Zero-sum repeated games : accelerated algorithms and tropical best-approximation

Omar Saadi

INSTITUT POLYTECHNIQUE DE PARIS

NNT : 2021IPPAX102

Thèse de doctorat

Zero-sum repeated games: accelerated algorithms and tropical best approximation

Thèse de doctorat de l'Institut Polytechnique de Paris
préparée à l'École polytechnique

École doctorale n°574 École doctorale de mathématiques Hadamard (EDMH)
Spécialité de doctorat : Mathématiques appliquées

Thèse présentée et soutenue à Palaiseau, le Vendredi 17 Décembre 2021, par

OMAR SAADI

Composition du Jury :

M. Vineet Goyal Associate Professor, Columbia University (IEOR)	Rapporteur
M. William M. McEneaney Professor, University of California, San Diego (Dept. of Mechanical and Aerospace Engineering)	Rapporteur
M. Omid Amini Professeur Hadamard, École Polytechnique (CMLS)	Président du jury
M. Petros Maragos Professor, National Technical University of Athens (School of Electrical and Computer Engineering)	Examinateur
Mme Zheng Qu Assistant professor, The University of Hong Kong (Department of Mathematics)	Examinateur
Mme Josephine Yu Associate Professor, Georgia Institute of Technology (School of Mathematics)	Examinateur
Mme Marianne Akian Directeur de recherche, Inria et École Polytechnique (CMAP)	Directrice de thèse
M. Stéphane Gaubert Directeur de recherche, Inria et École Polytechnique (CMAP)	Directeur de thèse

Acknowledgments

I would like to express my supreme gratitude to my advisors, Stéphane Gaubert and Marianne Akian for their wonderful guidance during my thesis. It was a great chance for me to work and spend time with them during these three years. Stéphane's enthusiasm and large knowledge in mathematics was a constant source of inspiration and motivation for me. The same has to be said about Marianne's great rigor and concern for accuracy and precision that encouraged me to explore the ideas and concepts always deeper. I thank them for the countless hours and meetings that we spent working together. You offered me ideal conditions to carry out this thesis. Thank you.

I want to thank Vineet Goyal and William M. McEneaney for the great honor they have conferred upon me by reviewing this manuscript. I also wish to thank Omid Amini, Petros Maragos, Zheng Qu and Josephine Yu for expressing their interest in my work as well as accepting to take part in the jury.

I thank all the people with whom I have been able to work, write papers or simply discuss various topics in applied mathematics and who have helped me move forward in my research. I thank in particular Zheng Qu for all the discussions and the collaboration that we had together, and for inviting me to Hong Kong even if it did not finally occur due to the pandemic. I thank also Yang Qi for our coffees, discussions and long working hours that were always a pleasure to have with him.

I would also like to thank the members of the TROPICAL team for their encouragement and support. I thank especially Xavier Allamigeon, Hanieh Tavakolipour and Marin Boyet for all the enjoyable and helpful discussions and exchanges.

My thanks also go to the other CMAP doctoral students, especially to Florian Bourgey, Mathilde Boissier, Hugo Girardon, Mehdi Talbi, Abdelhak Touiti, Wanqing Wang, Manon Rivoire, Nicolas Vandame, Benoît Tran, Maxime Grangereau, Paulin Jacquot, with whom I shared an important part of my life and have all given me a hand at one time or another.

I thank the administrative teams of CMAP and INRIA notably Nasséra Naar, Alexandra Noiret and Hanadi Dib, for their kindness and helpfulness.

I am also deeply and truly grateful to my parents and my brother for their unconditional love and support, as well as to all the members of my family and my friends.

Contents

CHAPTER 1

Introduction

1.1 Context of this work

1.1.1 Zero-sum stochastic games and Markov Decision Processes

Zero-sum stochastic games are classical models, introduced by Shapley [Sha53, NS03]. They allow to study sequential dynamic interaction between two agents evolving in an uncertain environment. The agents involved are called *players*. Their interests are opposed in the sense that one player pays the outcome of the game to the other. So that, one player, the minimizer (MIN), wants to minimize the outcome of the game, and the other player, the maximizer (MAX), wants to maximize it. The game is played in steps, and it evolves in a set of states S. At each step $k \geqslant 1$, a reward r_k is generated. It depends on the current state of the game s_k and the pair of actions (a_k, b_k) chosen by the two players. The actions of the two players determine also the transition probability to the next state. In a *discounted* stochastic game, the total payment generated is given by $\sum_{k\geqslant 1} \gamma^k r_k$, were $\gamma \in [0, 1)$ is a discount factor. Solving a stochastic game consists in finding the *value* of the game v_i, when the game starts from any given state $i \in S$. The scalar v_i is the value that each player can guarantee regardless of the actions of the other player. In the case where the two players are "patient", in the sense that they don't differentiate current and future payments ($\gamma = 1$), the game is said to be *undiscounted*. It corresponds to the limit of a discounted stochastic game when γ tends to 1. In this case, the total payment that player MIN wants to minimize and player MAX wants to maximize is the $\limsup$ or $\liminf$ of a mean $\frac{1}{k}\sum_{l=1}^{k} r_l$ as $k \to \infty$. In this case, the problem is called a *mean payoff zero-sum stochastic game* [LL69a, MN81a], and, the main quantity of interest is the mean payoff per time unit χ_i that each player can guarantee, when the

game starts from the state $i \in S$.

In the above general framework, when the players play simultaneously at each step, we say that the game is in *imperfect information*. In computer science, the case of games in *perfect information* was particularly studied [RF91, Con92, FV12, HMZ13]. In this case, the players play one after the other. The term *turn-based* is sometimes used as a synonym of perfect information. This case can be seen as a special case of imperfect information, and it is fundamentally easier. Indeed, we have the property that when the set of states and sets of actions are finite, and the input data of the game is rational, the value of the turn-based game is also rational. However, in the imperfect information case, this value is an algebraic number [BK76].

In the special case where there is only one player, or equivalently the other player has no influence on the game, the perfect information zero-sum game becomes a *Markov decision process* (MDP) [Bel57, How60, Put14, Ber17]. In this framework also, the agent aims to make a trade-off between the current payment that he receives and the evolution of the state of the system that is influenced by his actions, and that will impact his future payments. Similarly to games, we also distinguish discounted and undiscounted Markov decision processes. Stochastic games and Markov decision processes intervene in various fields ranging from reinforcement learning [VOW12, Lit94], finance [BR11], economy [Ami03], to health care [BH13] and ecology [Wil09].

Dynamic programming and Shapley operator

Dynamic programming is one of the main approaches used to solve Markov decision processes and stochastic games. It allows one to transform a game to a fixed point problem involving an operator T called *Shapley operator* in the case of games or *Bellman operator* in the case of Markov decision processes.

In the special case of a finite set of states $S = \{1, \cdots, n\}$, the Shapley operator $T : \mathbb{R}^n \to \mathbb{R}^n$ is a map whose ith coordinate is given by

$$T_i(v) = \min_{a \in A_i} \max_{b \in B_i} \{r_i^{ab} + \gamma \sum_{j \in S} P_{ij}^{ab} v_j\}, \quad i \in S,\ v \in \mathbb{R}^n \ .$$

For each $i \in S$, A_i is a set representing the possible actions of player MIN in state i, and B_i is a set representing the possible actions of player MAX in state i.We denote by $E := \{(i, a, b) \mid i \in S, a \in A_i, b \in B_i\}$ the set of all admissible triples state-actions. For all $(i, a, b) \in E$, P_i^{ab} is an element of $\Delta(S)$ the set of probability measures on S; we shall identify $P_i^{ab} = (P_{ij}^{ab})_{j \in S}$ to a row vector in $\mathbb{R}^n$, where P_{ij}^{ab} is the transition probability to the next state j, given the current state i and the actions taken $a \in A_i, b \in B_i$. For all $(i, a, b) \in E$, r_i^{ab} is a reward (real number) that MIN pays to MAX, and γ is a discount factor. We mention that the imperfect information case can be reduced to this case, when the sets of actions A_i and B_i are simplices.

In the discounted case $\gamma < 1$, it is known that the value vector $v = (v_i)_{i \in S}$ does exist. It is characterized as the unique solution of the fixed point problem

$$v = T(v) \ .$$

In the undiscounted case, where γ is identically 1, the main quantity of interest is the *mean payoff vector* [BK76]:

$$\chi(T) := \lim_{k \to \infty} T^k(0)/k \ .$$

The entry $\chi_i(T)$ represents the mean payoff per time unit, if the initial state is i. Here, the mean payoff is defined by considering a family of games in finite horizon k as k tends to infinity. There are alternative

approaches, in which the mean payoff is defined as the value of an infinite horizon game [LL69a]. The property of the *uniform value* established in [MN81a] entails that the different natural approaches lead to the same notion of mean payoff.

The analysis of the mean payoff problem is simplified when the following *non-linear eigenproblem* has a solution:

$$\eta e + v = T(v), \qquad \eta \in \mathbb{R},\ v \in \mathbb{R}^n \ , \tag{1.1}$$

where $e := (1 \cdots 1)^\top \in \mathbb{R}^n$ is the unit vector. The scalar η is called the *ergodic constant*, whereas the vector v, which is not unique, is called *bias* or *potential*. When this equation is solvable, we have $\chi(T) = \eta e$, i.e., the mean payoff is independent of the initial state, and it is equal to the ergodic constant. See e.g. [AGH18] for background.

Value iteration and policy iteration [How60, Put14] are two fundamental dynamic programming methods. They are used to solve turn-based stochastic games, both in the discounted and the undiscounted case. In the undiscounted case, one needs to modify the value iteration algorithm to include a normalization term, needed to get a bounded sequence, this is called *relative value iteration* [Whi63]. For discounted problems with a fixed discount factor, value iteration allows one to find an optimal policy in a time which is polynomial [Tse90] but not strongly polynomial [FH14]. Ye showed that policy iteration runs in strongly polynomial time for Markov decision processes with a fixed discount factor [Ye11a]. This result was subsequently extended to two-player zero-sum games by Hansen, Miltersen and Zwick [HMZ13]. However, for two player mean payoff games, Friedmann [Fri09b, Fri11] showed that policy iteration can take an exponential time, and Fearnley [Fea10] showed that the same is true for Markov decision processes. Hence, mean payoff problems are in the hardest class. However, some special mean payoff problems have been reduced to problems with a *fixed* discount factor, leading to parametrized complexity results, see [FABG13, FH13, AG13, Sch16]. We mention that deterministic mean payoff games and stochastic turn-based mean payoff games are among the problems in the complexity class NP ∩ co-NP [Con92, ZP96] for which no polynomial time algorithm is known.

Even for problems with a fixed discount, value iteration and policy iteration appear to be too slow, or unadapted, for huge scale instances. Algorithms based on Monte-Carlo simulations can lead to improved scalability. In a recent progress, Sidford et al. [SWWY18] combined value iteration algorithm with sampling and variance reduction techniques. They obtained an algorithm for discounted infinite-horizon MDPs that, remarkably, is sublinear in a certain relevant regime of the parameters. This result use in an essential way the discounted nature of the problem.

1.1.2 Nesterov's acceleration: from gradient descent to fixed point problems

Another idea recently used to deal with large scale instances of Markov decision processes consists on applying Nesterov's acceleration to value iteration. Nesterov's acceleration technique comes from the framework of gradient descent, where the goal is to minimize a function. Applying it to fixed point problems is tempting, but it is a challenging problem.

Nesterov proposed in [Nes83, Nes04] to accelerate the gradient descent scheme for the minimization of a μ-strongly convex function $f : \mathbb{R}^n \to \mathbb{R}$ whose gradient is of Lipschitz constant L, by adding an inertial step:

$$x_{k+1} = y_k - h\nabla f(y_k) \ , \tag{1.2a}$$

$$y_{k+1} = x_{k+1} + \alpha(x_{k+1} - x_k) \ , \tag{1.2b}$$

where $h > 0$, and $\alpha \in [0, 1]$ are parameters. Let x_* be the minimizer of f. When $\alpha = 0$, the scheme (1.2) reduces to gradient descent. With the step $h = 1/L$, the gradient descent converges

linearly with a rate $1 - 2\mu/(L+\mu)$. Indeed, we have $\|x_k - x_*\|^2 \leqslant \big(1 - 2\mu/(L+\mu)\big)^k \|x_0 - x_*\|^2$, and $f(x_k) - f(x_*) \leqslant \frac{L}{2}\big(1 - 2\mu/(L+\mu)\big)^k \|x_0 - x_*\|^2$, for all $k \geqslant 1$, see Theorem 2.1.14 in [Nes04]. Moreover, Theorem 2.2.3, *ibid.*, implies that if we choose

$$\alpha = \frac{1 - \sqrt{\mu/L}}{1 + \sqrt{\mu/L}}, \tag{1.3}$$

still with $h = 1/L$, the scheme (1.2) converges linearly with a rate $1 - \sqrt{\mu/L}$. Indeed, with α given by (1.3), we have $f(x_k) - f(x_*) \leqslant 2(1 - \sqrt{\mu/L})^k (f(x_0) - f(x_*))$ for all $k \geqslant 1$. Note that when the condition number L/μ is large, i.e. $L/\mu \gg 1$, the rate $1 - \sqrt{\mu/L}$ improves over $1 - 2\mu/(L+\mu)$, whence the scheme (1.3) is commonly known as *accelerated gradient descent*.

The idea of applying this Nesterov's type of scheme to the case of fixed point problems was recently studied in [IH19, GGC19]. Given the Bellman operator $T : \mathbb{R}^n \to \mathbb{R}^n$ of an MDP, the value iteration algorithm consists in starting from a vector $v_0 = (0, \cdots, 0) \in \mathbb{R}^n$, and iterating $v_{k+1} = T(v_k)$ for $k \geqslant 0$. A Nesterov scheme applied to this value iteration is of the following form

$$x_{k+1} = T(y_k) \ , \tag{1.4a}$$

$$y_{k+1} = x_{k+1} + \alpha(x_{k+1} - x_k) \ . \tag{1.4b}$$

The 0-player case, with a finite number n of states, is already of interest. In this case, the Bellman operator is an affine map $T : \mathbb{R}^n \to \mathbb{R}^n$ of the form $T(x) = g + Px$, where $g = (g_i) \in \mathbb{R}^n$ represents the payments and $P = (P_{ij}) \in \mathbb{R}^{n \times n}$ is a substochastic matrix, i.e. a matrix with nonnegative entries such that the sum of each row is less than or equal to 1. In this 0-player case, value iteration has an asymptotic (geometric) convergence rate given by the spectral radius of P. In many applications, this spectral radius is of the form $\gamma = 1 - \epsilon$ where ϵ is *small*. When the matrix P is *symmetric* , an algorithm with an improved rate $1 - \Omega(\epsilon^{1/2})$ can be obtained by specializing the accelerated gradient algorithm of Nesterov [Nes83]. The latter algorithm applies to the minimization of a smooth strictly convex function f, which, in the quadratic case, reduces to an affine fixed point problem with a symmetric matrix P, see [FB15]. In contrast, developing accelerated algorithms for problems of non-symmetric type is a challenging question, which has been studied recently in [IH19, GGC19]. In these works, the theoretical convergence results apply to matrices with a real spectrum, showing that the original choice of parameters for Nesterov's method in the symmetric case still yields an acceleration in this setting.

Apart from being applied to Markov decision processes, fixed point iteration also includes as a special case the proximal point method [Roc76b], when the mapping T corresponds to the resolvent of a maximal monotone operator. The proximal point method covers a list of pivotal algorithms in optimization such as the proximal gradient descent, the augmented Lagrangian method (ALM) [Roc76a] and the alternating directional method of multipliers (ADMM) [EB92]. The development of accelerated proximal point method has thus attracted a lot of attention [CMY15, AP19, Att21] and a recent paper [Kim19] constructed a new algorithm achieving $\|x_k - T(x_k)\| \leqslant O(1/k)$ through the performance estimation problem (PEP) approach [DT14]. In a more general setting when T is a nonexpansive mapping in a Euclidean norm, a version of Halpern's iteration was recently shown to yield a residual $\|x_k - T(x_k)\| \leqslant O(1/k)$ [Lie21], also via the PEP approach. These results improve over the worst case bound $\|x_k - T(x_k)\| \leqslant O(1/\sqrt{k})$ of the Krasnoselski-Mann's iteration for a nonexpansive mapping (in arbitrary norm) [BB96].

There is also a large body of literature on (quasi-)Newton type methods for solving nonlinear equations [rFS09, IS14, WN11], which can be naturally employed for solving fixed point problems and yield fast asymptotic convergence rate. It is well-known that such methods converge only when close enough to the solution. Some papers proposed various safe-guard conditions to globalize the convergence [TP19, ZOB20] and do not provide a rate of convergence.

1.1.3 Best approximation problems in tropical geometry

The *max-plus semiring* $(\mathbb{R}_{\max}, \oplus, \odot)$ is the set of real numbers, completed by $-\infty$ and equipped with the addition $(a, b) \mapsto a \oplus b := \max(a, b)$ and the multiplication $(a, b) \mapsto a \odot b := a + b$. Nowadays, the term *tropical semiring* is used to refer to the max-plus semiring, and also to the min-plus semiring, where the classical addition is replaced by the minimum. For some introductions to tropical geometry see [RGST05, IMS09, MS15b]. Tropical semirings were studied in relation with various problems, including discrete event systems [BCOQ92, CGQ99, HOVDW14], graph algorithms [GM84, But03b], Hamilton-Jacobi-Bellman partial differential equations [McE07, GMQ11, Qu14], and more recently machine learning methods [CM18, ZNL18, MCT21]. We recall that a subset $\mathcal{C}$ of $(\mathbb{R}_{\max})^n$ is a *tropical (convex) cone* or equivalently a *tropical submodule* of $(\mathbb{R}_{\max})^n$, if it satisfies $x, y \in \mathcal{C}$ and $\lambda \in \mathbb{R}_{\max}$ implies $\lambda + x \in \mathcal{C}$ and $x \vee y \in \mathcal{C}$, where $\lambda + x \in (\mathbb{R}_{\max})^n$ denotes the vector with entries $\lambda + x_i$, for $i \in [n]$, and $x \vee y = \sup(x, y)$ denotes the vector with entries $\max(x_i, y_i)$, for $i \in [n]$.

Tropical linear approximation

Several "best approximation" problems have been studied in tropical geometry. The simplest one consists in finding the nearest point in a (closed) tropical module, in the sense of Hilbert's metric. The solution is given by the tropical projection [CGQ04], see also [AGNS11]. The best approximation in the space of ultrametrics, which can be formulated in terms of approximation by a tropical module in view of its application to phylogenetics, has been thoroughly studied [CF00, LSTY17, Ber20]. Another important special case is the best approximation of a point by a tropical linear space [Ard04, JSY07]. These problems concern the approximation of a *single* point.

An interesting problem is to approximate a set of points by a *tropical linear space*. The *tropical Grassmannian* $\mathrm{Gr}_{k,n}^{\mathrm{trop}}$ can be defined as the image by a non-archimedean valuation of the Grassmannian $\mathrm{Gr}_{k,n}(\mathbb{K})$ over an (algebraically closed) non-archimedean field $\mathbb{K}$, under the Plücker embedding, see [SS04, FR15]. In this way, an element of $\mathrm{Gr}_{k,n}^{\mathrm{trop}}$ is represented by its *tropical Plücker coordinates* $p = (p_I) \in (\mathbb{R} \cup \{-\infty\})^{\binom{n}{k}}$. This vector yields a *tropical linear space* $L(p)$, defined by

$$L(p) = \bigcap_I \{x \in (\mathbb{R}_{\max})^n \mid \max_{i \in I}(p_{I \setminus \{i\}} + x_i) \text{ is achieved at least twice}\} \ ,$$

where the intersection is taken over all subsets of $[n]$ of cardinality $k + 1$.

When $k = n - 1$, $L(p)$ is a called a *tropical hyperplane*. So a tropical hyperplane is a set of vectors of the form

$$\mathcal{H}_a = \{x \in (\mathbb{R} \cup \{-\infty\})^n, \qquad \max_{1 \leqslant i \leqslant n} (a_i + x_i) \text{ is achieved at least twice}\} \ . \tag{1.5}$$

Such a hyperplane is parametrized by the vector $a = (a_1, \ldots, a_n) \in (\mathbb{R} \cup \{-\infty\})^n$, which is required to be non-identically $-\infty$. Tropical hyperplanes are the simplest examples of tropical linear spaces and tropical hypersurfaces [EKL06]. Tropical hyperplanes arise in tropical convexity [CGQ04, DS04], since closed tropical convex sets can be described as intersections of tropical half-spaces. A further motivation for considering tropical hyperplanes, arises from the study of pricing problems: tropical hypersurfaces have been used in [BK19] to represent the influence of prices on the decision of agents buying bundles of elementary products. The "unit demand" case (bundles of cardinality one) is modeled by tropical hyperplanes.

We call *tropical linear regression*, the best approximation of a given set of points by a tropical hyperplane. The tropical linear regression problem is not only of theoretical interest. We shall see in

Chapter 4 that it allows one to quantify the "distance to equilibrium" of a market model, and to infer hidden preferences of a decision maker.

Also, it is a general principle that regression (best approximation) is somehow dual to separation (best classification). Support vector machines (SVMs) are classical learning algorithms used to analyze data for classification. A tropical analogue of SVMs was introduced in [GJ08], and further studied in [TWY20]. Like classical SVMs, a tropical SVM is a discriminative classifier that allows to separate data points into sectors (half-spaces). These half-spaces are defined by the tropical hyperplane maximizing the minimal Hilbert's distance from the data points to itself. In [MCT21], Maragos et al. present applications of tropical geometry to machine learning, including neural networks, graphical models, and nonlinear regression. In particular, they study neural networks with piecewise linear activations from the point of view of tropical geometry. Note that piecewise linear functions correspond to tropical polynomials. They study also the approximation (regression) of data by piecewise linear functions.

Tropical low-rank approximation

Classical low-rank approximation, allows to reduce dimensionality and extract a concise linear structure from a given possibly high dimensional data set. It is commonly used in algorithms in data science. A basic tool for dimension reduction and low-rank approximation is Principal Component Analysis (PCA) [Pea01, Hot33], based on the properties of singular value decomposition (also known as Eckart-Young decomposition) [EY36]. In particular, it provides a best approximation of a given rank, with respect to the Frobenius norm, or statistically speaking in the least squares sense. These approaches are useful when the data in question has a predominantly linear structure. However, in the case of intrinsically non-linear systems, the approximation becomes more challenging. Hence, developing a tropical (max-plus) analogue of low-rank approximation, which provides, with its max-plus structure, non-linear approximation of matrices and more generally of tensors, will be of great interest. Like in the classical case, such tropical approximation is most effective when the data have a compatible structure – i.e., a predominantly tropically linear structure.

