix size n, there is an almost contractive set of automaton matrices that has the largest possible SIA index among stochastic matrices of size $n \times n$. We wonder whether representing the largest SIA index as the solution of an optimization problem on the space of stochastic matrix sets can lead to another proof of this result.

Indeed, finding the set of stochastic matrices that has the largest SIA index is an optimization problem. Optimization problems often have their optimal solution on the vertices of the admissible set[3] and automaton matrices are the vertices of the polytope of stochastic matrices[4]. A formulation of this optimization problem could potentially unify and generalize similar results appearing in [GGJ18, BJO15].

The main obstacle to this is the characterization of the admissible set of this optimization problem. The admissible set would be the set of almost contractive sets, whose vertices are not simply the set of almost contractive sets of automaton matrices. Indeed, the set

$$\left\{ \begin{pmatrix} 0 & 0.5 & 0.5 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{pmatrix} \right\}$$

provides an example of almost contractive set that is not a convex combination of almost contractive sets of automaton matrices.

[3] for example, when the admissible set and the objective function are convex

[4] Any stochastic matrix is equal to a convex combination of automaton matrices, and no automaton matrix is a convex combination of other automaton matrices.

Part III

Conclusion

Conclusion

We have studied contractive sets in Part I. We have developed in Chapter 4 an algorithm to decide whether a set is contractive. To obtain this algorithm, we have elaborated innovative techniques using invariant polyhedra. In particular, we have analyzed the polyhedron $\mathcal{P}$ that is invariant for P-preserving matrices, a class that contains all stochastic matrices. We concluded this chapter by answering an open question mentioned in [BO14]: the complexity of deciding whether a set of two undirected matrices is contractive.

In Chapter 5, we have developed further our analysis of invariant polyhedra and we have obtained a tight finiteness bound. This bound characterizes the length of products that should be analyzed in order to determine convergence of all infinite products.

In Chapter 6, we have explored the generality of the approach developed in Chapters 4 and 5. We have seen in the first section of this chapter that this approach is very general and can be applied to a different setting and polyhedron, namely to the study of primitive sets of matrices. In the second section, we have shown that not all symmetric polytopes (and thus not all symmetric polyhedra) have a lattice of double-faces that has the Sperner property. Thus, the results of Chapters 4 and 5 are still applicable to any polyhedron but the computation of the finiteness bound will be more difficult in some cases.

A remaining open question is the existence of other applications in which these developments could be useful. Indeed, our theory is on the one hand very general (i.e., it applies to all systems that have an invariant polyhedron) but on the other hand, we have only been able to exhibit two relevant applications.

In Chapter 4, we have left open the exact complexity of deciding whether a set is contractive. We have proven the existence of a decision algorithm with singly exponential time complexity and singly exponential space complexity, while earlier work has proved NP-hardness of the problem [BO14]. We still do not know whether the problem is in PSPACE and whether there exists an algorithm that has a polynomial space complexity.

In Chapter 5, the tight finiteness bound that we obtained depends on the dimension of the matrices but does not depend on the number of matrices in the set. It would be interesting to analyze whether finer bounds can be obtained when they depend on the number of matrices. A good starting point would be to analyze the case of sets of two matrices.

In Part II, we considered almost contractive sets. We studied the shortest SIA, Sarymsakov, scrambling and positive-column products of almost contractive sets. We defined indices that characterize the lengths of these shortest products and we defined for each matrix size n quantities $\mathrm{sia}(n)$, $\mathrm{sar}(n)$, $\mathrm{scr}(n)$ and $\mathrm{pc}(n)$ that are the largest indices among all sets of $n \times n$ stochastic matrices. We have shown that $\mathrm{sar}(n) = \mathrm{scr}(n) = \mathrm{pc}(n)$, reducing our study to two remaining relevant quantities $\mathrm{sia}(n)$ and $\mathrm{pc}(n)$. We obtained lower and upper bounds on them by establishing a link with automata theory and the Černý conjecture. We indeed showed that the largest SIA and positive-column indices $\mathrm{sia}(n)$ and $\mathrm{pc}(n)$ can be obtained with sets of automaton matrices. We have shown that the automata that are known to have a large reset threshold (i.e., the automata that are tight for the Černý conjecture) have in fact a small SIA index. This has led us to conjecture that the SIA index is always small and in particular, that $\mathrm{sia}(S) \leq 2n$. We have then supported this conjecture with an computer search for the largest SIA index among almost contractive sets.