An important case in which such a tropical linear structure arises is the numerical solution of optimal control problems. These problems can be solved using Bellman's dynamic programming principle [Bel52]. The latter shows that the value function is the solution of the so called dynamic programming equation. In the continuous space and time case, this equation takes the form of a partial differential equation called the Hamilon-Jacobi-Bellman (HJB) equation [FR12]. To avoid the curse of dimensionality from which suffer the grid based methods in solving HJB equation, *max-plus basis methods* have been developed (see [FM00, McE07, GMQ11, MKH11, Qu14]). In these works, the value function is approximated as a supremum of finitely many basis functions and the supremum is propagated forward in time. Max-plus decomposition methods have allowed to attenuate the curse of dimensionality, for classes of HJB equations [McE07, SMGJ14]. A key issue, in the efficient implementation of max-plus methods, is to select a "basis" of functions with a prescribed cardinality that best approximates a given collection of functions. A version of this is known as the pruning problem [GMQ11, Qu14]. It is equivalent to a problem of low-rank approximation in an infinite dimensional space (of functions). A discrete analogue of the pruning problem has been discussed in [TM19, TTM20].

Other more recent methods allowing to solve the HJB equation are based on tensor decomposition. Dolgov, Kalise and Kunish [DKK21] propose a method based on a classical tensor train approximation for the value function together with a Newton-like iterative method to solve the resulting nonlinear system. Oster, Sallandt and Schneider [OSS19] propose to use low-rank hierarchical tensor product approximation/tree-based tensor formats, together with high-dimensional quadrature, e.g. Monte-Carlo, to solve HJB equation, overcoming computational infeasibility. Recently, deep learning methods were

also used to solve HJB equation by trying to find a feedback control law in the form of a neural network [KGNZ19] in the case of deterministic problems. Although, generating data for training the neural network and validating its accuracy remains challenging.

These recent developments motivate the study of the tropical analogue of low-rank approximation for matrices and tensors. We will show that approximating a matrix by a tropical product of two matrices (having a low-rank) is equivalent to approximately embedding a set of points (the columns of the matrix) by a column space of a matrix, i.e. a tropical cone with a given number of vertices.

Hook proposes in [Hoo17] heuristic algorithms using a local descent method to find low-rank approximation of matrices. In [YZZ19], Yoshida et al. study tropical principal component analysis (PCA) by looking for a tropical polytope with a fixed number of vertices that minimizes the sum of tropical distances (associated to Hilbert's semi-norm) between each data point and its tropical projection into that tropical polytope. They develop a randomized heuristic method to solve this problem with a focus on the special case of tropical polytopes with three vertices. In [PYZ20], Page et al. study also tropical PCA, but with a focus on applying it to dimension reduction over the space of phylogenetic trees. They developed a stochastic optimization method using a Markov Chain Monte Carlo approach to estimate tropical principal components over the space of phylogenetic trees.

Tropical low-rank factorization is closely related to classical nonnegative factorization (since tropical numbers behave as nonnegative numbers). Moreover, it includes as a special case the problem of factorization for Boolean matrices. Deciding whether a Boolean matrix admits a factorization of a given rank is an NP-hard problem, see e.g. [MN20, Sect. 4.1]. Since the Boolean semiring can be embedded in the tropical semiring, tropical low-rank factorization is also NP-hard. Given the difficulty, it is of interest to identity tractable subclasses, and to develop efficient heuristic algorithms for tropical low-rank approximation.

1.2 Contributions

In this thesis, we develop accelerated algorithms for Markov decision processes and more generally for zero-sum stochastic games. We also address best approximation problems arising in tropical geometry. These two fields are closely related. As we will show in the thesis, we can compute classes of tropical best approximation problems by reducing them to solving zero-sum games.

Accelerated value iteration

In Chapter 2, we study the extension of Nesterov's acceleration from gradient descent aiming to minimize a given function to a class of fixed point problems – which cannot be interpreted in terms of convex programming. We study here the affine fixed point problem $x = g + Px$ where the matrix P is non symmetric, possibly not substochastic, with a complex spectrum and a spectral radius of the form $\gamma = 1 - \epsilon$, where ϵ is small. Theorem 2.7, one of our main results, states that a modification of Nesterov's scheme [Nes83] applied to fixed point iteration does converge with an asymptotic rate $1 - \epsilon^{1/2}$ if the spectrum of P is contained in an explicit region of the complex plane. This region is obtained as the image of the disk of radius $1 - \epsilon$ by a rational function of degree 2. We also show that the incorporation of a Krasnosel'skiň-Mann type damping [Man53, Kra55] (see Equation (2.4a)) enlarges the admissible spectrum region of P for acceleration, see Theorem 2.13. Moreover, we introduce a new scheme (2.8), of order $d \geqslant 2$, and show in Theorem 2.19 that it leads to a multiply accelerated asymptotic rate of $1 - \epsilon^{1/d}$, but under a more demanding condition on the spectrum of P, see Theorem 2.19. This theorem also shows that this condition is tight. However, slightly more flexible conditions suffice to guarantee a rate of $1 - \Omega(\epsilon^{1/d})$, as shown by Theorem 2.27.

We subsequently apply the proposed schemes and theoretical results, concerning the affine "0-player case" , to solve non-linear fixed point problems arising from Markov decision processes. We use policy iteration, which allows a reduction to a sequence of affine fixed point problems, still benefiting of acceleration for the solution of each affine problem. This leads to an accelerated policy iteration algorithm (see Algorithm 1), which produces an approximate solution with a precision of order $((1+\gamma)\delta+\delta')/(1-\gamma)^2$ where γ is the maximal discount factor, δ is the accuracy of each inner affine problem and δ' is the accuracy of the policy improvement, see Proposition 2.29.

In Section 2.5, we show the performance of the simple and multiple acceleration schemes, on classes of instances in which the spectral conditions for acceleration are met. In Section 2.5.1, we consider a framework of random matrices that shows distributions of eigenvalues [BCC08] that are compatible with the spectral conditions required for the convergence of the simple and multiple acceleration schemes proposed here. In Section 2.5.2, we show the performance of the accelerated schemes in solving a Hamilton-Jacobi-Bellman equation in the case of small drifts. This example illustrates the usefulness of Theorem 2.9 that allows to have a more tolerant accelerable region on the complex plane while still benefiting from an accelerated asymptotic rate of $1-\Omega(\epsilon^{1/2})$.

Deflated value iteration

In Chapter 3, our aim is to develop accelerated value iteration algorithms for well structured huge scale instances of *mean payoff* problems. To do so, we develop further a general method, first introduced in our previous work [AG13], allowing one to reduce a class of mean payoff problems to discounted problems. Our first main results are Theorem 3.5 and Corollary 3.6, which characterize the best contraction rate of the Shapley operator of a zero-sum game, with respect to all possible weighted sup-norms, as the Collatz-Wielandt number of a certain convex order-preserving positively homogeneous map which we call the "Clarke recession function". This is a key ingredient to obtain our subsequent complexity estimates. This is also of independent interest.

Then, we provide in Theorem 3.25 the reduction from mean payoff problems to discounted problems. This applies to the subclass of two-player games in which there is a distinguished state c to which all other states have access, for all policies of the two players. This reduction combines a scaling argument (a combinatorial version of Doob's h-transform arising in the boundary theory of Markov processes [Dyn69]) and a deflation technique: to a mean payoff problem, we associate a discounted problem, with a *state-dependent* discount rate (Theorem 3.25). To compute this reduction, we need first to solve the dynamic programming equation of a stochastic shortest path (one-player) problem, in which a player wants to maximize the expected hitting time to the distinguished state c. We obtain an explicit contraction rate for the reduced problem in terms of the maximal expected hitting times, which appears in our complexity bounds.

This approach leads to a new algorithm to solve the mean payoff problem, that we call *deflated value iteration* (Algorithm 2). This algorithm is based on two steps, the first step is to compute the value of the stochastic shortest path problem above, and the second one is to solve the reduced problem. Both are solved by using value iteration. We also give a complexity bound in Theorem 3.28, and we compare numerically deflated value iteration with the classical relative value iteration in Section 3.6.

This reduction technique, allows us also to propose a sublinear algorithm solving mean payoff stochastic games, obtained as follows. We solve the mean payoff problem by calling twice a variant of the algorithm of Sidford's et al. [SWWY18]: we call first this variant to compute the parameters of the reduction, and we call it a second time to solve the discounted game obtained after the reduction. We also note that the present variant includes an extension of the algorithm of [SWWY18] for one player to the two-player case. However, this extension is an easier matter–the main novelty here is rather the

reduction from the mean payoff problem to the discounted case and the resulting complexity bounds.

Tropical linear regression

In Chapter 4, we show that tropical linear regression is tractable, theoretically, and to some extent, computationally. Our main result is a strong duality theorem, Theorem 4.23, showing that the infimum of the distance of a set of points $\mathcal{V}$ to a tropical hyperplane coincides with the supremum of the radii of Hilbert's balls included in the tropical convex cone generated by the elements of $\mathcal{V}$. This provides optimality certificates which can be interpreted geometrically as collections of n "witness" points among the elements of $\mathcal{V}$. Our approach also entails that tropical linear regression is polynomial-time equivalent to solving mean payoff games.

We subsequently study variants of the tropical linear regression problem, involving in particular the *signed* notion of tropical hyperplane, obtained by requiring the maximum in (1.5) to be achieved by two indices i, j belonging to prescribed disjoint subsets I, J of $\{1, \cdots, n\}$. We also establish a strong duality theorem in this setting, and provide reductions to mean payoff games for these variants.

We finally illustrate tropical linear regression by an application to an auction model. We consider a market governed by an invitation to tender procedure. We suppose that a decision maker selects repeatedly bids made by firms, based on the bid prices, which are ultimately made public (after the decision is taken), and also on other criteria (assessments of the technical quality of each firm or of environmental impact) or influence factors (like bribes). This is a variant of the classical "first-price sealed-bid auction" [Kri02], with a bias induced by the secret preference. Here, we define the market to be at equilibria if for each invitation, there are at least two best offers. Hence, in the simplest model (unit demand), the set of equilibria prices can be represented by a tropical hyperplane. We distinguish two versions of this problem, one in which only the prices are public, and the other, in which the identities of the winners of the successive invitations are also known. In both cases, we show that solving a tropical linear regression problem allows an observer to quantify the distance of such a market to equilibrium, and also to infer secret preference factors. This solves, in the special case of unit-demand, an inverse problem, consisting in identifying the agent preferences and utilities in auction models, like the one of [BK19]. This might be of interest to a regulation authority wishing to quantify anomalies, or to a bidder, who, seeing the history of the market, would wish to determine how much he should have bidded to win a given invitation or to get the best price for an invitation that he won, thus avoiding the "winner's curse".

Tropical low-rank approximation

In Chapter 5, we establish general properties of tropical low-rank approximation, and identify classes of low-rank approximation problems that are efficiently (in particular, polynomial-time) solvable. We first give a geometric interpretation to the low-rank matrix approximation problem, in terms of approximation of a collection of points by a tropical submodule with few generators. Then, we study the tropical low-rank approximation in the case of rank one and rank two. We introduce a notion of outer radius of a column space of a matrix, and we characterize, in Theorem 5.13, the outer radius of a given column space as the eigenvalue of some specific matrix. We show also, in Theorem 5.16, that the tropical best rank-one approximation of a given matrix is equal to half the outer radius of its column space. We provide also a strongly polynomial algorithm that gives a rank-one approximation of 3-way tensors. We provide an algebraic interpretation of the outer radius as a skew singular value. This yields a tropical analogue of a classical result in matrix theory, showing the error in spectral norm of the best rank-one approximation is given by the second singular value. We extend the best tropical rank-one approximation

to the case of kernels. In dimension three, we give a linear-time algorithm that allows to compute the best rank-two approximation, based on signed tropical linear regression.

1.3 Organization of the manuscript

The manuscript is divided in two parts. Part I (Chapters 2 and 3) contains the results on accelerated algorithms for Markov decision processes and zero-sum stochastic games. Part II (Chapters 4 and 5) contains the results on tropical best approximation. Each chapter of this thesis was written as an independent chapter. The reader may read each of them essentially independently of the others.

- In Chapter 2, we propose an accelerated version of value iteration (AVI) allowing to solve affine fixed point problems with non self-adjoint matrices, alongside with an accelerated version of policy iteration (API) for MDP, building on AVI. This acceleration extends Nesterov's accelerated gradient algorithm to a class of fixed point problems – which cannot be interpreted in terms of convex programming.
- In Chapter 3, we develop a deflated version of value iteration to solve the mean payoff version of stochastic games. This method allows one to transform a mean payoff problem to a discounted one under the hypothesis of existence of a distinguished state that is accessible from all other states and under all policies. Combining this deflation method with variance reduction techniques, we derive a sublinear algorithm solving mean payoff stochastic games.
- In Chapter 4, we solve a tropical linear regression problem consisting in finding the best approximation of a set of points by a tropical hyperplane. We show that the value of this regression problem coincides with the maximal radius of a Hilbert's ball included in a tropical polyhedron, and that this problem is polynomial-time equivalent to mean payoff games. We apply these results to an inverse problem from auction theory.
- In Chapter 5, we study a tropical analogue of low-rank approximation for matrices. We establish general properties of tropical low-rank approximation, and identify classes of low-rank approximation problems that are polynomial-time solvable.

Chapter 2 is based on the preprint [AGQS20], accepted pending minor revision for SIAM Matrix Analysis journal (SIMAX). Chapter 3 is an extended version of a CDC conference article [AGQS19]. Chapter 4 is based on the preprint [AGQS21].

Part I

Accelerated algorithms

CHAPTER 2

Multiply Accelerated Value Iteration for Non-Symmetric Affine Fixed Point Problems and application to Markov Decision Processes

Abstract

In this chapter, we analyze a modified version of Nesterov accelerated gradient algorithm, which applies to affine fixed point problems with non self-adjoint matrices, such as the ones appearing in the theory of Markov decision processes with discounted or mean payoff criteria. We characterize the spectra of matrices for which this algorithm does converge with an accelerated asymptotic rate. We also introduce a dth-order algorithm, and show that it yields a multiply accelerated rate under more demanding conditions on the spectrum. We subsequently apply these methods to develop accelerated schemes for non-linear fixed point problems arising from Markov decision processes. This is illustrated by numerical experiments. This chapter is based on the preprint [AGQS20], accepted pending minor revision for SIAM Matrix Analysis journal (SIMAX).

2.1 Introduction

The dynamic programming method reduces optimal control and repeated zero-sum game problems to fixed point problems involving non-linear operators that are order preserving and sup-norm nonexpansive, see [Bel57, Put14] for background. The 0-player case, with a finite number n of states, is already of interest. In this case, the involved operator is $T : \mathbb{R}^n \to \mathbb{R}^n$ of the form $T(x) = g + Px$, where $g = (g_i) \in \mathbb{R}^n$ and $P = (P_{ij}) \in \mathbb{R}^{n\times n}$ is a substochastic matrix, i.e. a matrix with nonnegative entries such that the sum of each row is less than or equal to 1. The scalar g_i is an instantaneous payment received in state i, whereas P_{ij} is the transition probability from i to j. The difference $1 - \sum_j P_{ij}$ is the probability that the process terminates, when in state i. If v is a fixed point of T, the entry v_i yields the expected cost-to-go from the initial state i. More generally, in the one player case (Markov decision processes), one needs to solve a non-linear fixed point problem, described in Section 2.5, in which the operator T is now a supremum of affine operators $x \mapsto g + Px$.

The standard method to obtain the fixed point of T is to compute the sequence $x_k = T(x_{k-1})$, this is known as *value iteration* [Bel57]. In the 0-player case, value iteration has an asymptotic (geometric) convergence rate given by the spectral radius of P. In many applications, this spectral radius is of the form $1 - \epsilon$ where ϵ is *small*. E.g., ϵ may represent a discount rate. We look for *accelerated* fixed point algorithms, with a convergence rate $1 - \Omega(\epsilon^{1/d})$ for some $d \geqslant 2$, i.e. a convergence rate that is smaller than $1 - c\epsilon^{1/d}$ for some constant $c > 0$.

In the special case of 0-player problems with a *symmetric* matrix P, an algorithm with a rate $1 - \Omega(\epsilon^{1/2})$ can be obtained by specializing the accelerated gradient algorithm of Nesterov [Nes83]. The latter algorithm applies to the minimization of a smooth strictly convex function f, which, in the quadratic case, reduces to an affine fixed point problem with a symmetric matrix P, see [FB15]. In contrast, developing accelerated algorithms for problems of non-symmetric type is a challenging question, which has been studied recently in [IH19, GGC19].

We study here the affine fixed point problem $x = g + Px$ where the matrix P is non symmetric, and possibly not substochastic. Theorem 2.7, one of our main results, states that a modification of Nesterov's scheme [Nes83] does converge with an asymptotic rate $1 - \epsilon^{1/2}$ if the spectrum of P is contained in an explicit region of the complex plane, obtained as the image of the disk of radius $1 - \epsilon$ by a rational function of degree 2. We also show that the incorporation of a Krasnosel'skiň-Mann type damping [Man53, Kra55] (see Equation (2.4a)) enlarges the admissible spectrum region of P for acceleration, see Theorem 2.13. Moreover, we introduce a new scheme (2.8), of order $d \geqslant 2$, and show in Theorem 2.19 that it leads to a multiply accelerated asymptotic rate of $1 - \epsilon^{1/d}$, but under a more demanding condition on the spectrum of P, see Theorem 2.19. This theorem also shows that this condition is tight. However, slightly more flexible conditions suffice to guarantee a rate of $1 - \Omega(\epsilon^{1/d})$, as shown by Theorem 2.27.

We subsequently apply the proposed schemes and theoretical results, concerning the affine "0-player case" , to solve non-linear fixed point problems arising from Markov decision processes. We use policy iteration, which allows a reduction to a sequence of affine fixed point problems, still benefiting of acceleration for the solution of each affine problem. This leads to an accelerated policy iteration algorithm (see Algorithm 1), which produces an approximate solution with a precision of order $((1+\gamma)\delta + \delta')/(1-\gamma)^2$ where γ is the maximal discount factor, δ is the accuracy of each inner affine problem and δ' is the accuracy of the policy improvement, see Proposition 2.29.

In Section 2.5, we show the performance of the simple and multiple acceleration schemes, on classes of instances in which the spectral conditions for acceleration are met. In Section 2.5.1, we consider a framework of random matrices that shows distributions of eigenvalues [BCC08] that are compatible with the spectral conditions required for the convergence of the simple and multiple acceleration schemes

proposed here. In Section 2.5.2, we show the performance of the accelerated schemes in solving a Hamilton-Jacobi-Bellman equation in the case of small drifts. This example illustrates the usefulness of Theorem 2.9 that allows to have a more tolerant accelerable region on the complex plane while still benefiting from an accelerated asymptotic rate of $1 - \Omega(\epsilon^{1/2})$.

The recent works [IH19, GGC19] also deal with generalizations of Nesterov's accelerated algorithm to solve fixed point problems. Their theoretical convergence results apply to matrices with a real spectrum, showing that the original choice of parameters for Nesterov's method in the symmetric case still yields an acceleration in this setting. In contrast, we allow a complex spectrum and characterize the region of the complex plane containing spectra of matrices for which the acceleration is valid (see Theorem 2.7 and Theorem 2.13). Also, a main novelty of the present work is the analysis of multiple accelerations (2.8). The idea of applying Nesterov's acceleration to Markov decision processes appeared in [GGC19], in which a considerable experimental speedup is reported on random instances. The algorithm there coincides with one of the algorithms studied here – 2-accelerated value iteration for Markov decision processes. It is an open problem to establish the convergence of this method for large enough classes of Markov decision processes. The characterization of the set of "accelerable" 0-player problems that we provide here explains why this problem is inherently difficult: in the 0-player problem, the convergence conditions are governed by fine spectral properties which have no known non-linear analogue in the one-player case.

Apart from being applied to Markov decision processes, fixed point iteration also includes as a special case the proximal point method [Roc76b], when the mapping T corresponds to the resolvent of a maximal monotone operator. The proximal point method covers a list of pivotal algorithms in optimization such as the proximal gradient descent, the augmented Lagrangian method (ALM) [Roc76a] and the alternating directional method of multipliers (ADMM) [EB92]. The development of accelerated proximal point method has thus attracted a lot of attention [CMY15, AP19, Att21] and a recent paper [Kim19] constructed a new algorithm achieving $\|x_k - T(x_k)\| \leqslant O(1/k)$ through the performance estimation problem (PEP) approach [DT14]. In a more general setting when T is a nonexpansive mapping in a Euclidean norm, a version of Halpern's iteration was recently shown to yield a residual $\|x_k - T(x_k)\| \leqslant O(1/k)$ [Lie21], also via the PEP approach. These results improve over the worst case bound $\|x_k - T(x_k)\| \leqslant O(1/\sqrt{k})$ of the Krasnoselski-Mann's iteration for a nonexpansive mapping (in arbitrary norm) [BB96]. The acceleration results in the above cited works do not overlap with ours as they only apply to nonexpansive mappings in a Euclidean norm. Moreover, in this chapter we consider strictly contractive mapping and thus focus on linear instead of sublinear convergence guarantees.

There is also a large body of literature on (quasi-)Newton type methods for solving nonlinear equations [rFS09, IS14, WN11], which can be naturally employed for solving fixed point problems and yield fast asymptotic convergence rate. It is well-known that such methods converge only when close enough to the solution. Some papers proposed various safe-guard conditions to globalize the convergence [TP19, ZOB20] and do not provide a rate of convergence. We formally characterize the spectrum condition and the faster convergence rate of accelerated value iteration for affine fixed point problem.

This chapter is organized as follows. In Section 2.2 we introduce the accelerated value iteration (AVI) of any degree $d \geqslant 2$. In Section 2.3 we provide a formal analysis of AVI of degree 2. In Section 2.4 we analyze AVI of arbitrary degree $d \geqslant 2$ and also present the application to Markov decision processes. In Section 2.5 we provide numerical experimental results.

2.2 Accelerated Value Iteration

Nesterov proposed in [Nes83, Nes04] to accelerate the gradient descent scheme for the minimization of a μ-strongly convex function $f : \mathbb{R}^n \to \mathbb{R}$ whose gradient is of Lipschitz constant L, by adding an inertial step:

$$x_{k+1} = y_k - h\nabla f(y_k) \ , \tag{2.1a}$$

$$y_{k+1} = x_{k+1} + \alpha(x_{k+1} - x_k) \ , \tag{2.1b}$$

where $0 < h$, and $\alpha \in [0, 1]$ are parameters. Let x_* be the minimizer of f. When $\alpha = 0$, (2.1) reduces to gradient descent. With the step $h = 1/L$, the gradient descent converges linearly with a rate $1 - 2\mu/(L + \mu)$. Indeed, we have $\|x_k - x_*\|^2 \leqslant \big(1 - 2\mu/(L + \mu)\big)^k \|x_0 - x_*\|^2$, and $f(x_k) - f(x_*) \leqslant \frac{L}{2}\big(1 - 2\mu/(L + \mu)\big)^k \|x_0 - x_*\|^2$, for all $k \geqslant 1$, see Theorem 2.1.14 in [Nes04]. Moreover, Theorem 2.2.3, *ibid.*, implies that if we choose

$$\alpha = \frac{1 - \sqrt{\mu/L}}{1 + \sqrt{\mu/L}}, \tag{2.2}$$

still with $h = 1/L$, the scheme (2.1) converges linearly with a rate $1 - \sqrt{\mu/L}$. Indeed, with α given by (2.2), we have $f(x_k) - f(x_*) \leqslant 2(1 - \sqrt{\mu/L})^k (f(x_0) - f(x_*))$ for all $k \geqslant 1$. Note that when the condition number L/μ is large, i.e. $L/\mu \gg 1$, the rate $1 - \sqrt{\mu/L}$ improves over $1 - 2\mu/(L + \mu)$, whence the scheme (2.2) is commonly known as *accelerated gradient descent*.

We consider the fixed point problem for the operator

$$T(x) = g + Px \ . \tag{2.3}$$

Here, we allow the vector x and the matrix P to have complex entries, requiring only the spectral radius of the matrix P to be strictly less than 1. In the application to MDPs, the vector x will be real and the matrix P will be nonnegative. By abuse of notation, we denote by x_* the unique fixed point of T. We study the *Accelerated Value Iteration algorithm (AVI)* for computing a fixed point of the operator T. It makes a Krasnosel'skiň-Mann (KM) type damping of parameter $0 < \beta \leqslant 1$, replacing T by $(1 - \beta)I + \beta T$, followed by a Nesterov acceleration step:

$$x_{k+1} = (1 - \beta)y_k + \beta T(y_k) \ , \tag{2.4a}$$

$$y_{k+1} = x_{k+1} + \alpha(x_{k+1} - x_k) \ . \tag{2.4b}$$

When $\alpha = 0$ and $\beta = 1$, the scheme (2.4) reduces to the standard fixed point iteration algorithm:

$$x_{k+1} = g + Px_k \ . \tag{2.5}$$

When the spectral radius of P is smaller than $1 - \epsilon$ for some $\epsilon \in (0, 1)$, the standard fixed point scheme converges with an asymptotic rate no greater than $1 - \epsilon$ to the unique fixed point, meaning that for any norm $\|\cdot\|$

$$\limsup_{k\to\infty} \|x_k - x_*\|^{1/k} \leqslant 1 - \epsilon.$$

By analogy with accelerated gradient descent, we aim at accelerating the standard fixed point scheme by finding appropriate parameters α and β so that

$$\limsup_{k\to\infty} \|x_k - x_*\|^{1/k} \leqslant 1 - \sqrt{\epsilon}, \tag{2.6}$$

for matrices P with spectral radius bounded by $1 - \epsilon$.

Remark 2.1. If P is symmetric, the iteration (2.4) can be recovered by applying the accelerated gradient descent scheme (2.1) to the quadratic function $f(x) \equiv \frac{1}{2}x^\top(I-P)x - g^\top x$. The damping parameter β corresponds to the step h. However, Nesterov's results only apply to the case when f is a strongly convex function. This requires in particular $I-P$ to be symmetric positive definite. In particular all the eigenvalues of P must be real and smaller than 1.

The scheme (2.4) for fixed point iteration has been considered recently by [IH19, GGC19]. Moreover, inspired by the momentum method [Pol64, GFJ15] for improving gradient descent, [GGC19] also proposed a momentum fixed point method described as follows:

$$x_{k+1} = (1-\beta)x_k + \beta T(x_k) + \alpha(x_k - x_{k-1}). \tag{2.7}$$

Asymptotic rate analysis for (2.4) (2.7) follows from [GGC19] when the spectrum of P is real.

As discussed in the introduction, our main results apply to complex spectra, and also to higher degree of acceleration.

In the scheme (2.4), y_{k+1} is generated from a linear combination of the last two iterates. We now consider the following *Accelerated Value Iteration of degree* d *(dA-VI)*, in which y_{k+1} is a linear combination of the last d iterates for any $d \geqslant 2$,

$$x_{k+1} = (1-\beta)y_k + \beta T(y_k) \ , \tag{2.8a}$$

$$y_{k+1} = (1+\alpha_{d-2}+\cdots+\alpha_0)x_{k+1} - \alpha_{d-2}x_k - \cdots - \alpha_0 x_{k-d+2} \ . \tag{2.8b}$$

We will show how to select the parameters $\alpha = (\alpha_0, \cdots, \alpha_{d-2})$ to obtain an acceleration of order d, in the sense that

$$\limsup_{k\to\infty} \|x_k - x_*\|^{1/k} \leqslant 1 - \epsilon^{1/d}. \tag{2.9}$$

Remark 2.2. The idea of accelerating the vanilla KM fixed point method by extrapolating a finite number of previous steps goes back to the work of Anderson in 1965 [And65]. The algorithm known as Anderson Acceleration (AA) chooses dynamically the extrapolation coefficients, while the coefficients $\alpha = (\alpha_0, \cdots, \alpha_{d-2})$ in dA-VI (2.8) remain constant for all the iterations. The theoretical analysis of AA and of its variants is still under development. In particular, the theoretical convergence rate of AA seems to be missing in the literature, except in the special case when T corresponds to the gradient descent mapping of a strongly convex and smooth function [SBD17]. When T takes the form of (2.3), this requires P to be symmetric, see Remark 2.1. In [ZOB20], a modified AA, interleaving KM updates by using safe-guarding steps, is shown to be globally converging, but the convergence rate is not analyzed. As shown later, the d-AVI (2.8) does not need any safe-guard checking and will converge with accelerated asymptotic rate as in (2.9) under some conditions on the spectrum of P.

Remark 2.3. The computational cost of one iteration of the classical Value Iteration algorithm (2.5) is $O(n^2)$. In comparison, the computational cost of one iteration of the dA-VI algorithm (2.8) is $O(n(n+d))$. Regarding the space complexity, the classical Value Iteration needs to store two vectors (x_{k+1}, x_k) each of size n, so it needs a $2n$ space of memory. In comparison, the dA-VI algorithm needs to store $d+1$ vectors $(y_{k+1}, x_{k+1}, \cdots, x_{k-d+2})$ each of size n, so it needs $(d+1)n$ space of memory. We notice that in practice the degree d that we will use is small ($\leqslant 4$), therefore the computational cost of one iteration of dA-VI and its space complexity are similar to the ones of the classical Value Iteration algorithm. Moreover, the asymptotic convergence rate $1-\epsilon^{1/d}$ allows the dA-VI algorithm to converge in a number of iterations smaller than the Value Iteration algorithm (see the numerical experiments in Section 2.5).

2.3 Analysis of Accelerated Value Iteration of degree 2

In this section, we analyse the AVI scheme (2.4). We will show that with an appropriate choice of the acceleration parameter α, and under an assumption on the shape of the complex spectrum of P, the asymptotic rate can indeed be improved up to $1-\sqrt{\epsilon}$. We also show that the damping parameter β will allow us to enlarge the convergence region, while keeping the acceleration properties. We deal separately with the special $d = 2$ case, since it is more elementary, easier to compare with existing acceleration schemes, and since it gives insight on the generalization to the higher degree case which will be done in Section 2.4.

2.3.1 The spectrum of the AVI iteration

We define $P_\beta := (1-\beta)I + \beta P$. Then, the AVI algorithm (2.4) can be written as the second order iteration

$$y_{k+1} = \beta g + (1+\alpha)P_\beta y_k - \alpha P_\beta y_{k-1}. \tag{2.10}$$

Considering $z_k = y_k - x_*$, the iteration becomes $z_{k+1} = (1+\alpha)P_\beta z_k - \alpha P_\beta z_{k-1}$. This is equivalent to:

$$\begin{pmatrix} z_{k+1} \\ z_k \end{pmatrix} = \begin{pmatrix} (1+\alpha)P_\beta & -\alpha P_\beta \\ I & 0 \end{pmatrix} \begin{pmatrix} z_k \\ z_{k-1} \end{pmatrix} \tag{2.11}$$

Without loss of generality we first deal with the case with no damping, i.e., $\beta = 1$. The discussion for general $\beta \in (0,1]$ can be found in Section 2.3.2 . Then, the matrix appearing in (2.11) becomes

$$Q_\alpha := \begin{pmatrix} (1+\alpha)P & -\alpha P \\ I & 0 \end{pmatrix} \quad . \tag{2.12}$$

The asymptotic rate of convergence of the sequence (z_k) in the system (2.11), when it is converging, and thus of the sequence (y_k) in the AVI scheme (2.4) is determined by the spectral radius of Q_α. Recall that we want to improve this asymptotic rate, thus it suffices to find appropriate values of α such that the spectral radius of Q_α is as small as possible.

We first relate the eigenvalues of Q_α with those of P. We introduce the following rational function of degree 2, defined on $\mathbb{C} \setminus \{\alpha/(1+\alpha)\}$ by

$$\phi_\alpha(z) := \frac{z^2}{(1+\alpha)z - \alpha} \quad .$$

The following is a standard property of block-companion matrices, we provide the proof for completeness.

Lemma 2.4. *If $\alpha \neq 0$ then λ is an eigenvalue of Q_α if and only if there exists an eigenvalue δ of P such that $\delta = \phi_\alpha(\lambda)$. In other words,*

$$\operatorname{spec} Q_\alpha = \phi_\alpha^{-1}(\operatorname{spec} P).$$

Proof. Let λ be an eigenvalue of Q_α. There exists a non-zero vector $\left(\begin{smallmatrix} z_1 \\ z_0 \end{smallmatrix}\right) \in \mathbb{R}^{2n}$ such that $Q_\alpha \left(\begin{smallmatrix} z_1 \\ z_0 \end{smallmatrix}\right) = \lambda \left(\begin{smallmatrix} z_1 \\ z_0 \end{smallmatrix}\right)$. This is equivalent to $(1+\alpha)Pz_1 - \alpha P z_0 = \lambda z_1$ and $z_1 = \lambda z_0$, or equivalently $(\lambda(1+\alpha) - \alpha)Pz_0 = \lambda^2 z_0$ and $z_1 = \lambda z_0$. We have $z_0 \neq 0$, because otherwise $z_1 = \lambda z_0 = 0$. We notice that $\lambda(1+\alpha) - \alpha \neq 0$, because otherwise $\lambda = \frac{\alpha}{1+\alpha} \neq 0$ and $\lambda^2 z_0 = 0$, then $z_0 = 0$, which is not true. Therefore $Pz_0 = \phi_\alpha(\lambda) z_0$ which allows to conclude. $\square$

2.3.1.a The case of real eigenvalues We now explain how to select α optimally. We first suppose that the spectrum of P is real and nonnegative, i.e., $\operatorname{spec} P \subset [0, 1-\epsilon]$ for some $\epsilon \in (0,1)$. We denote by $\mathcal{B}(z,r)$ the closed disk of the complex plane with center z and radius r. We consider the minimax problem

$$\min_{\alpha>0} \max_{P:\operatorname{spec} P\subset[0,1-\epsilon]} \rho(Q_\alpha) \tag{2.13}$$

where ρ denotes the spectral radius, and the matrix Q_α, depending on P, is defined by (2.12).

Lemma 2.5. *The solution α^* of the minimax problem* (2.13) *is given by*

$$\alpha^* = \frac{1-\sqrt{\epsilon}}{1+\sqrt{\epsilon}}. \tag{2.14}$$

It guarantees that $\operatorname{spec} Q_{\alpha^*} \subset \mathcal{B}(0, 1-\sqrt{\epsilon})$*, for all matrices P such that* $\operatorname{spec} P \subset [0, 1-\epsilon]$.

Proof. By Lemma 2.4, $\lambda \in \operatorname{spec} Q_\alpha$ if and only if there exists $\delta \in \operatorname{spec} P \subset [0, 1-\epsilon]$, such that $\delta = \frac{\lambda^2}{(1+\alpha)\lambda-\alpha}$. This can be written as a second degree equation in λ:

$$\lambda^2 - (1+\alpha)\delta\lambda + \alpha\delta = 0. \tag{2.15}$$

The discriminant of this equation is $\Delta = \delta^2(1+\alpha)^2 - 4\alpha\delta = \delta(1+\alpha)^2(\delta-\alpha')$, where $\alpha' := \frac{4\alpha}{(1+\alpha)^2}$. We note that the function $\alpha \mapsto \frac{4\alpha}{(1+\alpha)^2}$ is a strictly increasing bijection from $[0,1]$ to itself, with inverse function $\alpha \mapsto \frac{1-\sqrt{1-\alpha}}{1+\sqrt{1-\alpha}}$. Hence $\alpha' \geqslant 1-\epsilon$ if and only if $\alpha \geqslant \alpha^*$.

Claim 1. For fixed α, the maximal modulus of the solutions of (2.15) is increasing with δ.

Proof of Claim 1. If $\Delta \leqslant 0$, i.e. $\delta \leqslant \alpha'$, then the solutions of (2.15) are complex conjugate $\lambda_\pm = \frac{1}{2}(\delta(1+\alpha) \pm i\sqrt{\delta(1+\alpha)^2(\alpha'-\delta)})$, and we have $\lambda_+\lambda_- = \alpha\delta$. Then $|\lambda_+| = |\lambda_-| = \sqrt{\alpha\delta}$, which is increasing in δ.

If $\Delta \geqslant 0$, i.e. $\delta \geqslant \alpha'$, the solutions of (2.15) are real:

$$\lambda_\pm = \frac{1}{2}(\delta(1+\alpha) \pm \sqrt{\delta(1+\alpha)^2(\delta-\alpha')}).$$

Then $\max(|\lambda_+|, |\lambda_-|) = \frac{1}{2}(\delta(1+\alpha) + \sqrt{\delta(1+\alpha)^2(\delta-\alpha')})$ is strictly increasing in δ. □

Claim 1 shows that

$$\max_{P:\operatorname{spec} P\subset[0,1-\epsilon]} \rho(Q_\alpha) = \max\left\{|\lambda| : \lambda^2 - (1+\alpha)(1-\epsilon)\lambda + \alpha(1-\epsilon) = 0.\right\} \tag{2.16}$$

The discriminant of the second order equation in (2.16) is

$$\Delta = (1-\epsilon)^2(1+\alpha)^2 - 4\alpha(1-\epsilon) = (1-\epsilon)(1+\alpha)^2(1-\epsilon-\alpha').$$

If $\alpha \geqslant \alpha^*$, then $\alpha' \geqslant 1-\epsilon$ and $\Delta \leqslant 0$. In this case $|\lambda_+| = |\lambda_-| = \sqrt{\alpha(1-\epsilon)}$ is increasing in $\alpha \in \left[\frac{1-\sqrt{\epsilon}}{1+\sqrt{\epsilon}}, 1\right]$. If $\alpha \leqslant \alpha^*$, then $\alpha' \leqslant 1-\epsilon$ and $\Delta \geqslant 0$. In this case

$$\max(|\lambda_+|, |\lambda_-|) = \frac{1}{2}\left((1-\epsilon)(1+\alpha) + \sqrt{(1-\epsilon)^2(1+\alpha)^2 - 4\alpha(1-\epsilon)}\right).$$

Claim 2. The function $F : \alpha \to (1-\epsilon)(1+\alpha) + \sqrt{(1-\epsilon)^2(1+\alpha)^2 - 4\alpha(1-\epsilon)}$ is strictly decreasing on $\left[0, \frac{1-\sqrt{\epsilon}}{1+\sqrt{\epsilon}}\right]$.

Proof of Claim 2. We have

$$F'(\alpha) = 1 - \epsilon + \frac{2(1-\epsilon)^2(1+\alpha) - 4(1-\epsilon)}{2\sqrt{(1-\epsilon)^2(1+\alpha)^2 - 4\alpha(1-\epsilon)}} = \frac{(1-\epsilon)h(\alpha)}{\sqrt{(1-\epsilon)^2(1+\alpha)^2 - 4\alpha(1-\epsilon)}},$$

where $h(\alpha) = \sqrt{(1-\epsilon)^2(1+\alpha)^2 - 4\alpha(1-\epsilon)} + (1-\epsilon)(1+\alpha) - 2$. It is easy to check that $h(\alpha) = \sqrt{(2-(1-\epsilon)(1+\alpha))^2 - 4\epsilon} - (2-(1-\epsilon)(1+\alpha)) < 0$ for all $\alpha \in [0, \frac{1-\sqrt{\epsilon}}{1+\sqrt{\epsilon}}]$. Since $2-(1-\epsilon)(1+\alpha) \geqslant 0$ for all $\alpha \in \left[0, \frac{1-\sqrt{\epsilon}}{1+\sqrt{\epsilon}}\right]$, we deduce that $h(\alpha) < 0$ and hence $F'(\alpha) < 0$ for all $\alpha \in \left[0, \frac{1-\sqrt{\epsilon}}{1+\sqrt{\epsilon}}\right]$. □

We conclude that the best choice of α which minimizes the maximum of the spectral radius of Q_α corresponding to all P with spectrum in $[0, 1-\epsilon]$ is α^* given in (2.14), and it allows to have $\operatorname{spec} Q_{\alpha^*} \subset \mathcal{B}(0, 1-\sqrt{\epsilon})$ for all such matrix P. □

Remark 2.6. If P is symmetric, then the quadratic function f in Remark 2.1 is a strongly convex function with $L = 1$ and $\mu = \epsilon$. In this special case the α^* in Lemma 2.5 coincides with the inertial parameter (2.2) in Nesterov's constant-step method. The same choice of step has been proposed, for nonsymmetric matrices with real spectrum, in [IH19].

2.3.1.b The case of complex eigenvalues Now we do not assume anymore that P has a real spectrum. We will show that the best acceleration rate achievable in the case of a real spectrum, obtained by choosing $\alpha = \alpha^*$ as in Lemma 2.5, is still achievable in the case of a complex spectrum satisfying a geometric condition.

Consider the following simple closed curve Γ_ϵ defined by the parametric equation:

$$\theta \mapsto \frac{(1-\epsilon)e^{2i\theta}}{2e^{i\theta} - 1} \quad , \quad \theta \in (0, 2\pi] \ .$$

Denote by Σ_ϵ the compact set delimited by the curve Γ_ϵ. We show in Figure 2.1 the curve Γ_0 and the enclosed region Σ_0. It is easy to see that Γ_ϵ (resp. Σ_ϵ) is a scaling of Γ_0 (resp. Σ_0) by $1-\epsilon$. Moreover, we have

$$\left| \frac{e^{2i\theta}}{2e^{i\theta} - 1} \right| = \frac{1}{|2e^{i\theta} - 1|} \leqslant \frac{1}{|2e^{i\theta}| - 1} = 1, \tag{2.17}$$

and thus the curve Γ_ϵ is included in the disk $\mathcal{B}(0, 1-\epsilon)$. It follows that

$$\Sigma_\epsilon \subset \mathcal{B}(0, 1-\epsilon). \tag{2.18}$$

Theorem 2.7. *Let $\epsilon \in (0,1)$, P be a $n \times n$ complex matrix and Q_α be defined as in* (2.12) *with $\alpha = (1-\sqrt{\epsilon})/(1+\sqrt{\epsilon})$. If $\operatorname{spec} P \subset \Sigma_\epsilon$, then $\operatorname{spec} Q_\alpha \subset \mathcal{B}(0, 1-\sqrt{\epsilon})$.*

Proof. To show that $\operatorname{spec} P \subset \Sigma_\epsilon \Rightarrow \operatorname{spec} Q_\alpha \subset \mathcal{B}\left(0, 1-\sqrt{\epsilon}\right)$, we will prove the contrapositive

$$\operatorname{spec} Q_\alpha \cap \mathbb{C} \setminus \mathcal{B}\left(0, 1-\sqrt{\epsilon}\right) \neq \emptyset \Rightarrow \operatorname{spec} P \cap \mathbb{C} \setminus \Sigma_\epsilon \neq \emptyset. \tag{2.19}$$

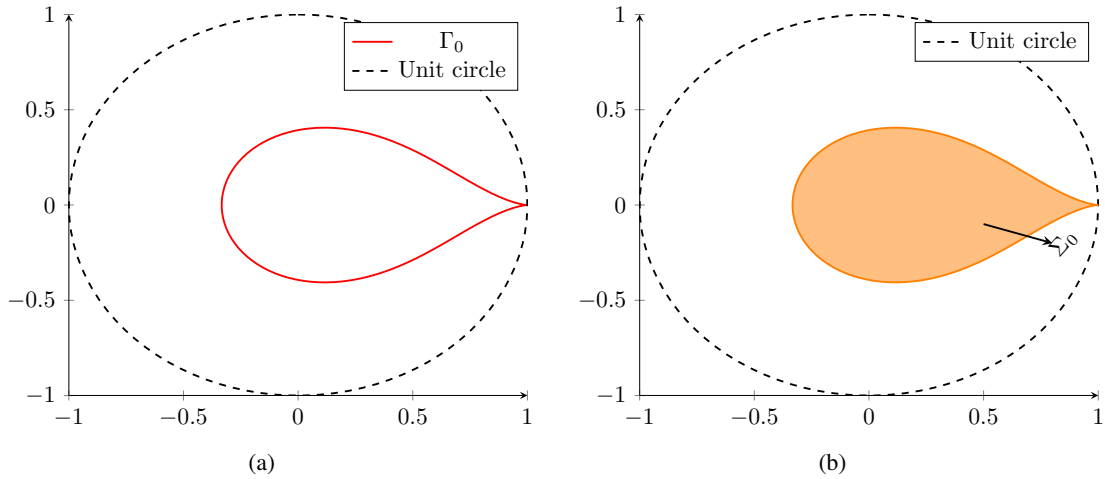


Figure (2.1) Illustration of the curve Γ_0 (Figure 2.1(a)) and its enclosed region Σ_0 (Figure 2.1(b)).