This conjecture, and the link with automata theory has also possible consequences for the Černý conjecture. Indeed, a proof of our conjecture would yield an upper bound of $2n(n-1)$ on the reset threshold of automata, which would be much better that the best known bound of $\frac{n^3-n}{6}$.

Contractive set can be recognized using a polynomial-time algorithm described in [PV12]. We have proven, however, that the indices are hard to compute. That is, the existence of a SIA product can be decided in polynomial time but the existence of an SIA of length $\leq \ell$ for a given ℓ is an NP-hard problem. Furthermore, even approximating the indices is NP-hard.

An important remaining challenge from Chapter 8 is of course our conjecture on the SIA index. There is a large gap between the best lower bound n and the best upper bound $\frac{n^3-n}{6}$ that we were able to obtain on $\mathrm{sia}(n)$ the largest SIA index among all sets of $n \times n$ matrices.

Another open question was mentioned in the conclusion of this chapter.

We have seen that the largest SIA, Sarymsakov, scrambling or positive-column index among all sets of $n \times n$ matrices can always be obtained with a set of automaton matrix and we wonder whether there is a general principle unifying these results. As already suggested, a formulation of the indices as solution of optimization problems might provide this unifying principle.

We would like to conclude by observing that many questions are still open. We have only asked two questions on products of stochastic matrices: the convergence of all infinite products and the convergence of one infinite product. However, many other questions can be asked about switching systems with stochastic transition matrices, such as the convergence in the case of state-dependent switching (where the transition matrix depends on the state of the system) or in the case of constrained switching (where only a set of sequences of transition matrices are possible).

Bibliography

[AK09] M. Akelbek and S. Kirkland. Primitive digraphs with the largest scrambling index. *Linear Algebra and its Applications*, 430(4):1099 – 1110, 2009.

[AS03] D. Angeli and E. D. Sontag. Monotone Control Systems. *IEEE Transactions on Automatic Control*, 48:1684–1698, 2003.

[ATB86] M. Athans, J. Tsitsiklis, and D. Bertsekas. Distributed asynchronous deterministic and stochastic gradient optimization algorithms. *IEEE Transactions on Automatic Control*, 1986.

[AVG13] D. S. Ananichev, M. V. Volkov, and V. V. Gusev. Primitive digraphs with large exponents and slowly synchronizing automata. *Journal of Mathematical Sciences*, 192(3):263–278, 2013.

[Bar88] N. Barabanov. Lyapunov indicators of discrete inclusions i-iii. *Automation and Remote Control*, 49:152–157, 283–287, 558–565, 1988.

[BHOT05] V.D. Blondel, J.M. Hendrickx, A. Olshevsky, and J.N. Tsitsiklis. Convergence in multiagent coordination, consensus, and flocking. In *Proceedings of the* 44^{th} *IEEE Conference on Decision and Control*, 2005.

[BHT07] V. D. Blondel, J. M. Hendrickx, and J. N. Tsitsiklis. On the 2R conjecture for multi-agent systems. In *Proceedings of the European Control Conference*, 2007.

[BHT09] V. D. Blondel, J. M. Hendrickx, and J. N. Tsitsiklis. On Krause's multi-agent consensus model with state-dependent connectivity. *IEEE Transactions on Automatic Control*, 54:2586–2597, 2009.

[Bir57] Garrett Birkhoff. Extensions of jentzsch's theorem. *Transactions of the American Mathematical Society*, 85(1):219–227, 1957.

[BJMP12] B. Bamieh, M. R. Jovanovic, P. Mitra, and S. Patterson. Coherence in large-scale networks: dimension dependent limitations of local feedback. *IEEE Transactions on Automatic Control*, 57(9):2235–2249, 2012.

[BJO15] V. D. Blondel, R. M. Jungers, and A. Olshevsky. On primitivity of sets of matrices. *Automatica*, 61:80–88, 2015.

[BL90] R. A. Brualdi and B. Liu. Generalized exponents of primitive directed graphs. *Journal of Graph Theory*, 14(4):483–499, 1990.

[BO14] V. D. Blondel and A. Olshevsky. How to decide consensus? A combinatorial necessary and sufficient condition and a proof that consensus is decidable but NP-hard. *SIAM Journal on Control and Optimization*, 52(5):2707 – 2726, 2014.