We consider an eigenvalue $\lambda \in \operatorname{spec} Q_\alpha \cap \mathbb{C} \setminus \mathcal{B}(0, 1-\sqrt{\epsilon})$ so that $\lambda = r(1-\sqrt{\epsilon})e^{i\bar{\theta}}$ for some $\bar{\theta} \in (0, 2\pi]$ and $r > 1$. The associated eigenvalue of P is

$$\delta_r(\bar{\theta}) := \frac{\lambda^2}{(1+\alpha)\lambda - \alpha} = \frac{(1-\sqrt{\epsilon})^2 r^2 e^{2i\bar{\theta}}}{\frac{2}{1+\sqrt{\epsilon}} r(1-\sqrt{\epsilon})e^{i\bar{\theta}} - \frac{1-\sqrt{\epsilon}}{1+\sqrt{\epsilon}}} = \frac{(1-\epsilon)r^2 e^{2i\bar{\theta}}}{2re^{i\bar{\theta}} - 1}.$$

It is easy to check from $r > 1$ that

$$|\delta_r(0)| = \frac{(1-\epsilon)r^2}{2r-1} > 1-\epsilon,$$

which together with (2.18) implies that

$$\delta_r(0) \notin \Sigma_\epsilon.$$

Suppose that $\delta_r(\bar{\theta}) \in \Sigma_\epsilon$. Since the curve Γ_ϵ is the boundary of the compact set Σ_ϵ, there must be a $\theta \in (0, 2\pi)$ such that

$$\delta_r(\theta) \in \Gamma_\epsilon.$$

In other words, there is $u, v \in \mathbb{C}$ such that $|u| = |v| = 1$ and $(1-\epsilon)\frac{u^2}{2u-1} = (1-\epsilon)\frac{r^2v^2}{2rv-1}$. Then $r^2(2u-1)v^2 - 2ru^2v + u^2 = 0$. We consider v as the unknown variable in this equation. The discriminant is $\Delta = 4r^2u^2(u-1)^2$ and then

$$v \in \left\{ \frac{2ru^2 \pm 2ru(u-1)}{2r^2(2u-1)} \right\} = \left\{ \frac{u}{r}, \frac{u}{r(2u-1)} \right\}.$$

Since $|u| = |v|$, it is impossible that $v = \frac{u}{r}$. If $v = \frac{u}{r(2u-1)}$, then by taking the module we have $|2u-1| = \frac{1}{r}$, which is absurd because $|2u-1| \geqslant |2u| - 1 = 1 > \frac{1}{r}$. We thus conclude that $\delta_r(\bar{\theta}) \in \mathbb{C} \setminus \Sigma_\epsilon$ and (2.19) is proved. □

Remark 2.8. In Theorem 2.19, we will give an analysis for acceleration of arbitrary degree $d \geqslant 2$, which recovers Theorem 2.7 for the case $d = 2$ and in addition shows that if $\operatorname{spec} Q_\alpha \subset \mathcal{B}(0, 1-\sqrt{\epsilon})$, then $\operatorname{spec} P \subset \Sigma_\epsilon$.

For any $r \geqslant 1$, denote by $\Gamma_\epsilon(r)$ the simple closed curve defined by the parametric equation $\theta \mapsto \delta_r(\theta), \theta \in (0, 2\pi]$, and denote by $\Sigma_\epsilon(r)$ the region enclosed by $\Gamma_\epsilon(r)$. We have the following stronger result.

Theorem 2.9. *Let $\epsilon \in (0,1)$, P be a $n \times n$ complex matrix, Q_α be defined as in* (2.12) *with $\alpha = (1-\sqrt{\epsilon})/(1+\sqrt{\epsilon})$ and $r \geqslant 1$. If $\operatorname{spec} P \subset \Sigma_\epsilon(r)$, then $\operatorname{spec} Q_\alpha \subset \mathcal{B}(0, r(1-\sqrt{\epsilon}))$.*

An ingredient of the proof of Theorem 2.7 was to show that for any $r > 1$, the curve $\Gamma_\epsilon(r)$ does not intersect with the curve Γ_ϵ. In a similar way, we can prove the above Theorem 2.9, by showing that the curve $\Gamma_\epsilon(r)$ does not intersect with the curve $\Gamma_\epsilon(r')$ for any distinct $r \geqslant 1$ and $r' \geqslant 1$.

Remark 2.10. Let $0 < \gamma \leqslant 1$. An equivalent statement of Theorem 2.9 is as follows: if $\operatorname{spec} P \subset \Sigma_\epsilon\left(\frac{1-\gamma\sqrt{\epsilon}}{1-\sqrt{\epsilon}}\right)$, then $\operatorname{spec} Q_\alpha \subset \mathcal{B}(0, 1-\gamma\sqrt{\epsilon})$. This implies that the asymptotic rate can be of order $1 - \Omega(\sqrt{\epsilon})$ if the spectrum of P is sufficiently close to Σ_ϵ. We illustrate this result in Figure 2.2 with the example of $\epsilon = 0.01$ and $\gamma = 0.5$.

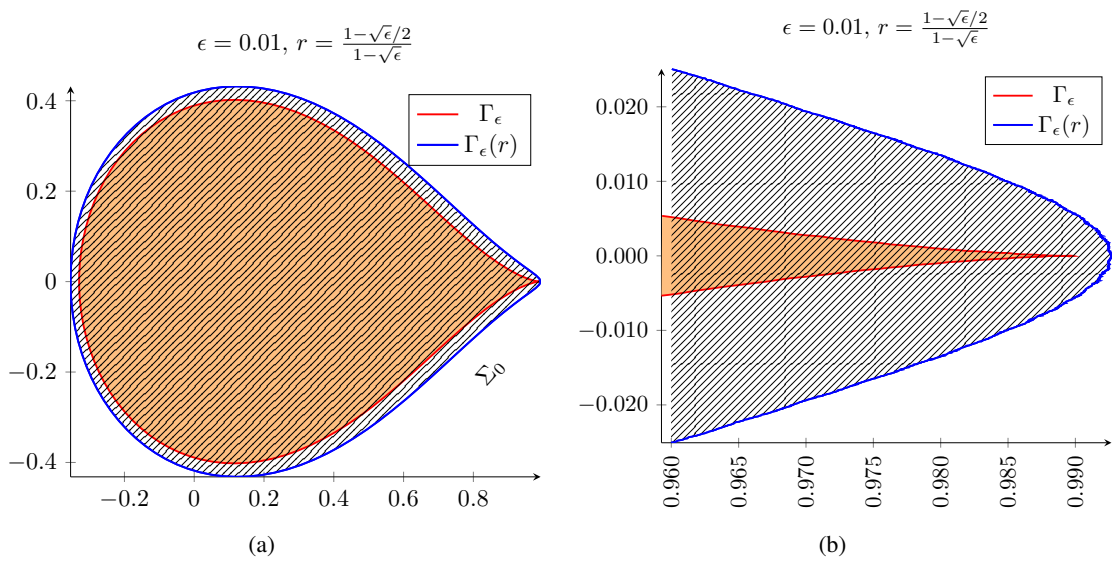


Figure (2.2) Illustration of the curve Γ_ϵ (the curve in red) and its enclosed region Σ_ϵ (the region in orange), and the curve $\Gamma_\epsilon(r)$ (the curve in blue) and its enclosed region $\Sigma_\epsilon(r)$ (the dashed region). Figure 2.2(a) is a zoom of Figure 2.2(b).

2.3.2 Enlargement of the accelerable region by damping

In this subsection we consider the effect of the Krasnosel'skiĭ-Mann damping parameter $\beta \in (0, 1]$ The following corollary, which is immediate from Theorem 2.7, determines the accelerable region for the spectrum of the initial matrix P.

Corollary 2.11. *If there is $\beta \in (0,1]$ such that $\operatorname{spec} P_\beta \subset \Sigma_\epsilon$, then AVI algorithm* (2.4) *with the parameters β and $\alpha = \frac{1-\sqrt{\epsilon}}{1+\sqrt{\epsilon}}$ converges with an asymptotic rate no greater than $1-\sqrt{\epsilon}$, i.e.,* (2.6) *holds.*

Based on Corollary 2.11, we now look for a radius $r > 0$ such that if $\operatorname{spec} P \subset \mathcal{B}(0, r) \cup [-1+\epsilon, 1-\epsilon]$, then there is a scaling parameter $\beta \in (0, 1]$ such that $\operatorname{spec} P_\beta \subset \Sigma_{\epsilon'}$, for some $\epsilon' > 0$, with the goal of achieving an accelerated asymptotic rate $1 - \Omega(\sqrt{\epsilon})$.

We start by giving a disk and a part of the real line which are contained in Σ_0.

Lemma 2.12. *We have*

$$\mathcal{B}\left(\frac{1}{3},\frac{1}{3}\right)\cup\left[-\frac{1}{3},1\right]\subset\Sigma_0.$$

Proof. The boundary of Σ_0 intersects with the real axis at $(1,0)$ and $(-1/3,0)$. Thus $\left[-\frac{1}{3},1\right]\subset\Sigma_0$. For any $\theta\in(0,2\pi]$, we have

$$\begin{aligned}
&\left|\frac{e^{2i\theta}}{2e^{i\theta}-1}-\frac{1}{3}\right|^2-\frac{1}{9}=\left|\frac{3e^{2i\theta}-2e^{i\theta}+1}{3(2e^{i\theta}-1)}\right|^2-\frac{1}{9}\\
&=\frac{|3e^{2i\theta}-2e^{i\theta}+1|^2-|2e^{i\theta}-1|^2}{9|2e^{i\theta}-1|^2}\\
&=\frac{14-16\cos(\theta)+6\cos(2\theta)-5+4\cos(\theta)}{9|2e^{i\theta}-1|^2}\\
&=\frac{4\cos^2(\theta)-4\cos(\theta)+1}{9|2e^{i\theta}-1|^2}=\frac{(2\cos(\theta)-1)^2}{9|2e^{i\theta}-1|^2}\geqslant 0.
\end{aligned}$$

Thus the boundary of Σ_0 does not intersect the interior of the disk $\mathcal{B}(\frac{1}{3},\frac{1}{3})$. Since $0\in\Sigma_0\cap\mathcal{B}(\frac{1}{3},\frac{1}{3})$, the disk $\mathcal{B}(\frac{1}{3},\frac{1}{3})$ is entirely contained in Σ_0. □

The following result shows that if the spectrum of the initial matrix P belongs to a "flying saucer" shaped region of the complex plane (see Figure 2.3(a) for illustration), the AVI algorithm does converge with an asymptotic rate $1-\Omega(\sqrt{\epsilon})$.

Theorem 2.13. *If* $\operatorname{spec}P\subset\mathcal{B}(0,\frac{1-\epsilon}{2})\cup[-1+\epsilon,1-\epsilon]$*, then by setting* $\beta=\frac{2}{3-\epsilon}$ *and*

$$\alpha=\frac{1-\sqrt{2\epsilon/(3-\epsilon)}}{1+\sqrt{2\epsilon/(3-\epsilon)}},\tag{2.20}$$

the iterates of algorithm AVI (2.4) *satisfy*

$$\limsup_{k\to\infty}\|x_k-x_*\|^{1/k}\leqslant 1-\sqrt{\frac{2\epsilon}{3}}.$$

Proof. For any $\beta\in(0,1]$, the spectrum of P_β is the image of the spectrum of P by the homothety $H^\beta:=z\mapsto 1-\beta+\beta z$ of center 1 and ratio β. Note that $\beta=\frac{2}{3-\epsilon}$ satisfies

$$1-\beta=\frac{\beta(1-\epsilon)}{2}=\frac{1-\beta\epsilon}{3}.$$

Hence the image of $\mathcal{B}(0,\frac{1-\epsilon}{2})\cup[-1+\epsilon,1-\epsilon]$ by the homothety H^β is

$$\mathcal{B}\left(\frac{1-\beta\epsilon}{3},\frac{1-\beta\epsilon}{3}\right)\cup\left[-\frac{1-\beta\epsilon}{3},1-\beta\epsilon\right].$$

See Figure 2.3(b) for an illustration. In view of Lemma 2.12, this region is contained in $\Sigma_{\beta\epsilon}$. It follows that $\operatorname{spec}P_\beta\subset\Sigma_{\beta\epsilon}$ and the statement follows by applying Corollary 2.11. □

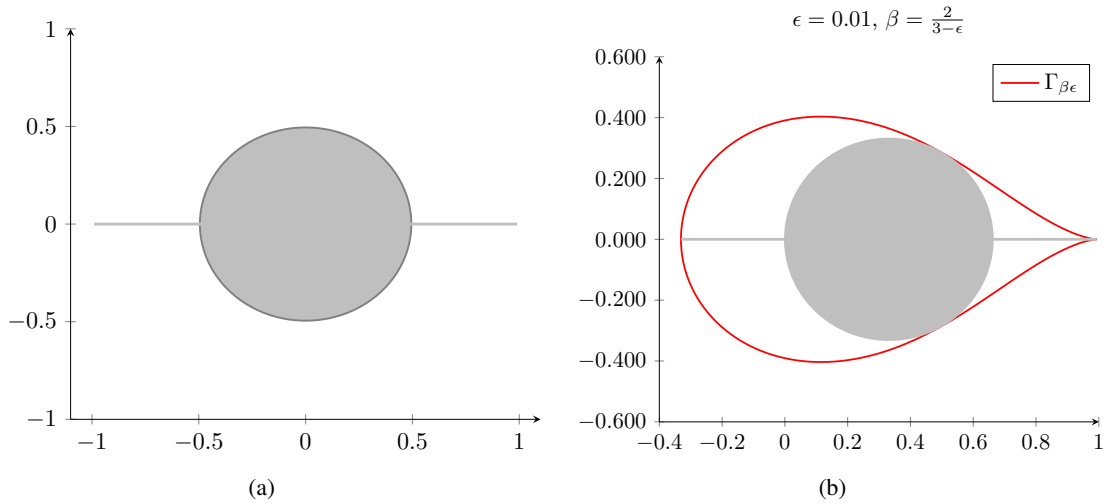


Figure (2.3) (a): the region $\mathcal{B}\left(0, \frac{1-\epsilon}{2}\right) \cup [-1+\epsilon, 1-\epsilon]$ (the flying saucer shaped region in grey). (b): the curve $\Sigma_{\beta\epsilon}$ (the curve in red) and the image of $\mathcal{B}\left(0, \frac{1-\epsilon}{2}\right) \cup [-1+\epsilon, 1-\epsilon]$ by the homothety H^β.

Remark 2.14. For $0 < \epsilon < \frac{1}{3}$, the flying saucer shaped region $\mathcal{B}\left(0, \frac{1-\epsilon}{2}\right) \cup [-1+\epsilon, 1-\epsilon]$ can not be included in Σ_0 and Corollary 2.11 is not applicable. However, the homothety H^β with $\beta = \frac{2}{3-\epsilon}$ sends this region inside $\Sigma_{\beta\epsilon}$, whence an accelerated asymptotic rate.

Remark 2.15. In the special case when $\operatorname{spec} P \subset [-1+\epsilon, 1-\epsilon]$, a similar result was established in [GGC19]. Translated with our notations, Theorem 5.1 in [GGC19] proved an asymptotic rate $1 - \sqrt{\epsilon/(2-\epsilon)}$ by setting $\alpha = 1/(2-\epsilon)$ and $\beta = \frac{1-\sqrt{\epsilon/(2-\epsilon)}}{1+\sqrt{\epsilon/(2-\epsilon)}}$ in (2.4).

We complement Theorem 2.13 by showing the optimality of the radius $\frac{1-\epsilon}{2}$ in the sense described by the following lemma.

Lemma 2.16. *The largest radius $r \geqslant 0$, for which there exists $\beta \in (0,1]$ such that $H^\beta(\mathcal{B}(0,r)) \subset \Sigma_0$, is $r = \frac{1}{2}$ and it corresponds to the choice $\beta = \frac{2}{3}$.*

Proof. Applying the homothety H^β to $\mathcal{B}(0,r)$ leads to the ball $\mathcal{B}(1-\beta, \beta r)$. We thus look for the largest r such that $\mathcal{B}(1-\beta, \beta r) \subset \Sigma_0$ for some $\beta \in (0,1)$. We notice that $r > 1$ is not possible, because for any $r > 1$ we have $1 + \beta(r-1) > 1$, which is outside Σ_0.

Now, we suppose that $0 \leqslant r \leqslant 1$ and there is $\beta \in (0,1)$ such that $\mathcal{B}(1-\beta, \beta r) \subset \Sigma_0$. We consider the line (D_r) of the complex plane passing through the point of coordinates $(1,0)$ and tangent to the upper half of the circle $\mathcal{B}(0,r)$. This line is given by the equation

$$y = \frac{r}{\sqrt{1-r^2}}(1-x).$$

Note that (D_r) is invariant by the homothety H^β and thus is also tangent to $\mathcal{B}(1-\beta, \beta r)$. Thus (D_r) must intersect with Σ_0 at a point other than $(1,0)$, see Figure 2.4 for an illustration. The curve Γ_0 is given by

$$\theta \mapsto \frac{e^{2i\theta}}{2e^{i\theta}-1} = \frac{2\cos(\theta) - \cos(2\theta)}{5 - 4\cos(\theta)} + i\frac{2\sin(\theta) - \sin(2\theta)}{5 - 4\cos(\theta)}, \quad \theta \in (0, 2\pi].$$

Let $\theta \in (0, 2\pi)$ such that $(x_0, y_0) := \left(\frac{2\cos(\theta)-\cos(2\theta)}{5-4\cos(\theta)}, \frac{2\sin(\theta)-\sin(2\theta)}{5-4\cos(\theta)}\right)$ lies in $(D_r) \cap \Sigma_0$. Then

$$\frac{r}{\sqrt{1-r^2}} = \frac{y_0}{1-x_0} = \frac{2\sin(\theta)(1-\cos(\theta))}{5-6\cos(\theta)+2\cos^2(\theta)-1} = \frac{\sin(\theta)}{2-\cos(\theta)} \quad .$$

We can easily prove that:

$$\max_{\theta\in(0,2\pi)} \frac{\sin(\theta)}{2-\cos(\theta)} = \frac{\sqrt{3}}{3} \quad .$$

Hence,

$$\frac{r}{\sqrt{1-r^2}} \leqslant \frac{\sqrt{3}}{3},$$

which implies that $r \leqslant 1/2$.

When $r = 1/2$, we let $\beta = 2/3$. Then the image of $\mathcal{B}(0, r)$ by the homothety H^β is $\mathcal{B}(1/3, 1/3)$, which by Lemma 2.12 is contained in Σ_0. □

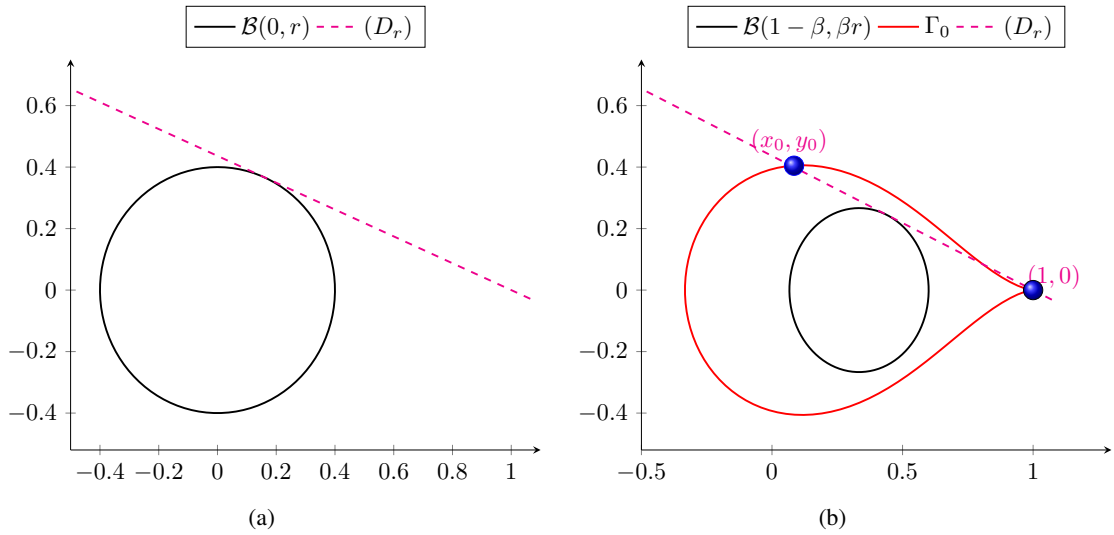


Figure (2.4) (a): the line (D_r) is tangent to the boundary of $\mathcal{B}(0, r)$. (b): the line (D_r) is tangent to the boundary of $\mathcal{B}(1-\beta, \beta r)$, which is contained in Σ_0. The line (D_r) intersects Σ_0 at $(1, 0)$ and (x_0, y_0).

2.4 Analysis of Accelerated Value Iteration of degree d

In this section we consider the acceleration scheme dA-VI (2.8) of any order $d \geqslant 2$. Hereinafter, $\alpha = (\alpha_0, \cdots, \alpha_{d-2}) \in \mathbb{R}^{d-1}$ denotes the vector of parameters required in (2.8b). We shall extend the previous results for $d = 2$ to arbitrary $d \geqslant 2$. That is, with an appropriate choice of α, and under an assumption on the shape of the complex spectrum of P, the asymptotic rate of (2.8) can be $1 - \epsilon^{1/d}$. We refer to Remark 2.2 for a discussion on the connection between the dA-VI (2.8) and Anderson acceleration.

2.4.1 Parameters

We show how to select the parameters $\alpha = (\alpha_0, \cdots, \alpha_{d-2})$ in (2.8b) to obtain an acceleration of any order $d \geqslant 2$. For the sake of simplicity we let $\beta = 1$. Then $z_k = y_k - x_*$ satisfies the following system of linear equations:

$$\begin{pmatrix} z_{k+1} \\ z_k \\ \vdots \\ z_{k-d+2} \end{pmatrix} = Q_{\alpha,d} \begin{pmatrix} z_k \\ z_{k-1} \\ \vdots \\ z_{k-d+1} \end{pmatrix}$$

where

$$Q_{\alpha,d} := \begin{pmatrix} (1+\alpha_{d-2}+\cdots+\alpha_0)P & -\alpha_{d-2}P & \cdots & -\alpha_0 P \\ I & 0 & \cdots & 0 \\ \vdots & \ddots & & \vdots \\ 0 & \cdots & I & 0 \end{pmatrix},$$

We introduce the following rational function of degree d defined by

$$\phi_{\alpha,d}(\lambda) = \frac{\lambda^d}{U(\lambda)}, \tag{2.21}$$

where $U(\cdot) : \mathbb{C} \to \mathbb{C}$ is the polynomial of degree $d-1$ given by:

$$U(\lambda) = (1+\alpha_{d-2}+\cdots+\alpha_0)\lambda^{d-1} - \alpha_{d-2}\lambda^{d-2} - \cdots - \alpha_0 \ .$$

The polynomial U satisfies $U(1) = 1$. The following standard result, which is proved as Lemma 2.4 above, relates the eigenvalues of $Q_{\alpha,d}$ with those of P.

Lemma 2.17. *λ is an eigenvalue of Q_α if and only if there exists an eigenvalue δ of P such that $\delta = \phi_{\alpha,d}(\lambda)$. In other words,*

$$\operatorname{spec} Q_{\alpha,d} = \phi_{\alpha,d}^{-1}(\operatorname{spec} P).$$

We want to choose the vector of parameters α that leads to the smallest possible spectral radius for $Q_{\alpha,d}$, in order to obtain the smallest asymptotic rate for (2.8), like in the case of AVI (i.e. $d = 2$).

Lemma 2.18. *The best choice of the parameters $\alpha_0, \cdots, \alpha_{d-2}$ that minimizes the maximum of the moduli of the preimages of $1-\epsilon$ by $\phi_{\alpha,d}$ is:*

$$\alpha_i = \binom{d}{i} \frac{(\epsilon^{1/d}-1)^{d-i}}{(1-\epsilon)}, \quad \forall \quad i = 0, \cdots, d-2 \ , \tag{2.22}$$

and it corresponds to the following rational function

$$\phi_d^*(\lambda) = \frac{(1-\epsilon)\lambda^d}{\lambda^d - (\lambda - (1-\epsilon^{1/d}))^d} \ . \tag{2.23}$$

Proof. It is easy to verify that with the choice of $\alpha_0, \ldots, \alpha_{d-2}$ in (2.22),

$$U(\lambda) = \frac{1}{1-\epsilon}\left(\lambda^d - \left(\lambda - (1-\epsilon^{1/d})\right)^d\right) \ ,$$

and thus it leads to the rational function (2.23). In addition, $\phi_d^*(\lambda) = 1-\epsilon$ iff $\left(\lambda - (1-\epsilon^{1/d})\right)^d = 0$, from which we deduce that the maximal moduli of the preimages of $1-\epsilon$ by ϕ_d^* is $1-\epsilon^{1/d}$.

Let $\lambda_1, \lambda_2, \cdots, \lambda_d$ be the solutions of $\phi_{\alpha,d}(\lambda) = 1 - \epsilon$ satisfying $\max_i |\lambda_i| \leqslant 1 - \epsilon^{1/d}$. Then $\lambda^d - (1-\epsilon)U(\lambda) = \prod_i(\lambda - \lambda_i)$ for all $\lambda \in \mathbb{C}$. By taking $\lambda = 1$ we obtain that $\prod_i(1-\lambda_i) = \epsilon$. We have

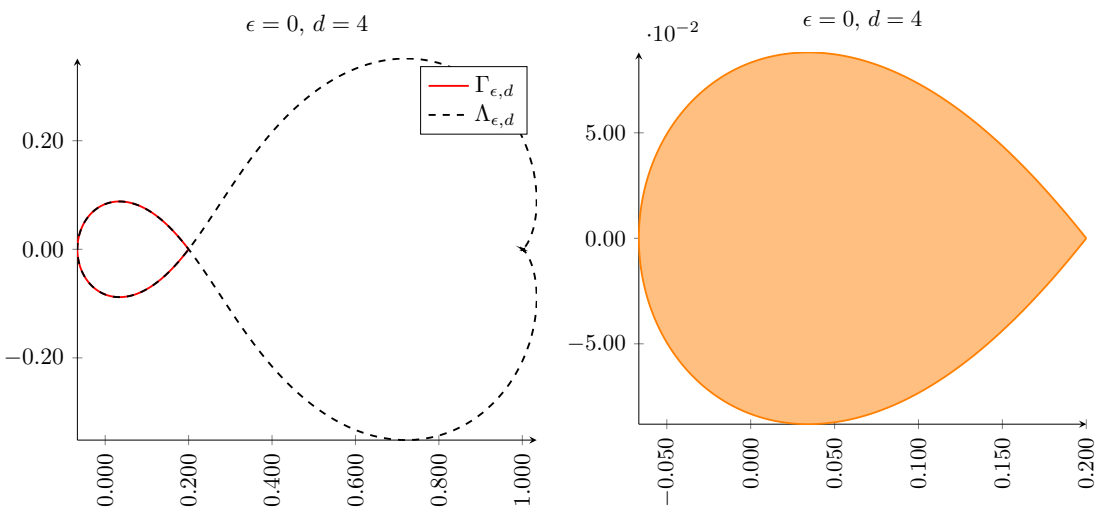
$$\epsilon \leqslant (1 - \max_i |\lambda_i|)^d \leqslant \prod_i (1 - |\lambda_i|) \leqslant \prod_i |1 - \lambda_i| = \epsilon \ .$$

Therefore for all i, $\lambda_i = 1 - \epsilon^{1/d}$ and $\phi_{\alpha,d}$ is exactly ϕ^*_d. □

In the following, we consider the scheme (2.8) implemented with the special choice of the parameters $\{\alpha_i : i = 0, \cdots, d-2\}$ given in (2.22). We want to generalize the characterization of the accelerable region Σ_ϵ for the AVI algorithm to get the largest accelerable region for the dA-VI algorithm. For this purpose, for any $\epsilon \geqslant 0$ and $d \geqslant 2$, let $\Gamma_{\epsilon,d}$ be the simple closed curve defined by the parametric equation:

$$\theta \mapsto \frac{(1-\epsilon)e^{id\theta}}{e^{id\theta} - (e^{i\theta} - 1)^d}, \quad \theta \in \left(\pi - \frac{2\pi}{d}, \pi + \frac{2\pi}{d}\right] \ . \tag{2.24}$$

See an illustration in Figure 2.5 for $\epsilon = 0$ and $d = 4$. Denote by $\Sigma_{\epsilon,d}$ the compact set delimited by



(a) curve $\Gamma_{\epsilon,d}$ (b) region $\Sigma_{\epsilon,d}$

Figure (2.5) Ilustration of the curve $\Gamma_{\epsilon,d}$ (the curve in red in Figure 2.5(a)) and its enclosed region $\Sigma_{\epsilon,d}$ (Figure 2.5(b)). The dashed curve in Figure 2.5(a) corresponds to $\left\{\frac{(1-\epsilon)e^{id\theta}}{e^{id\theta}-(e^{i\theta}-1)^d} : \theta \in (0, 2\pi])\right\}$.

the simple closed curve $\Gamma_{\epsilon,d}$. The following theorem identifies conditions on the spectrum of the initial matrix P which guarantee that the dA-VI algorithm converges asymptotically with a rate $1 - \epsilon^{1/d}$.

Theorem 2.19. *Choosing the parameters* $(\alpha_0, \ldots, \alpha_{d-1})$ *as in* (2.22)*, we get that*

$$\operatorname{spec} Q_{\alpha,d} \subset \mathcal{B}(0, 1 - \epsilon^{1/d}),$$

if and only if $\operatorname{spec} P \subset \Sigma_{\epsilon,d} \cup \{1 - \epsilon\}$.

The proof is given in the next subsection.

2.4.2 Proof of Theorem 2.19

Lemma 2.20. $\operatorname{spec} Q_{\alpha,d} \subset \mathcal{B}(0, 1-\epsilon^{1/d})$ *if and only if*

$$\operatorname{spec} P \subset \{(1-\epsilon)z : \psi_d^{-1}(z) \subset \mathcal{B}(0,1)\} \ , \tag{2.25}$$

where ψ_d *is the rational function defined by*

$$\psi_d(\lambda) = \frac{\lambda^d}{\lambda^d - (\lambda-1)^d} \ .$$

Proof. We note from Lemma 2.17 that $\operatorname{spec} Q_{\alpha,d} \subset \mathcal{B}(0, 1-\epsilon^{1/d})$ if and only if

$$\operatorname{spec} P \subset \{z : (\phi_d^*)^{-1}(z) \subset \mathcal{B}(0, 1-\epsilon^{1/d})\} \ . \tag{2.26}$$

We note the following property:

$$\phi_d^*((1-\epsilon^{1/d})\lambda) = \frac{(1-\epsilon)\lambda^d}{\lambda^d - (\lambda-1)^d} = (1-\epsilon)\psi_d(\lambda) \ , \tag{2.27}$$

Hence (2.26) is equivalent to (2.25). □

We next give a description of the following set.

$$\mathcal{S} := \left\{ z \in \mathbb{C} : \psi_d^{-1}(z) \subset \mathcal{B}(0,1) \right\}. \tag{2.28}$$

We shall need to define

$$\mathcal{Q} := \bigcap_{k=0}^{d-1} e^{\frac{2k\pi i}{d}} H,$$

where

$$H := \{w \in \mathbb{C} : \operatorname{Re}(w) \leqslant 1/2\},$$

is the half-plane containing all the complex numbers with real part smaller than $1/2$, and $e^{\alpha i}H$ denotes the halfspace obtained by rotating H of angle α.

Lemma 2.21.

$$\mathcal{S} = \left\{ \frac{1}{1-\frac{1}{z^d}} : z \in \mathcal{Q} \right\} \cup \{1\}. \tag{2.29}$$

Proof. We define two self-maps of the extended complex plane $\bar{\mathbb{C}}$:

$$f_1(\lambda) := \frac{\lambda}{\lambda - 1} \ , \tag{2.30}$$

$$f_2(\lambda) := \lambda^d \ . \tag{2.31}$$

Note that

$$f_1(\lambda) = 1 + \frac{1}{\lambda - 1},$$

which entails that f_1 is an inversion of center 1. In particular, $f_1 \circ f_1(\lambda) = \lambda$ for any $\lambda \in \bar{\mathbb{C}}$. It is easy to see that

$$\psi_d(\lambda) = f_1 \circ f_2 \circ f_1(\lambda), \ \ \forall \lambda \in \bar{\mathbb{C}}. \tag{2.32}$$

Hence we know that

$$\begin{aligned}\mathcal{S} &= \left\{z \in \mathbb{C} : \psi_d^{-1}(z) \subset \mathcal{B}(0,1)\right\} \\ &= \left\{z \in \mathbb{C} : f_1^{-1}\left(f_2^{-1}\left(f_1^{-1}(z)\right)\right) \subset \mathcal{B}(0,1)\right\} \\ &= \left\{z \in \mathbb{C} : f_2^{-1}\left(f_1^{-1}(z)\right) \subset f_1(\mathcal{B}(0,1))\right\} \\ &= \left\{f_1(w) \in \mathbb{C} : f_2^{-1}(w) \subset f_1(\mathcal{B}(0,1))\right\},\end{aligned} \tag{2.33}$$

where the second equality used (2.32), the third equality relies on the bijection property of f_1 and the last equality applies the change of variable $w = f_1^{-1}(z)$.

Now we characterize the set $f_1(\mathcal{B}(0,1))$. Note that $z = f_1(w)$ if and only if $\frac{1}{z} + \frac{1}{w} = 1$. Thus there is $w \in \mathcal{B}(0,1)$ such that $z = f_1(w)$ if and only if $|1 - \frac{1}{z}| \geqslant 1$. We then deduce that

$$f_1(\mathcal{B}(0,1)) = \left\{w \in \bar{\mathbb{C}} : |w-1| \geqslant |w|\right\}. \tag{2.34}$$

Note that

$$\{w \in \mathbb{C} : |w-1| \geqslant |w|\} = \{w \in \mathbb{C} : \mathrm{Re}(w) \leqslant 1/2\} = H.$$

Indeed, it is known that a circle passing through the center of an inversion is sent to a line by this inversion, and the disk delimited by the circle is send to a half-plane. We conclude that

$$f_1(\mathcal{B}(0,1)) = H \cup \{\infty\}. \tag{2.35}$$

Plugging (2.35) into (2.33) we obtain

$$\mathcal{S} = \left\{f_1(w) \in \mathbb{C} : f_2^{-1}(w) \subset H \cup \{\infty\}\right\}. \tag{2.36}$$

It remains to characterize the set

$$\left\{w \in \bar{\mathbb{C}} : f_2^{-1}(w) \subset H \cup \{\infty\}\right\} = \left\{w \in \mathbb{C} : f_2^{-1}(w) \subset H\right\} \cup \{\infty\}.$$

Define:

$$\bar{\mathcal{Q}} := \left\{z \in \mathbb{C} : f_2^{-1}(f_2(z)) \subset H\right\}. \tag{2.37}$$

It is easy to see that:

$$\left\{w \in \mathbb{C} : f_2^{-1}(w) \subset H\right\} = \left\{f_2(z) : z \in \bar{\mathcal{Q}}\right\}$$

It follows that

$$\left\{w \in \bar{\mathbb{C}} : f_2^{-1}(w) \subset H \cup \{\infty\}\right\} = \left\{f_2(z) : z \in \bar{\mathcal{Q}}\right\} \cup \{\infty\}. \tag{2.38}$$

Finally plugging (2.38) into (2.36) we obtain that

$$\mathcal{S} = \left\{f_1(f_2(z)) \in \mathbb{C} : z \in \bar{\mathcal{Q}}\right\} \cup \{1\} = \left\{\frac{1}{1 - \frac{1}{z^d}} : z \in \bar{\mathcal{Q}}\right\} \cup \{1\}. \tag{2.39}$$

Since for any $z \in \mathbb{C}$,

$$f_2^{-1}(f_2(z)) = \left\{z, e^{-\frac{2\pi i}{d}} z, \ldots, e^{-\frac{2(d-1)\pi i}{d}} z\right\},$$

we obtain

$$\bar{\mathcal{Q}} = \left\{z \in \mathbb{C} : \left\{z, e^{-\frac{2\pi i}{d}} z, \ldots, e^{-\frac{2(d-1)\pi i}{d}} z\right\} \subset H\right\}. \tag{2.40}$$

Therefore, $\bar{\mathcal{Q}}$ is actually the intersection of d halfspaces obtained by rotating H of angles $\frac{2k\pi}{d}$ for $k = 0, \ldots, d-1$. Namely,

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

□

Remark 2.22. For $d = 2$, $\mathcal{Q}$ is the set of complex numbers with real part in $[-1/2, 1/2]$. For $d \geqslant 3$, $\mathcal{Q}$ is a regular polygon with d vertices which circumscribes the disk $\mathcal{B}(0, 1/2)$, see Figure 2.6 for illustration from $d = 2$ to $d = 5$. In particular we have $\mathcal{B}(0, 1/2) \subset \mathcal{Q} \subset \mathcal{B}(0, 1/(2\cos(\pi/d)))$ and $\mathcal{Q}$ asymptotically approximates $\mathcal{B}(0, 1/2)$ when $d \to +\infty$. It follows that

$$\left\{ \frac{1}{1 - \frac{1}{z^d}} : z \in \mathcal{B}(0, 1/2) \right\} \cup \{1\} \subset \mathcal{S} \subset \left\{ \frac{1}{1 - \frac{1}{z^d}} : z \in \mathcal{B}(0, 1/(2\cos(\pi/d))) \right\} \cup \{1\} \ .$$

Note that for any $a > 1$,

$$\left\{ \frac{1}{1 - \frac{1}{z^d}} : z \in \mathcal{B}(0, a) \right\} = \left\{ \frac{1}{1 - \frac{1}{w}} : w \in \mathcal{B}(0, a^d) \right\} = \left\{ \frac{1}{z} : |z - 1| \geqslant a^d \right\}.$$

and thus,

$$\mathcal{B}(0, 1/(a^d + 1)) \subset \left\{ \frac{1}{1 - \frac{1}{z^d}} : z \in \mathcal{B}(0, a) \right\} \subset \mathcal{B}(0, 1/(a^d - 1)).$$

This allows to deduce the following estimation of the region $\mathcal{S}$.

$$\mathcal{B}\left(0, \frac{1}{2^d + 1}\right) \cup \{1\} \subset \mathcal{S} \subset \mathcal{B}\left(0, \frac{1}{(2\cos(\pi/d))^d - 1}\right) \cup \{1\}. \tag{2.41}$$

Next we characterize the boundary of the accelerable region $\mathcal{S}$. We denote by $\mathrm{Bd}\, S$ the boundary of a set S.

Proposition 2.23. *We have*

$$\mathcal{S} = \Sigma_{0,d} \cup \{1\} \tag{2.42}$$

where $\Sigma_{0,d}$ is the compact set of the complex plane delimited by the simple closed curve $\Gamma_{0,d}$ as defined in (2.24).

Proof. Since f_2 is holomorphic and thus open, it sends the interior of $\mathcal{Q}$ into the interior of $f_2(\mathcal{Q})$. It follows that $\mathrm{Bd}\, f_2(\mathcal{Q}) \subset f_2(\mathrm{Bd}\, \mathcal{Q})$. By the continuity of f_2, for any $z \in \mathrm{Bd}\, \mathcal{Q}$ and any $\epsilon > 0$, there is $\delta > 0$ such that

$$f_2(\mathcal{B}(z, \delta)) \subset \mathcal{B}(f_2(z), \epsilon).$$

Since $z \in \mathrm{Bd}\, \mathcal{Q}$, $\mathcal{B}(z, \delta) \cap \mathcal{Q}^c \neq \emptyset$ and thus $f_2(\mathcal{B}(z, \delta)) \cap f_2(\mathcal{Q}^c) \neq \emptyset$.we note from the definition (2.40) that

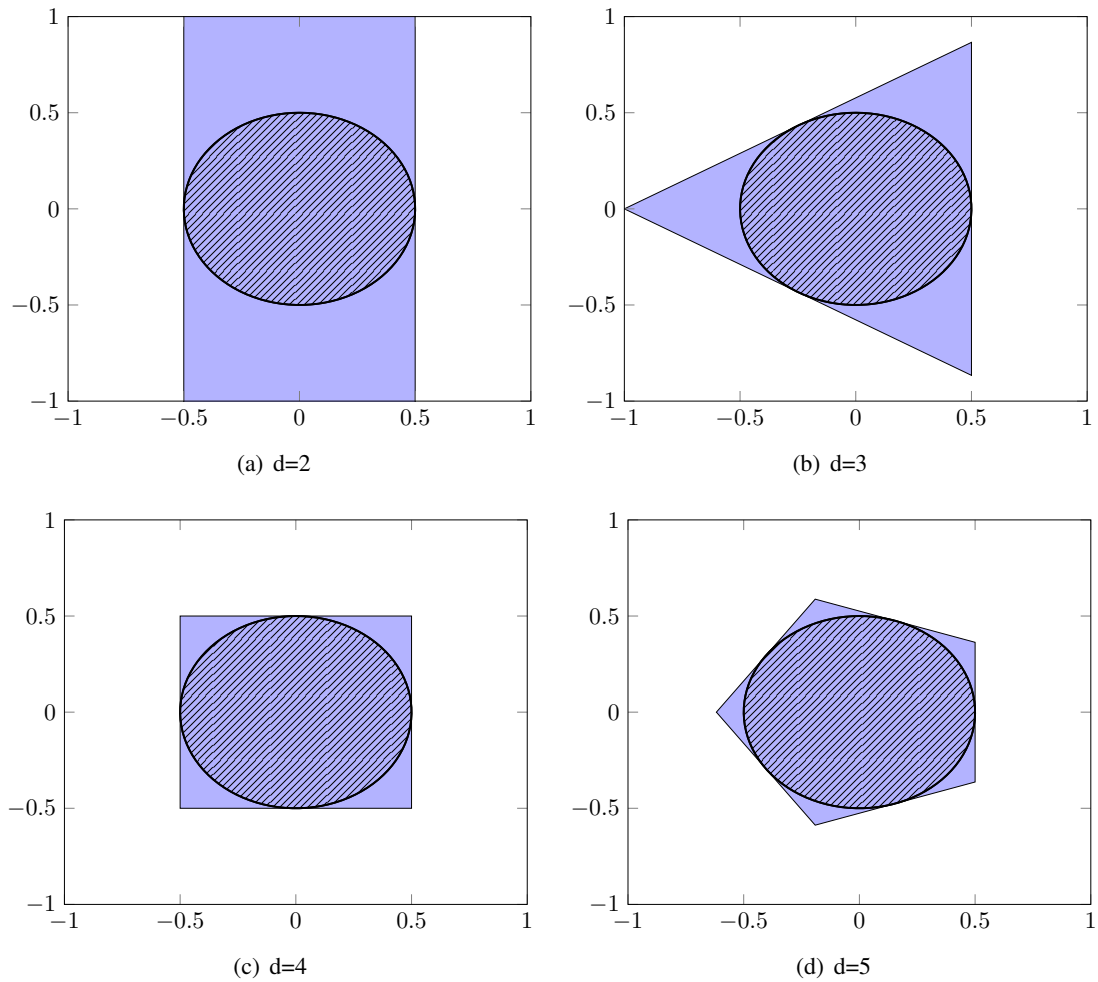
$$f_2(\mathcal{Q}) \cap f_2(\mathcal{Q}^c) = \emptyset.$$

Thereby $f_2(\mathcal{B}(z, \delta)) \cap (f_2(\mathcal{Q}))^c \neq \emptyset$ and $\mathcal{B}(f_2(z), \epsilon) \cap (f_2(\mathcal{Q}))^c \neq \emptyset$. This shows that $f_2(z) \in \mathrm{Bd}(f_2(\mathcal{Q}))$ and thus $f_2(\mathrm{Bd}\, \mathcal{Q}) \subset \mathrm{Bd}\, f_2(\mathcal{Q})$. We thus proved that

$$\mathrm{Bd}\, f_2(\mathcal{Q}) = f_2(\mathrm{Bd}\, \mathcal{Q}). \tag{2.43}$$

Since $f_1 : \bar{\mathbb{C}} \to \bar{\mathbb{C}}$ is a homeomorphism, we know that

$$\mathrm{Bd}\, f_1(f_2(\mathcal{Q})) = f_1(\mathrm{Bd}\, f_2(\mathcal{Q})) \overset{(2.43)}{=} f_1(f_2(\mathrm{Bd}\, \mathcal{Q})). \tag{2.44}$$



(a) d=2 (b) d=3

(c) d=4 (d) d=5

Figure (2.6) Ilustration of $\mathcal{Q}$ (the region in blue) and the circumscribed disk $\mathcal{B}(0, 1/2)$ (the dashed region).