[BR91] R. A. Brualdi and H. Ryser. *Combinatorial Matrix Theory*. Cambridge University Press, 1991.

[Čer64] J. Černý. Poznámka k homogénnym eksperimentom s konečnými automatami. *Matematicko-fysikalny Casopis SAV*, 14:208–216, 1964.

[CGHJ17] P.-Y. Chevalier, V. V. Gusev, J. M. Hendrickx, and R. M. Jungers. Sets of stochastic matrices with converging products: Bounds and complexity. *Submitted*, 2017.

[CHJ14] P.-Y. Chevalier, J. M. Hendrickx, and R. M. Jungers. A switched system approach to the decidability of consensus. In *Book of Abstracts of the International Symposium on the Mathematical Theory of Networks and Systems*, 2014.

[CHJ15a] P.-Y. Chevalier, J. M. Hendrickx, and R. M. Jungers. Efficient algorithms for the consensus decision problem. *SIAM Journal on Control and Optimization*, 2015.

[CHJ15b] P.-Y. Chevalier, J. M. Hendrickx, and R. M. Jungers. Reachability of consensus and synchronizing automata. *IEEE Control and Decision Conference*, 2015.

[CHJ16] P.-Y. Chevalier, J. M. Hendrickx, and R. M. Jungers. Tight bound for deciding convergence of consensus systems. *Systems and Control Letters*, 2016.

[CS82] J. E. Cohen and P. H. Sellers. Sets of nonnegative matrices with positive inhomogeneous products. *Linear Algebra and Its Applications*, 1982.

[CW08] D. Coppersmith and C. W. Wu. Conditions for weak ergodicity of inhomogeneous markov chains. *Statistics & Probability Letters*, 78(17):3082 – 3085, 2008.

[DS14] I. Dinur and D. Steurer. Analytical approach to parallel repetition. In *ACM Symposium on Theory of Computing (STOC)*, pages 624–633, 2014.

[Eck06] J. Eckhoff. Combinatorial properties of f-vectors of convex polytopes. *Nordisk Matematisk Tidsskrift*, 54:146–159, 2006.

[Eng97] K. Engel. *Sperner Theory*. Cambridge University Press, 1997.

[Epp90] D. Eppstein. Reset sequences for monotonic automata. *SIAM Journal on Computing*, 19(3):500–510, 1990.

[Fra82] P. Frankl. An extremal problem from two families of sets. *European Journal of Combinatorics*, 3:125–127, 1982.

[GGJ18] B. Gerencsér, V. V. Gusev, and R. M. Jungers. Primitive sets of nonnegative matrices and synchronizing automata. *SIAM Journal on Matrix Analysis and Applications*, 39(1):83–98, 2018.

[GH11] M. Gerbush and B. Heeringa. Approximating minimum reset sequences. In *Implementation and Application of Automata*, volume 6482 of *LNCS*, pages 154–162. Springer, 2011.

[GJT15] F. Gonze, R. M. Jungers, and A. N. Thrahtman. A note on a recent attempt to improve the pin-frankl bound. *Discrete Mathematics and Theoretical Computer Science*, 17(1):307–308, 2015.

[Grü03] B. Grünbaum. Convex polytopes. In *Graduate Text in Mathematics*. Springer, New York, 2003.

[GS15] P. Gawrychowski and D. Straszak. Strong inapproximability of the shortest reset word. In *Mathematical Foundations of Computer Science 2015*, pages 243–255. Springer Berlin Heidelberg, 2015.

[Gus13] V. V. Gusev. Lower bounds for the length of reset words in eulerian automata. *International Journal of Foundations of Computer Science*, 24(02):251–262, 2013.

[Har02] D. J. Hartfiel. *Nonhomogeneous Matrix Products*. World Scientific, 2002.

[HT13] J. M. Hendrickx and J. N. Tsitsiklis. Convergence of type-symmetric and cut-balanced consensus seeking systems. *IEEE Transactions on Automatic Control*, 58:214–218, 2013.

[JLM03] A. Jadbabaie, J. Lin, and A. S. Morse. Coordination of groups of mobile agents using nearest neighbor rules. *IEEE Transactions on Automatic Control*, 48:988–1001, 2003.

[Jun09] R. M. Jungers. *The Joint Spectral Radius: Theory and Applications*. Springer, 2009.

[Kah62] A. B. Kahn. Topological sorting of large networks. *Comm. ACM*, 5:558–562, 1962.