As mentioned in Remark 2.22, for $d = 2$, $\mathcal{Q}$ is the set of complex numbers with real part in $[-1/2, 1/2]$ and the boundary of $\mathcal{Q}$ can be described as follows:

$$\operatorname{Bd} \mathcal{Q} = \left\{ \frac{\pm(1 + i \tan\theta)}{2} : \theta \in \left(-\frac{\pi}{2}, \frac{\pi}{2}\right) \right\}.$$

For $d \geqslant 3$, $\mathcal{Q}$ is the regular convex polygon with boundary given by the simple closed curve:

$$\operatorname{Bd} \mathcal{Q} = \left\{ \frac{e^{\frac{2k\pi i}{d}}(1 + i \tan\theta)}{2} : \theta \in \left(-\frac{\pi}{d}, \frac{\pi}{d}\right], k \in \{0, \dots, d-1\} \right\}.$$

Since

$$\begin{aligned} \frac{1}{1 - e^{-i\theta}} &= \frac{1}{1 - \cos\theta + i \sin\theta} = \frac{1 - \cos\theta - i\sin\theta}{2 - 2\cos\theta} \\ &= \frac{1}{2} - \frac{i \sin\theta}{2(1 - \cos\theta)} = \frac{1}{2} + \frac{i \sin(\theta + \pi)}{2(1 + \cos(\theta + \pi))} \\ &= \frac{(1 + i \tan\frac{\theta + \pi}{2})}{2}, \end{aligned}$$

we obtain another representation of $\operatorname{Bd} \mathcal{Q}$:

$$\operatorname{Bd} \mathcal{Q} = \begin{cases} \left\{ \frac{\pm 1}{1 - e^{-i\theta}} : \theta \in (0, 2\pi) \right\} & \text{if } d = 2 \\ \left\{ \frac{e^{\frac{2k\pi i}{d}}}{1 - e^{-i\theta}} : \theta \in \left(\pi - \frac{2\pi}{d}, \pi + \frac{2\pi}{d}\right], k \in \{0, \dots, d-1\} \right\}, & \text{if } d \geqslant 3 \end{cases} \tag{2.45}$$

Plugging (2.45) into (2.44) we obtain that

$$\operatorname{Bd} f_1(f_2(\mathcal{Q})) = \begin{cases} \left\{ \frac{1}{1 - (1 - e^{-i\theta})^2} : \theta \in (0, 2\pi) \right\} & \text{if } d = 2 \\ \left\{ \frac{1}{1 - (1 - e^{-i\theta})^d} : \theta \in \left(\pi - \frac{2\pi}{d}, \pi + \frac{2\pi}{d}\right] \right\} & \text{if } d \geqslant 3 \end{cases} \tag{2.46}$$

Therefore, define the set

$$\Sigma_{0,d} := \begin{cases} f_1(f_2(\mathcal{Q})) \cup \{1\} & \text{if } d = 2 \\ f_1(f_2(\mathcal{Q})) & \text{if } d \geqslant 3 \end{cases} \tag{2.47}$$

Then we have (2.42) and

$$\operatorname{Bd} \Sigma_{0,d} = \begin{cases} \left\{ \frac{1}{1 - (1 - e^{-i\theta})^2} : \theta \in (0, 2\pi] \right\} & \text{if } d = 2 \\ \left\{ \frac{1}{1 - (1 - e^{-i\theta})^d} : \theta \in \left(\pi - \frac{2\pi}{d}, \pi + \frac{2\pi}{d}\right] \right\} & \text{if } d \geqslant 3 \end{cases} \tag{2.48}$$

which can be written as

$$\operatorname{Bd} \Sigma_{0,d} = \left\{ \frac{e^{id\theta}}{e^{id\theta} - (e^{i\theta} - 1)^d}, \quad \theta \in \left(\pi - \frac{2\pi}{d}, \pi + \frac{2\pi}{d}\right] \right\} .$$

for any $d \geqslant 2$. Finally the compactness of $\Sigma_{0,d}$ follows from the compactness of $\mathcal{S}$, which can be easily seen from the fact that $\mathcal{S} \subset \psi_d(\mathcal{B}(0,1))$ by the definition (2.28). □

Proof of Theorem 2.19. This follows directly from (2.25), (2.28) and (2.42). □

Remark 2.24. In Theorem 2.19, the region $\Sigma_{\epsilon,d}$ assured to be accelerable *does not* contain some part of the real interval $[0, 1 - \epsilon]$ for any $d \geqslant 3$. This is consistent with Theorem 2.2.12 of [Nes04] implying that for a linear recurrent scheme with finite memory calling the oracle T, the geometric convergence rate cannot be smaller than $1 - O(\kappa^{-1/2})$ where κ is a condition number, corresponding to ϵ^{-1} here.

2.4.3 Robustness of the acceleration

Note that the parameters $(\alpha_0, \dots, \alpha_{d-1})$ defined in (2.22) requires the knowledge of ϵ thus of the exact value of the spectral radius of P, which may be a restrictive assumption for practitioners. For example, in the stochastic shortest path problem analyzed in [BT91a], we do not know the spectral radii of the sub-stochastic matrices arising in the restricted contracting operator described in Proposition 1 of [BT91a]. In this section we evaluate how the small perturbation of ϵ will affect the order of convergence of the acceleration scheme. This in particular allows the use of an approximate value of ϵ to compute the parameters $(\alpha_0, \dots, \alpha_{d-1})$ while still achieving an asymptotic convergence rate of order $1 - \Omega(\epsilon^{1/d})$.

For $h \geqslant 0$, we are looking for the smallest radius $g_\epsilon(h) \geqslant 0$ such that $(\phi_d^*)^{-1}(\mathcal{B}(1-\epsilon, h)) \subset \mathcal{B}(1-\epsilon^{1/d}, g_\epsilon(h))$, and we enforce $g_\epsilon(h) \leqslant \epsilon^{1/d}$ to preserve the acceleration. First we make this analysis for ψ_d (i.e. $\epsilon = 0$).

Lemma 2.25. *For $h \geqslant 0$, the smallest nonnegative real number $g_0(h)$ such that*

$$\psi_d^{-1}(\mathcal{B}(1, h)) \subset \mathcal{B}(1, g_0(h))$$

is

$$g_0(h) = \frac{h^{1/d}}{(1+h)^{1/d} - h^{1/d}}, \quad \forall h \geqslant 0 \ .$$

Proof. For $h = 0$, it follows from $\psi_d^{-1}(1) = \{1\}$.

Now let $h > 0$, we want to have $\psi_d(\mathcal{B}(1, g_0(h))^c) \subset \mathcal{B}(1, h)^c$, i.e.:

$$|\lambda| > g_0(h) \Rightarrow |\psi_d(1+\lambda) - 1| > h, \ \ \forall \lambda \in \mathbb{C}. \tag{2.49}$$

We have $\psi_d(1+\lambda) = \frac{(1+\lambda)^d}{(1+\lambda)^d - \lambda^d} = 1 + \frac{\lambda^d}{(1+\lambda)^d - \lambda^d} = 1 + \frac{1}{(1+\frac{1}{\lambda})^d - 1}$, then

$$|\psi(1+\lambda) - 1| > h \Leftrightarrow \left| \left(1 + \frac{1}{\lambda}\right)^d - 1 \right| < \frac{1}{h}, \ \ \forall \lambda \in \mathbb{C}.$$

For any $\lambda \in \mathbb{C}$ we know that

$$\left| \left(1 + \frac{1}{\lambda}\right)^d - 1 \right| = \left| \sum_{k=1}^{d} \binom{d}{k} \frac{1}{\lambda^k} \right| \leqslant \sum_{k=1}^{d} \binom{d}{k} \frac{1}{|\lambda|^k} = \left(1 + \frac{1}{|\lambda|}\right)^d - 1. \tag{2.50}$$

Thus

$$|\lambda| > \frac{h^{1/d}}{(1+h)^{1/d} - h^{1/d}} \Leftrightarrow \left(1 + \frac{1}{|\lambda|}\right)^d - 1 < \frac{1}{h} \Rightarrow |\psi(1+\lambda) - 1| > h, \ \ \forall \lambda \in \mathbb{C}.$$

This allows to conclude because we have equality in (2.50) when $\lambda \in \mathbb{R}_+$. □

Lemma 2.26. *For any $a \in [0, 1]$ we have*

$$(\phi_d^*)^{-1}(\mathcal{B}(1-\epsilon, a\epsilon)) \subset \mathcal{B}(1-\epsilon^{1/d}, a^{1/d}\epsilon^{1/d}) \ .$$

Proof. By the property (2.27) and Lemma 2.25 we deduce that for $h \geqslant 0$, the smallest radius $g_\epsilon(h)$ such that $(\phi_d^*)^{-1}(\mathcal{B}(1-\epsilon, h)) \subset \mathcal{B}(1-\epsilon^{1/d}, g_\epsilon(h))$ is given by

$$g_\epsilon(h) = \frac{(1-\epsilon^{1/d})h^{1/d}}{(1+h-\epsilon)^{1/d} - h^{1/d}}, \quad \forall h \geqslant 0.$$

Note that

$$(1+h-\epsilon)^{1/d} - h^{1/d} \geqslant 1-\epsilon^{1/d}, \ \ \forall h \leqslant \epsilon.$$

Hence

$$g_\epsilon(h) \leqslant h^{1/d}, \ \ \forall h \leqslant \epsilon.$$

We achieve the proof by taking $h = a\epsilon$. □

The following theorem describes a d-accelerable region.

Theorem 2.27. *Let* $a \in [0,1[$, *if* $\operatorname{spec} P \subset \mathcal{B}\left(0, \frac{1-\epsilon}{2^d+1}\right) \cup \mathcal{B}\left(1-\epsilon, a\epsilon\right)$ *then with the choice of* α *specified in* (2.22) *we have,*

$$\operatorname{spec} Q_{\alpha,d} \subset \mathcal{B}(0, 1-\epsilon^{1/d}) \cup \mathcal{B}(1-\epsilon^{1/d}, a^{1/d}\epsilon^{1/d}) \ \ ,$$

so that the iterates of the dA-VI algorithm (2.8) *with* $\beta = 0$ *satisfy*

$$\limsup_{k\to\infty} \|x_k - x_*\|^{1/k} \leqslant 1 - (1-a^{1/d})\epsilon^{1/d}.$$

Proof. By combining Theorem 2.19, Equation (2.41) and Lemma 2.26. □

2.4.4 Application to Markov Decision Processes: Accelerated Policy Iteration

As an application, we consider the standard discounted Markov decision process (MDP) with state space $[n] := \{1, \dots, n\}$, see [Whi83, BT96] for background. For each state i, denote by $\mathcal{A}(i)$ the set of actions, $P^a_{i,j}$ the transition probability from state i to state j under action $a \in \mathcal{A}(i)$, and g^a_i the reward of choosing action $a \in \mathcal{A}(i)$ in state i. Let $1 > \gamma_i > 0$, for $i \in [n]$, be state-dependent discount factors. The associated dynamic programming operator $T : \mathbb{R}^n \to \mathbb{R}^n$ is given by:

$$T_i(x) := \max_{a\in\mathcal{A}(i)} \gamma_i \sum_{j\in[n]} P^a_{i,j} x_j + g^a_i, \ \ \forall i \in [n]\,. \tag{2.51}$$

We set $\gamma := \max_{i\in[n]} \gamma_i$.

The value of the discounted problem for this MDP starting from an initial state i is given by:

$$v_i := \max_{a_0,a_1,\cdots} \mathbb{E}[g^{a_0}_{X_0} + \gamma_{X_0} g^{a_1}_{X_1} + \gamma_{X_0}\gamma_{X_1} g^{a_2}_{X_2} + \cdots \mid X_0 = i] \ \ ,$$

where the maximum is taken over admissible sequences of random actions, and $X_0, X_1, \cdots$ denotes the random sequence of states generated by the actions.

We are interested in finding the value vector $v \in \mathbb{R}^n$ of this MDP which is a solution of the fixed point problem $v = T(v)$. The fixed point exists and is unique since T is a contraction of constant γ in the sup-norm.

A classical approach to solve this problem is to use value iteration, i.e., to compute the sequence $v^k = T(v^{k-1})$, which converges to the unique fixed point. It is tempting to apply directly accelerated value iteration to the non-linear problem $v = T(v)$. This approach was proposed in [GGC19], and it is experimentally effective on some instances. However, the convergence proof of accelerated value iteration uses inherently the affine character of the operator T, and it is not clear whether general enough convergence conditions can be given for Markov decision processes. An alternative approach, which we develop here, is to rely on *policy iteration* instead of value iteration, which will allow us to apply the idea of dth acceleration to solve MDP, but in an indirect manner, leading to convergence guarantees.

A *policy* is a map $\sigma : [n] \to \cup_{i\in[n]} A(i)$ such that $\sigma(i) \in A(i)$, it represents a state dependent decision rule. It determines a 0-player game, with an affine operator $T^\sigma : \mathbb{R}^n \to \mathbb{R}^n$,

$$T_i^\sigma(x) := \gamma_i \sum_{j\in[n]} P_{i,j}^{\sigma(i)} x_j + g_i^{\sigma(i)}, \quad \forall i \in [n]\,. \tag{2.52}$$

For a vector $x \in \mathbb{R}^n$, we define $\operatorname{top}(x) = \max_{i\in[n]} x_i$. We have $\|x\|_\infty = \max(\operatorname{top}(x), \operatorname{top}(-x))$. For $x, y \in \mathbb{R}^n$, we write $x \leqslant y$ to mean that $x_i \leqslant y_i$ for all $i \in [n]$. We denote by x_* the unique fixed point of the operator T and by x^σ the unique fixed point of the operator T^σ. We denote by $a^+ = \max(a, 0)$ the positive part of a real number. The following lemma presents some classical properties of the operators T and T^σ that are useful for our analysis.

Lemma 2.28. *Let $x, y \in \mathbb{R}^n$, σ a policy, $e = (1, \cdots, 1) \in \mathbb{R}^n$ the unit vector and $a \in \mathbb{R}^+$ a nonnegative real number, we have:*

$$(\operatorname{top}(T(x) - T(y)))^+ \leqslant \gamma(\operatorname{top}(x - y))^+, \tag{2.53}$$

$$\|x - x^\sigma\|_\infty \leqslant \frac{1}{1-\gamma}\|x - T^\sigma(x)\|_\infty, \tag{2.54}$$

$$\|x - x_*\|_\infty \leqslant \frac{1}{1-\gamma}\|x - T(x)\|_\infty, \tag{2.55}$$

$$T^\sigma(x + ae) \leqslant T^\sigma(x) + \gamma a e, \tag{2.56}$$

$$x \leqslant T^\sigma(x) + ae \Rightarrow x \leqslant x^\sigma + \frac{a}{1-\gamma} e, \tag{2.57}$$

$$x \leqslant T(x) + ae \Rightarrow x \leqslant x_* + \frac{a}{1-\gamma} e, \tag{2.58}$$

$$x \leqslant y \Rightarrow T^\sigma(x) \leqslant T^\sigma(y). \tag{2.59}$$

Property (2.59) follows from $P_{ij}^a \geqslant 0$, whereas (2.56) follows from $\sum_j P_{ij}^a = 1$. Property (2.53) means that T is a contraction of rate γ in the nonsymmetric norm $(x, y) \mapsto (\operatorname{top}(x - y))^+$. To see it we compute for $i \in [n]$, $T_i(x) - T_i(y) = \max_a\{\gamma_i \sum_{j\in[n]} P_{i,j}^a x_j + g_i^a\} - \max_a\{\gamma_i \sum_{j\in[n]} P_{i,j}^a y_j + g_i^a\} \leqslant (\gamma_i \sum_{j\in[n]} P_{i,j}^{a_x} x_j + g_i^{a_x}) - (\gamma_i \sum_{j\in[n]} P_{i,j}^{a_x} y_j + g_i^{a_x}) = \gamma_i \sum_{j\in[n]} P_{i,j}^{a_x}(x_j - y_j)$, where a_x is the action that maximizes the expression of $T_i(x)$. Property (2.54) (and similarly (2.55)) comes street forward from x^σ being a fixed point of T^σ and the latter being a γ-contraction in the sup norm.

To obtain property (2.57) (and similarly (2.58)), we apply k times the operator T^σ to both sides of the initial inequality and we use the properties (2.59) and (2.56) to obtain that $x \leqslant (T^\sigma)^{k+1}(x) + a\sum_{i=0}^k \gamma^i e$ and finely since T^σ is a strict contraction, we know that when k goes to infinity, $(T^\sigma)^{k+1}(x)$ converges to the fixed point x^σ.

Policy iteration computes a succession of policies $\sigma^1, \sigma^2, \ldots$. At each stage, it solves a 0-player fixed point problem, finding a vector v^k such that $v^k = T^{\sigma^k}(v^k)$. Then, the vector v^k is used to determine the new policy, by considering the maximizing actions in the expression of $T(v^k)$. When policy iteration is implemented in exact arithmetics, for a fixed $\gamma < 1$, the number of iterations is strongly polynomial [Ye11b]. Moreover, on ordinary instances, the number of iterations is often of a few units. Hence, the bottleneck, preventing to apply policy iteration to large scale Markov decision problems, is generally the solution of the affine problem $v^k = T^{\sigma^k}(v^k)$: algebraic methods, based on LU-factorization, are not adapted to large scale sparse instances, whereas standard iterative methods can be slow, since the contraction rate γ is typically close to 1. To address this difficulty, we present a version of policy iteration in which at each stage, v^k is computed by the dth accelerated scheme.

We consider the *Accelerated Policy Iteration of degree d (dA-PI)* presented in Algorithm 1. Using classical estimates on approximate value iteration, see [BT96, Ber11, SGG+15], we get the following convergence result.

Algorithm 1 Accelerated Policy Iteration of degree d (dA-PI).

1: Fix a target accuracy δ for value determination and δ' for policy improvement.
2: Initialization: select a starting policy σ^0, and set the initial values $x_{-1,0} = x_{-1,1} = \cdots = x_{-1,d-2} = y_{-1,d-2} = 0$
3: **for** $k = 0, 1, \cdots$ **do** the following:
4: (Accelerated value determination): Run the dA-VI (2.8) on the operator T^{σ^k} until having a residual smaller that δ: so first we initialize $x_{k,0}, x_{k,1}, \cdots, x_{k,d-2}$ by the last $d-1$ values of the sequence $(x_{k-1,l})_l$ and $y_{k,d-2}$ by the last value of the sequence $(y_{k-1,l})_l$, and for $l = d-2, \cdots$, we do the iterations of (2.8):

$$x_{k,l+1} = (1-\beta)y_{k,l} + \beta T(y_{k,l}) \ , \tag{2.60a}$$

$$y_{k,l+1} = (1 + \alpha_{d-2} + \cdots + \alpha_0)x_{k,l+1} - \alpha_{d-2}x_{k,l} - \cdots - \alpha_0 x_{k,l-d+2} \ , \tag{2.60b}$$

until $\|y_{k,l} - T^{\sigma^k}(y_{k,l})\|_\infty \leqslant \delta$. We denote the final $y_{k,l}$ by y_k.
5: (Policy improvement). We determine a policy σ^{k+1} such that $\|T(y_k) - T^{\sigma^{k+1}}(y_k)\|_\infty \leqslant \delta'$, and for each $i \in [n]$, we choose $\sigma^{k+1}(i) = \sigma^k(i)$ whenever possible.
6: **done**

Proposition 2.29. *Suppose that for any policy σ, $\operatorname{spec} P^\sigma \subset \Sigma_{\epsilon,d} \cup \{1-\epsilon\}$, and that we choose $\alpha = (\alpha_0, \cdots, \alpha_{d-2})$ as in* (2.22)*. Each iteration k of the dA-PI algorithm terminates, and we have :*

$$\limsup_{k\to\infty} \|y_k - x_*\|_\infty \leqslant \frac{(1+\gamma)\delta + \delta'}{(1-\gamma)^2} \ . \tag{2.61}$$

Moreover, if $\sigma^{k+1} = \sigma^k$ for some k, then $\|y_k - x_*\|_\infty \leqslant \frac{\delta+\delta'}{1-\gamma}$.

Proof. The termination of each iteration k comes from Theorem 2.19. For each k, we have from the algorithm $\|y_k - T^{\sigma^k}(y_k)\|_\infty \leqslant \delta$, then $y_k \leqslant T^{\sigma^k}(y_k) + \delta e \leqslant T(y_k) + \delta e$. Then by (2.58) we deduce that $y_k \leqslant x_* + \frac{\delta}{1-\gamma}e$. Therefore $\operatorname{top}(y_k - x_*) \leqslant \frac{\delta}{1-\gamma} \leqslant \frac{(1+\gamma)\delta+\delta'}{(1-\gamma)^2}$.

We have $y_k \leqslant T(y_k) + \delta e \leqslant T^{\sigma^{k+1}}(y_k) + (\delta+\delta')e$, then by (2.57) we get $y_k \leqslant x^{\sigma^{k+1}} + \frac{\delta+\delta'}{1-\gamma}e$. By (2.54) and the algorithm, we have $\|y_{k+1} - x^{\sigma^{k+1}}\|_\infty \leqslant \|y_{k+1} - T^{\sigma^{k+1}}(y_{k+1})\|_\infty/(1-\gamma) \leqslant \delta/(1-\gamma)$, then $x^{\sigma^{k+1}} \leqslant y_{k+1} + \frac{\delta}{1-\gamma}e$. We deduce that $y_k \leqslant y_{k+1} + \frac{2\delta+\delta'}{1-\gamma}e$. We apply $T^{\sigma^{k+1}}$ to both sides of this inequality and use (2.59) and (2.56) to get that $T(y_k) \leqslant T^{\sigma^{k+1}}(y_k) + \delta' e \leqslant T^{\sigma^{k+1}}(y_{k+1} + \frac{2\delta+\delta'}{1-\gamma}e) + \delta' e \leqslant T^{\sigma^{k+1}}(y_{k+1}) + \frac{(2\delta+\delta')\gamma}{1-\gamma}e + \delta' e \leqslant y_{k+1} + \delta e + \frac{(2\delta+\delta')\gamma}{1-\gamma}e + \delta' e = y_{k+1} + \frac{(1+\gamma)\delta+\delta'}{1-\gamma}e$. Therefore, $x_* - y_{k+1} \leqslant x_* - T(y_k) + \frac{(1+\gamma)\delta+\delta'}{1-\gamma}e = T(x_*) - T(y_k) + \frac{(1+\gamma)\delta+\delta'}{1-\gamma}e$. Then $(\operatorname{top}(x_* - y_{k+1}))^+ \leqslant (\operatorname{top}(T(x_*) - T(y_k)))^+ + \frac{(1+\gamma)\delta+\delta'}{1-\gamma}$, and by using (2.53) we deduce that $(\operatorname{top}(x_* - y_{k+1}))^+ \leqslant \gamma(\operatorname{top}(x_* - y_k))^+ + \frac{(1+\gamma)\delta+\delta'}{1-\gamma}$. By iterating this inequality, we deduce that for iteration k, $(\operatorname{top}(x_* - y_k))^+ \leqslant \gamma^k(\operatorname{top}(x_* - y_0))^+ + \frac{(1+\gamma)\delta+\delta'}{1-\gamma}\sum_{i=0}^{k-1}\gamma^i$. Therefore, $\limsup_{k\to\infty}(\operatorname{top}(x_* - y_k))^+ \leqslant \frac{(1+\gamma)\delta+\delta'}{(1-\gamma)^2}$, and by using $\|x_* - y_k\|_\infty = \max(\operatorname{top}(y_k - x_*), (\operatorname{top}(x_* - y_k))^+)$ we end the proof of (2.61).

Now, if $\sigma^{k+1} = \sigma^k$ for some k, then $\|T(y_k) - T^{\sigma^k}(y_k)\|_\infty \leqslant \delta'$ and we know that $\|y_k - T^{\sigma^k}(y_k)\|_\infty \leqslant \delta$, then $\|y_k - T(y_k)\|_\infty \leqslant \delta + \delta'$. Therefore by (2.55), we get $\|y_k - x_*\|_\infty \leqslant \frac{\delta+\delta'}{1-\gamma}$. □

Remark 2.30. Proposition 2.29 should be compared with [BT96, Prop. 6.2] and Remark 5 and Eqn 22 of [SGG+15], which bound the same limsup by an expression of the form $(\delta' + 2\gamma\epsilon)/(1-\gamma)^2$, where

ϵ is a upper bound of $\|y_k - x^{\sigma^k}\|_\infty$. Here, ϵ is replaced by δ, which is an upper bound of the residual $\|y_k - T^{\sigma^k}(y_k)\|_\infty$.

Remark 2.31. Proposition 2.29 is only an asymptotic result. In contrast, when policy iteration is implemented exactly, the value vector v^k associated to the kth policy that is selected satisfies $\|v_k - x_*\|_\infty \leqslant \gamma^k \|v_0 - x_*\|_\infty$, see Lemma 6.5 of [HMZ13].

Remark 2.32. Since accelerated value iteration, and so, accelerated policy iteration, are implemented with a fixed precision arithmetics, one may wonder whether acceleration leads to numerical instabilities. In the numerical experiments which follows, no such instabilities were observed for the relevant values $d = 2, 4$ considered here. We verified the validity of the approximate solutions that we obtained using the inequality (2.55). Indeed, the residual $\|y_k - T(y_k)\|_\infty$, where y_k is the approximate solution gotten at the final iteration of the algorithm, can be evaluated in an accurate way (with a precision close to the machine precision) using only the last value y_k. So, if this residual is small, by the inequality (2.55), we can certify that $\|y_k - x^*\|_\infty$ is also small, so that we have a valid approximate solution. In all the experiments of Section 2.5, the algorithms are stopped with a residual of $< 10^{-10}$, and $1 - \gamma$ is $\geqslant 10^{-4}$, so, it is guaranteed that the true solution is approximated with a precision $< 10^{-6}$.

2.5 Numerical results

In this section, we show the numerical performance of the proposed dA-VI and dA-PI with $d = 2$ and $d = 4$. The acceleration parameters in all the examples follow (2.22); the parameter $\alpha = \frac{1-\sqrt{\epsilon}}{1+\sqrt{\epsilon}}$ for accelerations of degree 2, and the parameters $\alpha_0 = \frac{(1-\epsilon^{1/4})^4}{1-\epsilon}, \alpha_1 = \frac{-4(1-\epsilon^{1/4})^3}{1-\epsilon}$ and $\alpha_2 = \frac{6(1-\epsilon^{1/4})^2}{1-\epsilon}$ for accelerations of degree 4.

In all the examples below, we do the policy improvement at each iteration k of the dA-PI algorithm in an exact way by taking, for each $i \in [n]$, $\sigma^{k+1}(i) \in [m]$ to be a value achieving the maximum when evaluating (2.51) at $x = y_k$, i.e. $\delta' = 0$, and we let the accuracy of the value determination to be $\delta = 10^{-10}$.

2.5.1 Markov decision processes with random matrices

We consider the discounted MDP model of (2.51). We take a damping parameter $\beta = 1$ in what follows.

The instances used in Figures 2.7 to 2.12 are generated in the following way. We fix two integers n and m. For each $i \in [n]$, we take $\mathcal{A}(i) = [m]$ and randomly generate a probability vector $p_i^a = (P_{i,1}^a, \ldots, P_{i,n}^a)$ as follows: $P_{i,j}^a = \frac{X_{i,j}^a}{X_{i,1}^a + \cdots + X_{i,n}^a}$, where the $X_{i,j}^a$ are independent Bernoulli random variables of mean $p \in (0, 1)$. The discount factors γ_i are randomly chosen in the interval $[1 - 2\epsilon, 1 - \epsilon]$, independently for each $i \in [n]$.

Let $\lambda_1, \ldots \lambda_n$ be the eigenvalues of $\sqrt{n}P$. It is shown in [BCC08] that the counting probability measure $\frac{\delta_{\lambda_1} + \cdots + \delta_{\lambda_n}}{n}$, converges weakly as $n \to \infty$ to the uniform law on the disk $\{z \in \mathbb{C} : |z| \leqslant \sqrt{(1-p)/p}\}$. Moreover, Theorem 1.2, *ibid.* shows that the second modulus of an eigenvalue of P is of order $1/\sqrt{n}$. This explains the shape of the spectrum shown on the figures Figures 2.7 to 2.9, and explains also, along with (2.41), why the accelerated schemes of order 4 work in the large scale example of Figure 2.12 where we take $p = 0.0025$ with $n = 10^5$.

In Figure 2.7, we consider an instance where the matrices are randomly generated as above with $n = 30$, $m = 10$ and $p = 0.2$. In subfigure 2.7(b), we display the spectrum of one matrix $P_\gamma^\sigma := (\gamma_i P_{ij}^\sigma)_{ij}$. One can notice that this spectrum presents eigenvalues that are outside the simply and multiply accelerable regions delimited respectively by Γ_ϵ and $\Gamma_{\epsilon,4}$ (see Theorem 2.19). Therefore, the accelerated

policy iteration algorithms (dA-PI) cannot be applied for this instance. In accordance with that, the subfigure 2.7(a) shows that the accelerated value iteration algorithms 2A-VI and 4A-VI do not converge.

In Figure 2.8, we consider an instance with $n = 100$, $m = 10$ and $p = 0.2$. The subfigure 2.8(b) shows that the spectrum of the random matrices in this case is located in the simply accelerable region delimited by Γ_ϵ, but it is not included in the 4-accelerable region $\Gamma_{\epsilon,4}$. Therefore, we can apply the 2A-PI algorithm but not the 4A-PI in this case. The subfigure 2.8(a) shows that the simply accelerated schemes 2A-PI and 2A-VI have significantly better performances than value iteration algorithm. It shows also as expected that the acceleration of order 4 does not converge.

In Figure 2.9, we consider an instance with $n = 1500$, $m = 10$ and $p = 0.2$. The subfigure 2.9(b) shows that the spectrum of the random matrices in this case is located inside the accelerable regions of order 2 and 4 delimited respectively by Γ_ϵ and $\Gamma_{\epsilon,4}$. Therefore, we can apply both 2A-PI and 4A-PI in this case. The subfigure 2.9(a) shows that all the accelerated schemes converge in this case and that the multi-accelerated schemes have better performances than the simply accelerated ones.

In Figure 2.10, we consider an instance with $n = 4000$, $m = 10$, $p = 0.1$. In this example we take $\epsilon = 10^{-2}$ to allow the Value Iteration algorithm to have a visible improvement.

In Figure 2.11, we consider an instance with $n = 4 \times 10^4$, $m = 10$ and the matrices used are sparse with a parameter $p = 0.005$.

We observe that the classical Policy Iteration (PI) algorithm [How60, Put14], using LU decomposition to solve the linear 0-player problem at each iteration, is way faster than our iterative algorithms (dA-PI and dA-VI) in the case of small matrices like in Figures 2.7 and 2.8, but as the size of the matrices gets bigger our iterative algorithms become more competitive like in Figures 2.9 and 2.10, and even way faster than Policy Iteration like in Figure 2.11.

The Figure 2.12 represents a large scale analogue to the previous examples where the number of states is $n = 10^5$, and the matrices used are sparse with a parameter $p = 0.0025$. For this example, the classical Policy Iteration algorithm cannot be used because of memory saturation. However, the dA-PI algorithms 1 that we propose, with $d = 2$ and $d = 4$ here, work in this case and show significantly better performances than the classical Value Iteration algorithm. The dA-VI algorithms also show competitive performances in comparison with dA-PI algorithms. However, we expect in general that dA-PI becomes more competitive than dA-VI when the number of actions m is large, because the number of policies visited grow slowly with the number of actions (in the discounted case, a worst case almost linear bound for this number is given in [Sch13], based on [Ye11b], the convergence being generally faster on typical instances).

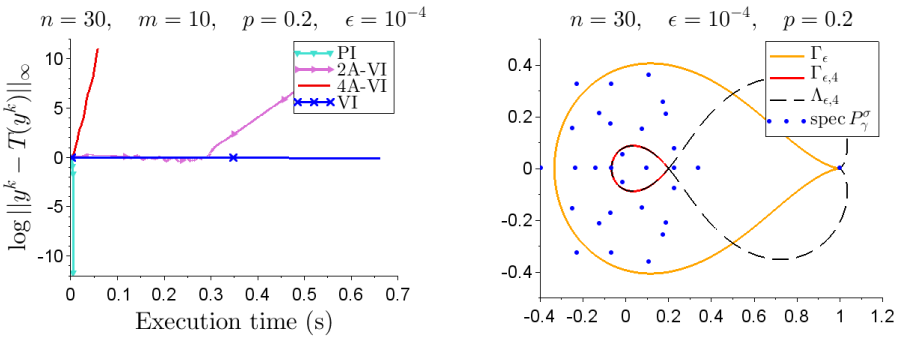
In particular, for all the examples in Figures 2.8 to 2.12, we notice that both 2A-PI and 4A-PI stop only after $k \leqslant 5$ iterations over policies because each one of them finds a policy σ^{k+1} equal to σ^k. The same phenomenon occurs in the second application shown in the next section (see Figures 2.14 and 2.16 below).

2.5.2 Hamilton-Jacobi-Bellman PDE

We now apply the accelerated schemes to solve a Hamilton-Jacobi-Bellman (HJB) equation arising from a controlled diffusion problem with a small drift.

2.5.2.a Description of the problem We consider an HJB equation in dimension $p \geqslant 1$, where v is a real-valued function defined on the torus $\mathbb{R}^p/\mathbb{Z}^p$, identified to $[0,1]^p$, assuming a cyclic boundary condition:

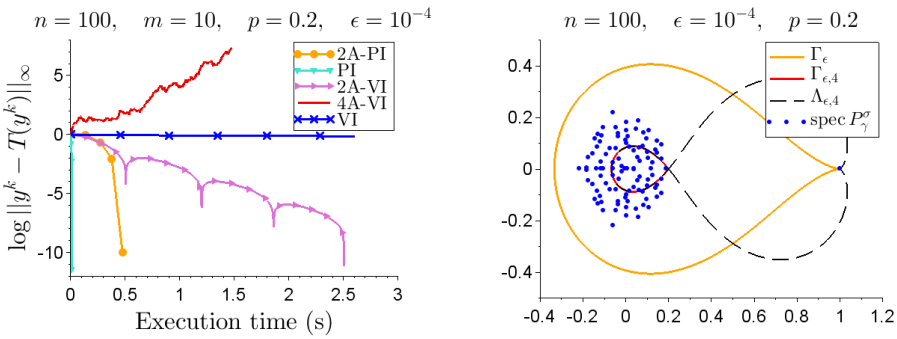
$$\max_{a\in[m]} \left(\frac{1}{2} \sum_{i=1}^{p} \sigma_i^2 \frac{\partial^2 v}{\partial x_i^2}(x) + \sum_{i=1}^{p} g_i(a,x) \frac{\partial v}{\partial x_i}(x) - \lambda v(x) + r(a,x) \right) = 0, \quad x \in [0,1]^p \ , \tag{2.62}$$



(a) PI, VI, 2A-VI and 4A-VI

(b) The spectrum of one matrix P_γ^σ where $\epsilon = 10^{-4}$.

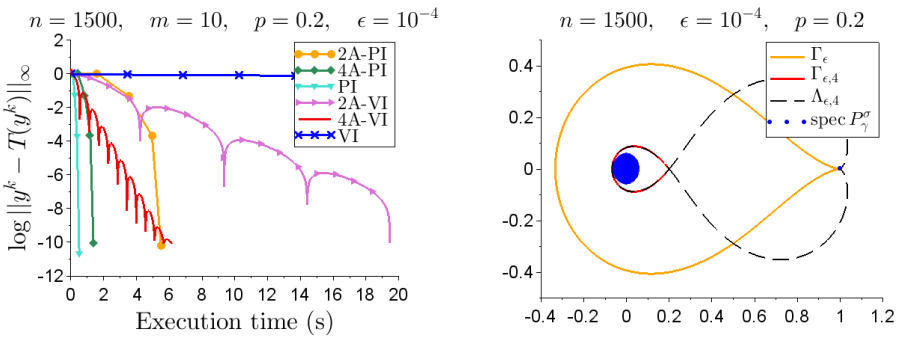
Figure (2.7) Markov Decision Process with random Markov matrices of size $n = 30$ and with $m = 10$ actions at each state.



(a) PI, 2A-PI, VI, 2A-VI and 4A-VI

(b) The spectrum of one matrix P_γ^σ where $\epsilon = 10^{-4}$.

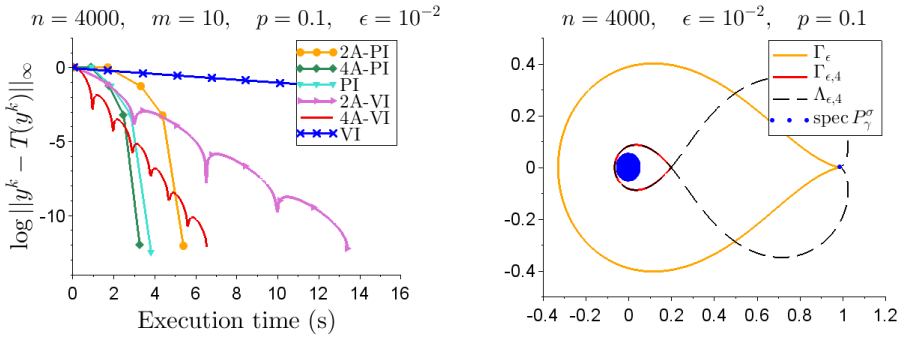
Figure (2.8) Markov Decision Process with random Markov matrices of size $n = 100$ and with $m = 10$ actions at each state.



(a) PI, 2A-PI, 4A-PI, VI, 2A-VI and 4A-VI

(b) The spectrum of one matrix P_γ^σ where $\epsilon = 10^{-4}$.

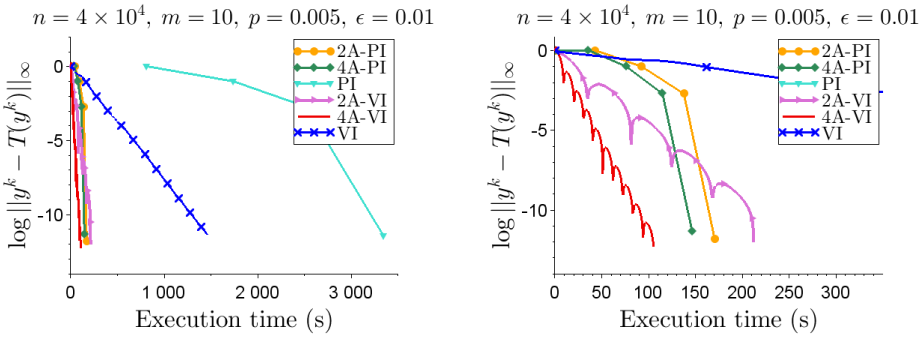
Figure (2.9) Markov Decision Process with random Markov matrices of size $n = 1500$ and with $m = 10$ actions at each state.



(a) PI, 2A-PI, 4A-PI, VI, 2A-VI and 4A-VI

(b) The spectrum of one matrix P_γ^σ where $\epsilon = 10^{-2}$.

Figure (2.10) Markov Decision Process with random Markov matrices of size $n = 4000$ and with $m = 10$ actions at each state.



(a) PI, 2A-PI, 4A-PI, VI, 2A-VI and 4A-VI

(b) Zoom on Figure 2.11(a).

Figure (2.11) Markov Decision Process with random Markov matrices of size $n = 4 \times 10^4$ and with $m = 10$ actions at each state.

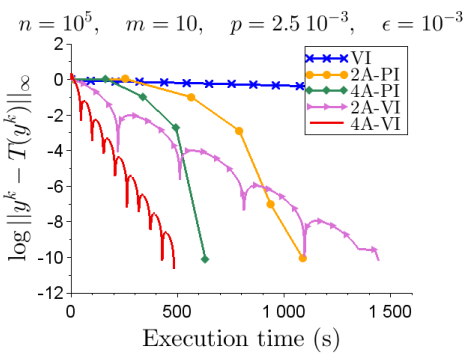


Figure (2.12) Markov Decision Process with random Markov matrices of size $n = 10^5$ and with $m = 10$ actions at each state.

where $[m] = \{1, \dots, m\}$ is the set of actions, the scalar $\sigma_i > 0$ represents the volatility in direction i, $g_i : [m] \times [0,1]^p \mapsto \mathbb{R}$ represents the drift in direction i that depends on the action a and the state x, $\lambda > 0$ is a dissipation parameter and $r : [m] \times [0,1]^p \mapsto \mathbb{R}$ is the function of rewards.

The function v is given by

$$v(x) = \sup_{a(\cdot)} \mathbb{E}\Big[\int_0^\infty \exp(-\lambda t) r(a(t), X_t) dt \mid X_0 = x \Big] \ ,$$

with $dX_t = g(a(t), X_t)dt + \mathrm{tdiag}(\sigma)dW_t$, where W_t is the standard Brownian motion on $\mathbb{R}^p$, $\mathrm{tdiag}(\sigma)$ is the diagonal matrix with entries $(\sigma_i)_{i\in[p]}$ and the supremum is taken over progressively measurable processes $a(t)$ with respect to the filtration of the Brownian motion W_t, see [FS06] for background.

For $x = (x_1, \cdots, x_p)$ and $i \in [p]$, we denote by $x_{\neq i}$ the $p-1$ entries of x that are different from i. For a scalar $g \in \mathbb{R}$, we denote $g^+ = \max(g, 0)$ and $g^- = \max(-g, 0)$.

We use a uniform grid $\Omega = \{h, 2h, \dots, Nh\}^p$ to discretize the space $[0,1]^p$, where N is a positive integer and $h = 1/N$. An upwind finite difference discretization of the HJB equation (2.62) leads to

$$\begin{aligned}\max_{a\in[m]} \Bigg(\frac{1}{2} \sum_{i=1}^p \sigma_i^2 \frac{v(x_{\neq i}, x_i + h) + v(x_{\neq i}, x_i - h) - 2v(x)}{h^2} \\ + \sum_{i=1}^p g_i(a,x)^+ \frac{v(x_{\neq i}, x_i + h) - v(x)}{h} \\ + \sum_{i=1}^p g_i(a,x)^- \frac{v(x_{\neq i}, x_i - h) - v(x)}{h} - \lambda v(x) + r(a,x) \Bigg) = 0, \quad x \in \Omega \ . \end{aligned} \tag{2.63}$$

This equation reduces to a finite dimensional dynamic programming equation of the form $V = T(V)$, with T as in (2.51), see [KKDD01] for background. We next recall this transformation, in order to apply our method.

We consider a discrete vector $V = (V_k)_{k\in[N]^p} \in \mathbb{R}^{N^p}$ such that for each index $k = (k_1, \cdots, k_p) \in [N]^p$, the kth entry of V is $V_k = v(hk)$.

The equation (2.63) can be rewritten in the following matrix form:

$$\max_{\tau\in[m]^{N^p}} (A_h^\tau V + r^\tau) = 0 \tag{2.64}$$

such that for a given policy $\tau : [N]^p \to [m]$, the matrix $A_h^\tau \in \mathbb{R}^{N^p \times N^p}$ has the kth row $(A_h)_{k\cdot}^{\tau(k)}$, $k \in [N]^p$, that represents the equation (2.63) for $x = hk \in \Omega$ and $a = \tau(k) \in [m]$, and where the vector r^τ has the kth entry $r_k^{\tau(k)} = r(\tau(k), hk)$.

We can easily see from (2.63) that the diagonal entries of each matrix A_h^τ are negative, while all the other entries are nonnegative, and this is due to the distinction of the positive and negative parts of the functions g_i that we did. We transform the problem (2.64) by introducing for each policy τ the matrix $P_h^\tau = \mathrm{I} + ch^2 A_h^\tau$, where c is a positive scalar that allows all the matrices P_h^τ to have nonnegative entries. The following lemma shows how such a scalar can be chosen.

Lemma 2.33. *If $c \leqslant c_0 := 1/(\sum_{i=1}^p \sigma_i^2 + h \max_{a\in[m],k\in[N]^p} \sum_{i=1}^p |g_i(a, hk)| + h^2\lambda)$, then for each policy τ, all the entries of the matrix P_h^τ are nonnegative.*

Moreover, we have $P_h^\tau e = (1 - ch^2\lambda)e$, where $e = (1, \cdots, 1) \in \mathbb{R}^{N^p}$, and then $\operatorname{spec} P_h^\tau \subset \mathcal{B}(0, 1-\epsilon)$ with $\epsilon = ch^2\lambda$.

Proof. By construction of P_h^τ, all its non-diagonal entries are nonnegative.

For $k \in [N]^p$, we can see from equation (2.63) that

$$(A_h^\tau)_{kk} = -\sum_{i=1}^{p} \sigma_i^2/h^2 - \sum_{i=1}^{p}(g_i(\tau(k), hk)^+ + g_i(\tau(k), hk)^-)/h - \lambda.$$

Therefore $(P_h^\tau)_{kk} = 1 - ch^2\lambda - c\sum_{i=1}^{p}\sigma_i^2 - ch\sum_{i=1}^{p}|g_i(\tau(k), hk)|$. Then if $c \leqslant c_0$, all the diagonal entries of P_h^τ are also nonnegative.

The property $P_h^\tau e = (1 - ch^2\lambda)e$ can be easily seen when we take v equal to the constant vector e in the equation (2.63), and since all the entries of P_h^τ are nonnegative, we deduce that its spectral radius is $1 - ch^2\lambda$ which ends the proof of the lemma. □

Remark 2.34. We notice that the parameter c used in the definition of P_h^τ plays the role of a Krasnosel'skiň-Mann damping (see (2.4a)). So if we divide c by 2, i.e. we take $c \leqslant c_0/2$, this ensures that all the eigenvalues of the matrix P_h^τ has a real part in the interval $[0, 1 - \epsilon]$.

Now, we can write the equation (2.63), as a fixed point problem that represents a 1-player game:

$$T(V) = V \tag{2.65}$$

where

$$T(V) = \max_{\tau \in [m]^n} (P_h^\tau V + r_h^\tau) \ .$$

with $r_h^\tau = ch^2 r^\tau$.

2.5.2.b Study of the eigenvalues for uncontrolled PDE with uniform drifts We will restrict the study of the eigenvalues of the matrices defining the problem (2.63), to the uncontrolled case where $m = 1$. We have only one matrix A_h, and $P_h = \mathrm{I} + ch^2 A_h$. We suppose also that the drift coefficients $g_i \in \mathbb{R}$ do not depend on the state x. Under this framework, we have the following lemma that gives an explicit expression of the eigenvalues of P_h.