[Lot97] M. Lothaire. *Combinatorics on Words*. Cambridge University Press, 1997.

[LW95] J. C. Lagarias and Y. Wang. The finiteness conjecture for the generalized spectral radius of a set of matrices. *Linear Algebra and its Applications*, 214:17–42, 1995.

[Lyn54] R.C. Lyndon. On Burnside's problem. *Transactions of the American Mathematical Society*, 77(4):202–215, 1954.

[Mar06] M. Margaliot. Stability analysis of switched systems using variational principles: an introduction. *Automatica*, 42:2059–2077, 2006.

[ME10] M. Mesbahi and M. Egerstedt. *Graph Theoretic Methods for Multiagent Networks*. Princeton University Press, 2010.

[NO09] A. Nedić and A. Ozdaglar. Distributed subgradient methods for multi-agent optimization. *IEEE Transactions on Automatic Control*, 54:48–61, 2009.

[OSFM07] R. Olfati-Saber, J. A. Fax, and R.M Murray. Consensus and cooperation in networked multi-agent systems. In *Proceedings of the IEEE*, 2007.

[OSS05] R. Olfati-Saber and J. S. Shamma. Consensus filters for sensor networks and distributed sensor fusion. In *Proceedings of the 44th IEEE Conference on Decision and Control*, 2005.

[Paz65] A. Paz. Definite and quasidefinite sets of stochastic matrices. *Proceedings of the American Mathematical Society*, 16(4):634–641, 1965.

[Paz71] A. Paz. *Introduction to Probabilistic Automata*. Academic Press, New York, 1971.

[Pin83] J.-É. Pin. On two combinatorial problems arising from automata theory. *Annals of Discrete Mathematics*, 17:535–548, 1983.

[PV12] V. Y. Protasov and A. S. Voynov. Sets of nonnegative matrices without positive products. *Linear Algebra and its Applications*, 437:749–765, 2012.

[Rys97] I. Rystsov. Reset words for commutative and solvable automata. *Theoretical Computer Science*, 172(1):273 – 279, 1997.

[Sch98] A. Schrijver. *Theory of Linear and Integer Programming*. John Wiley & sons, 1998.

[Sen81] E. Seneta. *Non-negative Matrices and Markov Chains*. Springer-Verlag, New York, 1981.

[Sen92] E. Senata. Coefficients of ergodicity: structure and applications. *Advances in Applied Probability*, 11:194–201, 1992.

[Smi95] H. Smith. *Monotone dynamical systems: An introduction to the theory of competitive and cooperative systems*. Mathematical Surveys and Monographs, 1995.

[Spe28] E. Sperner. Ein Satz über Untermengen einer endlichen Menge. *Mathematische Zeitschrift*, 1928.

[SS03] B. L. Shader and S. Suwilo. Exponents of nonnegative matrix pairs. *Linear Algebra and its Applications*, 363:275 – 293, 2003.

[SSR10] R. Sepulchre, A. Sarlette, and P. Rouchon. Consensus in Non-Commutative Spaces. *IEEE Control and Decision Conference*, 2010.

[Str05] D. W. Strook. *An Introduction to Markov Processes*. Springer, Berlin, Heidelberg, 2005.

[Szy18] M. Szykuła. Improving the upper bound on the length of the shortest reset words. *Arxiv preprint*, 2018.

[VCBJ+95] T. Vicsek, Cziroók, E. Ben-Jacob, I. Cohen, and O. Shochet. Novel type of phase transition in a system of self-deriven particles. *Physical Review Letters*, 75:1226, 1995.

[Vol08] Mikhail V Volkov. Synchronizing automata and the Černỳ conjecture. In *Language and Automata Theory and Applications*, pages 11–27. Springer, 2008.

[Wol63] J. Wolfowitz. Products of indecomposable, aperiodic, stochastic matrices. *Proceedings of the American Mathematical Society*, 15:733–736, 1963.

[XBK07] L. Xiao, S. Boyd, and S.-J. Kim. Distributed average consensus with least-mean-square deviation. *Journal of Parallel and Distributed Computation*, 67(1):33–46, 2007.

[XLC+15] W. Xia, J. Liu, M. Cao, K. H. Johansson, and T. Başar. Products of generalized stochastic Sarymsakov matrices. *IEEE Control and Decision Conference*, 2015.

[Zie95] B. Ziegler. Lectures on polytopes. In *Graduate Text in Mathematics*. Springer, New York, 1995.