Lemma 2.35. *The N^p eigenvalues of the matrix P_h are given as follows for each $k \in [N]^p$:*

$$\eta(k) = 1 - c\sum_{j=1}^{p}\sigma_j^2(1 - \cos(2\pi k_j h)) - c\lambda h^2 + 2ich\sum_{j=1}^{p}\sin(\pi k_j h)(g_j^+ e^{i\pi k_j h} - g_j^- e^{-i\pi k_j h}) \ .$$

Proof. For a given $k \in [N]^p$, we define the vector $V \in \mathbb{R}^{[N]^p}$ which $l \in [N]^p$ entry is given by $V_l = e^{2i\pi h\sum_{j=1}^{p} k_j l_j}$. From (2.63), we can verify that

$$(A_h V)_l = V_l\Bigg(\frac{1}{2}\sum_{j=1}^{p}\sigma_j^2\frac{e^{2i\pi h k_j} - 2 + e^{-2i\pi h k_j}}{h^2} + \sum_{j=1}^{p}\left(g_j^+\frac{e^{2i\pi h k_j} - 1}{h} - g_j^-\frac{1 - e^{-2i\pi h k_j}}{h}\right) - \lambda\Bigg). \tag{2.66}$$

Then this shows that

$$\mu(k) := \sum_{j=1}^{p}\sigma_j^2(\cos(2\pi k_j h) - 1)/h^2 - \lambda + 2i\sum_{j=1}^{p}\frac{\sin(\pi k_j h)}{h}(g_j^+ e^{i\pi k_j h} + g_j^- e^{-i\pi k_j h})$$

is an eigenvalue of the matrix A_h, and this allows to find all the N^p eigenvalues of A_h and therefore those of P_h also. □

Lemma 2.36. *The eigenvalues of the matrix P_h satisfy the following inequality:*

$$|\operatorname{Im}(\eta(k))| \leqslant \Big(\sum_{j=1}^{p} \frac{2g_j^2}{\lambda\sigma_j^2}\Big)^{\frac{1}{2}} \sqrt{\epsilon(1-\epsilon-\operatorname{Re}(\eta(k)))}, \quad k \in [N]^p \ .$$

Proof. From Lemma 2.35 and using that $g_j^+ - g_j^- = g_j$, $g_j^+ + g_j^- = |g_j|$ and $\epsilon = c\lambda h^2$, we deduce that the real and imaginary parts of the eigenvalue $\eta(k)$ are:

$$\operatorname{Im}(\eta(k)) = 2ch \sum_{j=1}^{p} g_j \sin(\pi k_j h)\cos(\pi k_j h) \ ,$$

$$\operatorname{Re}(\eta(k)) = 1 - \epsilon - 2c\sum_{j=1}^{p}(\sigma_j^2 + h|g_j|)(\sin(\pi k_j h))^2 \ .$$

By using Cauchy-Schwartz inequality, we have:

$$\sum_{j=1}^{p} |g_j \sin(\pi k_j h)| \leqslant \big(\sum_{j=1}^{p} \frac{g_j^2}{\sigma_j^2}\big)^{\frac{1}{2}} \big(\sum_{j=1}^{p} \sigma_j^2 (\sin(\pi k_j h))^2\big)^{\frac{1}{2}} \ .$$

and this implies the desired inequality. □

Recall that if the spectrum of a matrix is in the region $\Sigma_\epsilon(r)$ with the choice of r shown in Figure 2.2(b), the 2A-VI algorithm, applied to this matrix, converges with an asymptotic rate $1 - \sqrt{\epsilon}/2$ (see Remark 2.10). If follows from Lemma 2.36 that for a fixed value of ϵ, if the drift coefficients g_i are sufficiently small, the spectrum of the matrix P_h lies in a small neighborhood of the real segment $[0, 1-\epsilon]$, and so it satisfies the condition for acceleration with the latter asymptotic rate. Moreover, when ϵ is small, one can show using the same lemma that the acceleration conditions are met even for drift coefficients of order 1 (this involves a long and routine verification that we skip here). We illustrate these properties in the next section.

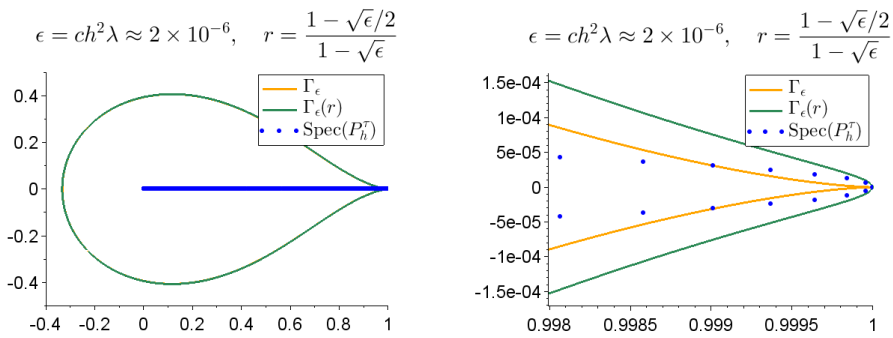
2.5.2.c Numerical results In Figures 2.13 and 2.14 we consider the HJB equation (2.62) in one dimension $p = 1$. We take the size of the discretization grid $N = 500$ with $m = 10$ actions at each state. We take the volatility $\sigma_1 = 1$ and the dissipation parameter $\lambda = 1$. We generate the drift values $g_1(a, x)$ at each state x and for each action a randomly uniformly in the interval $[0, 1]$ and we generate the rewards $r(a, x)$ randomly uniformly in $[0, 100]$. In subfigure 2.13(a) we display the spectrum of one matrix P_h^τ. The subfigure 2.13(b) shows a zoom on this spectrum around the point 1, where all the difficulty occurs. It shows that the eigenvalues of P_h^τ are not included in the peaked curve Γ_ϵ but are instead included in the more tolerant curve $\Gamma_\epsilon(r)$ with $r = (1 - \sqrt{\epsilon}/2)/(1 - \sqrt{\epsilon})$.

In Figure 2.14, we display the performance of value iteration, accelerated policy iteration and accelerated value iteration of degree 2.

Figures 2.15 and 2.16 display the analogue plots as Figures 2.13 and 2.14 with an HJB equation in dimension $p = 2$, with $N = 30$, $\sigma_1 = \sigma_2 = \sqrt{2}$, $\lambda = 2$, drifts $g_1(a, x)$ in the first direction generated uniformly randomly in $[0, 1]$, drifts $g_2(a, x)$ in the second direction generated uniformly randomly in $[-1, 0]$ and rewards $r(a, x)$ generated randomly uniformly in $[0, 100]$.

We see that for these examples the accelerated algorithms 2A-VI and 2A-PI converge and are faster than the classical Value Iteration algorithm.

We mention though that on these two examples the classical Policy Iteration algorithm is way faster than 2A-PI and 2A-VI, which is expected since the size of the matrices is small, as seen in Figures 2.7 to 2.12. However, when the size of the matrices gets bigger our iterative algorithms become faster than Policy Iteration like in the large scale example of Figure 2.11.



(a) Spectrum of P_h^τ and the accelerable regions. (b) Zoom on Figure 2.13(a) around 1.

Figure (2.13) Spectrum of the matrix P_h^τ and acceleration region, for the HJB PDE in dimension one

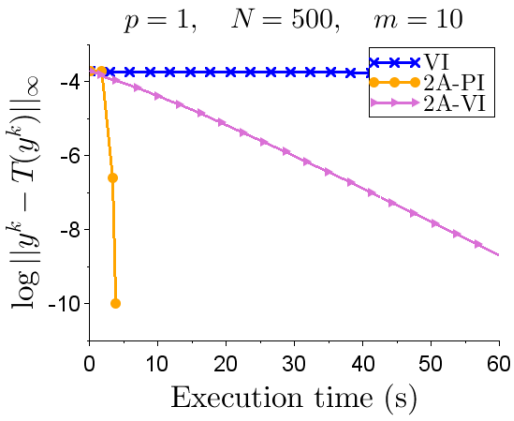
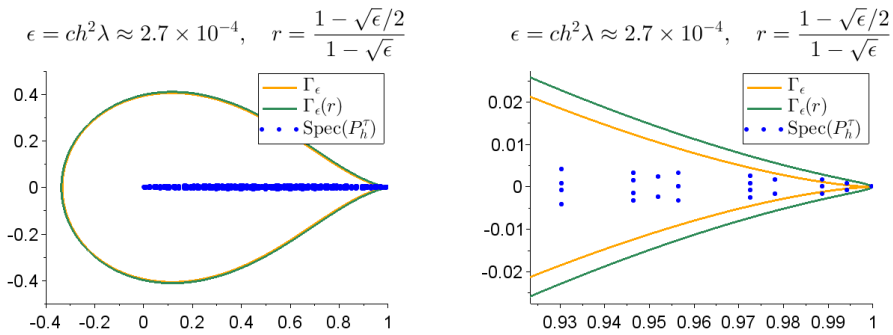


Figure (2.14) Solving HJB equation in one dimension with $N = 500$, $\lambda = 1$, $\sigma_1 = 1$, $g_1 \sim \mathcal{U}([0,1])$, $c = c_0/2 \approx 0.5$, $\epsilon = ch^2\lambda \approx 2 \times 10^{-6}$ and $r \sim \mathcal{U}([0,100])$.



(a) Spectrum of P_h^τ and the accelerable regions. (b) Zoom on Figure 2.15(a) around 1.

Figure (2.15) Spectrum of the matrix P_h^τ and acceleration region, for the HJB PDE in dimension two

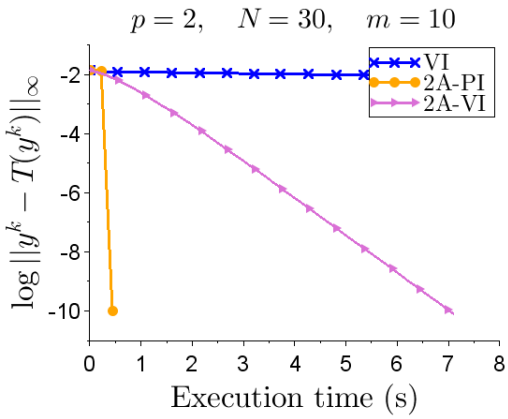


Figure (2.16) Solving HJB equation in two dimensions with $N = 30$, $\lambda = 2$, $\sigma_1 = \sigma_2 = \sqrt{2}$, $g_1 \sim \mathcal{U}([0,1])$, $g_2 \sim \mathcal{U}([-1,0])$, $c = c_0/2 \approx 0.12$, $\epsilon = ch^2\lambda \approx 2.7 \times 10^{-4}$ and $r \sim \mathcal{U}([0,100])$.

CHAPTER 3

Solving Mean Payoff Markov Decision Processes and Perfect Information Zero-sum Stochastic Games by Variance Reduced Deflated Value Iteration

Abstract

In this chapter, we introduce a deflated version of value iteration, which allows one to solve mean payoff problems, including both Markov decision processes and perfect information zero-sum stochastic games. This method requires the existence of a distinguished state which is accessible from all initial states and under all policies; it differs from the classical relative value iteration algorithm for mean payoff problems in that it does not need any primitivity or geometric ergodicity condition. Our method is based on a reduction from the mean payoff problem to a discounted problem by a Doob h-transform, combined with a deflation technique and non-linear spectral theory results (Collatz-Wielandt characterization of the eigenvalue). In this way, we provide a new method *Deflated Value Iteration* that allows to extend complexity results from the discounted to the mean payoff case. In particular, Sidford, Wang, Wu and Ye (2018) developed an algorithm combining value iteration with variance reduction techniques to sowelve

discounted Markov decision processes in sublinear time when the discount factor is fixed. We combine deflated value iteration with variance reduction techniques to obtain a sublinear algorithm for mean payoff stochastic games in which the first hitting times of a distinguished state are bounded a priori. This chapter is an extended version of the CDC conference article [AGQS19].

3.1 Introduction

Context. Markov decision processes, and more generally zero-sum two player stochastic games, are classical models to study sequential problems under uncertainty [Put14, NS03]. They appear in various applications ranging from engineering sciences, finance, economy, to health care or ecology. The dynamic programming method allows one to reduce the infinite horizon problem, in which players optimize a discounted payoff, to a fixed point problem, involving an order preserving and contracting map, called *Bellman* or *Shapley operator*. Value iteration and policy iteration [Put14] are two fundamental dynamic programming methods. For discounted problems with a fixed discount factor, value iteration allows one to find an optimal policy in a time which is polynomial [Tse90] but not strongly polynomial [FH14]. Ye showed that policy iteration runs in strongly polynomial time for a fixed discount factor [Ye11a]. This result was subsequently extended to two-player zero-sum games by Hansen, Miltersen and Zwick [HMZ13]. However, Friedmann [Fri09b, Fri11] showed that policy iteration can take an exponential time for two player games with mean payoff, and Fearnley [Fea10] showed that the same is true for Markov Decision Processes. Hence, problems with a vanishing discount and mean payoff problems are in the hardest class. However, some special mean payoff problems have been reduced to problems with a *fixed* discount factor, leading to parametrized complexity results, see [FABG13, FH13, AG13, Sch16].

Even for problems with a fixed discount, value iteration and policy iteration appear to be too slow, or unadapted, for huge scale instances. Algorithms based on Monte-Carlo simulations can lead to improved scalability. In a recent progress, Sidford et al. [SWWY18] combined value iteration algorithm with sampling and variance reduction techniques. They obtained an algorithm for discounted infinite-horizon MDPs that, remarkably, is sublinear in a certain relevant regime of the parameters. This result use in an essential way the discounted nature of the problem.

Here, our aim is to develop accelerated value iteration algorithms for well structured huge scale instances of *mean payoff* problems. To do so, we develop further a general method, first introduced in our previous work [AG13], allowing one to reduce a class of mean payoff problems to discounted problems.

Contributions. Our first main results are Theorem 3.5 and Corollary 3.6, which characterize the best contraction rate of the Shapley operator of a zero-sum game, with respect to all possible weighted sup-norms, as the Collatz-Wielandt number of a certain convex order-preserving positively homogeneous map which we call the "Clarke recession function". This is a key ingredient to obtain our subsequent complexity estimates. This is also of independent interest.

Then, we provide in Theorem 3.25 the reduction from mean payoff problems to discounted problems. This applies to the subclass of two-player games in which there is a distinguished state c to which all other states have access, for all policies of the two players. This reduction combines a scaling argument (a combinatorial version of Doob's h-transform arising in the boundary theory of Markov processes [Dyn69]) and a deflation technique: to a mean payoff problem, we associate a discounted problem, with a *state-dependent* discount rate (Theorem 3.25). To compute this reduction, we need first to solve the dynamic programming equation of a stochastic shortest path (one-player) problem, in which a player wants to maximize the expected hitting time to the distinguished state c. We obtain an explicit

contraction rate for the reduced problem in terms of the maximal expected hitting times, which appears in our complexity bounds.

This approach leads to a new algorithm to solve the mean payoff problem, that we call *deflated value iteration* (Algorithm 2). This algorithm is based on two steps, the first step is to compute the value of the stochastic shortest path problem above, and the second one is to solve the reduced problem. Both are solved by using value iteration. We also give a complexity bound in Theorem 3.28, and we compare numerically deflated value iteration with the classical relative value iteration in Section 3.6.

This reduction technique, allows us also to propose a sublinear algorithm solving mean payoff stochastic games, obtained as follows. We solve the mean payoff problem by calling twice a variant of the algorithm of Sidford's et al. [SWWY18]: we call first this variant to compute the parameters of the reduction, and we call it a second time to solve the discounted game obtained after the reduction. We also note that the present variant includes an extension of the algorithm of [SWWY18] for one player to the two-player case. However, this extension is an easier matter–the main novelty here is rather the reduction from the mean payoff problem to the discounted case and the resulting complexity bounds.

Comparison with other approaches. Our contraction results Theorem 3.5 and Corollary 3.6 imply that a contraction estimate previously computed in [AG13] is indeed optimal, if all the actions are "useful" in a natural sense.

We note also that in the case of mean payoff problems, the discount factor tends to 1 and the bounds of execution time for value iteration blow up which excludes to pass to the limit in the algorithm of [SWWY18]. Hence, we need the transformation of the mean payoff problem to a discounted one before applying value iteration.

Weighted sup-norms were already used by Bertsekas and Tsitsiklis to obtain contraction results for value iteration in the case of stochastic shortest path problems [BT91b].

Gupta, Jain and Glynn have recently developed a Monte-Carlo version of *relative value iteration* to solve mean payoff problems [GJG15]. The convergence analysis requires the Bellman operator to be a strict contraction in the "span seminorm". This is a demanding condition. For instance, in the 0-player case, this requires the transition matrix to be primitive, where our reduction holds in more general circumstances (the uniqueness of the final class suffices). Wang developed in [Wan17] an algorithm for mean payoff MDPs (one player), that has a sublinear bound. This algorithm depends on a mixing time and on a parameter τ measuring the distance between the invariant measures attached to different policies (the mixing times of [Wan17] should not be confused with the hitting times used here, the finiteness of the former implies the finiteness of the latter, but not vice versa). There are instances in which the distance τ is exponential in the input size, whereas the hitting time is linear.

Organization. This chapter is organized as follows. In Section 3.2, we recall basic notions about zero-sum games. In Section 3.3, we analyze the contraction rate of order-preserving maps with respect to weighted sup-norms. In Section 3.4, we present the main techniques allowing the reduction from the mean payoff case to the discounted case. Based on this reduction technique, we propose, in Section 3.5, a deterministic algorithm that we call *Deflated Value Iteration* (DVI) to solve mean payoff problems that satisfy our hypothesis. In Section 3.6, we give numerical results comparing DVI to the classical relative value iteration. In Section 3.7, we present an adaptation of the variance reduction algorithm of [SWWY18], allowing us to handle the operators obtained after the deflation and h-transform reduction. In Section 3.8, we derive sublinear bounds for classes of mean payoff problems. Examples are presented in Section 3.9.

3.2 Dynamic programming equations of zero-sum two-player games

3.2.1 Perfect information zero-sum stochastic games with general discount factor

We refer the reader to [NS03] for background on stochastic games. We next briefly recall the main notions and properties.

A perfect information two-player zero-sum stochastic game with general discount (SG) is described by the following data. We consider a finite state space $S := \{1, \dots, n\}$. For all $i \in S$, A_i is a finite set representing the possible actions of player MIN in state i, and $B_{i,a}$ is a finite set representing the possible actions of player MAX in state i, when player MIN just played action a. We denote by $E := \{(i, a, b) \mid i \in S, a \in A_i, b \in B_{i,a}\}$ the set of all admissible triples state-actions. For all $(i, a, b) \in E$, P_i^{ab} is an element of $\Delta(S)$ the set of probability measures on S; we shall identify P_i^{ab} to a row vector in $\mathbb{R}^n$, writing $P_i^{ab} = (P_{ij}^{ab})_{j \in S}$ where P_{ij}^{ab} is the transition probability to the next state j, given the current state i and the actions taken $a \in A_i, b \in B_{i,a}$. For all $(i, a, b) \in E$, r_i^{ab} is a reward (real number) that MIN pays to MAX, and γ_i^{ab} (real nonnegative number) is a discount factor. We define

$$R := \max_{(i,a,b) \in E} |r_i^{ab}| \in \mathbb{R}_{\geqslant 0}, \quad \Gamma := \max_{(i,a,b) \in E} \gamma_i^{ab} \in (0, \infty) \ ,$$

where $\mathbb{R}_{\geqslant 0} := \{x \in \mathbb{R} \mid x \geqslant 0\}$. We allow γ_i^{ab} to take values larger than 1. The term *turn-based* is sometimes used as a synonym of "perfect information". This is in contrast with the more general model of Shapley's imperfect information stochastic games in which two players play simultaneously with randomized actions, see e.g. [MN81a].

Recall that a *strategy* of a player is a decision rule which associates to a history of the game an admissible action of this player. A strategy σ of player Min, a strategy τ of player Max, and an initial state i, alltogether determine a random process $(i_\ell, a_\ell, b_\ell)_{\ell \geqslant 0}$ with values in E: i_ℓ represents the state at step ℓ, and a_ℓ, b_ℓ represent the actions of the two players at the same state. We require that $i_0 = i$. We denote by $\mathbb{E}_{i,\sigma,\tau}$ the expectation with respect to the probability measure governing this process. Given a finite horizon k, we consider the zero-sum game in which the payoff of player Max is given by

$$J_i^k(\sigma, \tau) = \mathbb{E}_{i,\sigma,\tau} \left(\sum_{\ell=0}^{k-1} \left(\prod_{m=0}^{\ell-1} \gamma_{i_m}^{a_m b_m} \right) r_{i_l}^{a_\ell b_\ell} \right) \ . \tag{3.1}$$

The *value* v_i^k of the k-stage game starting from i (see Proposition III.4.2. and Theorem IV.3.2 in [MSZ15b]) is defined as,

$$v_i^k := \inf_\sigma \sup_\tau J_i^k(\sigma, \tau) = \sup_\tau \inf_\sigma J_i^k(\sigma, \tau) \ , \tag{3.2}$$

where the infima and suprema are taken over the set of strategies of both players. By definition, the existence of the value requires the infimum and the supremum to commute. A pair of strategies σ^*, τ^* is said to be *optimal* if σ^* achieves the first infimum in (3.2) and if τ^* achieves the second supremum in (3.2).

For all $(i, a, b) \in E$, we set $M_{ij}^{ab} := \gamma_i^{ab} P_{ij}^{ab}$ and $M_i^{ab} := (M_{ij}^{ab})_{j \in S} \in \mathbb{R}^n$.

Definition 3.1. For a given SG, the Shapley operator T is the map $\mathbb{R}^n \to \mathbb{R}^n$ whose ith coordinate is given by

$$T_i(v) = \min_{a \in A_i} \max_{b \in B_{i,a}} \{r_i^{ab} + \sum_{j \in S} M_{ij}^{ab} v_j\}, \quad i \in S, \ v \in \mathbb{R}^n \ .$$

The following result is classical.

Theorem 3.2. *The value vector $v^k = (v_i^k)_{i\in S}$ does exist and satisfies*

$$v^k = T(v^{k-1}) \ , \qquad v^0 = 0 \ . \tag{3.3}$$

This follows from the proof of Theorem IV.3.2 of [MSZ15b], which deals with the undiscounted finite horizon problem. The extension to the present situation (with a state-dependent discounted factor) is straightforward. Note that the assumption that the discount factor be smaller than 1 is not needed for the well posedness of the finite horizon problem and for the validity of (3.3).

Similarly, one can consider the *infinite horizon discounted zero-sum game*, in which the payoff of player Max is now

$$J_i(\sigma,\tau) = \mathbb{E}_{i,\sigma,\tau}\left(\sum_{\ell=0}^{\infty}(\prod_{m=0}^{\ell-1}\gamma_{i_m}^{a_m b_m})r_{i_\ell}^{a_\ell b_\ell}\right) \ .$$

This payment is well defined, in particular, when, $\Gamma < 1$ since then the above series become absolutely convergent. Then, the value of the infinite horizon game and the notion of optimal strategies are defined in a similar manner to the finite horizon case. The value vector $v = (v_i)_{i\in S}$ does exist and it is characterized as the unique solution of the fixed point problem

$$v = T(v) \ ,$$

see again [MSZ15b] for background. We shall see later on that the assumption $\Gamma < 1$ can be relaxed: what matters is that the discount factor be smaller than one in an "average" sense.

In what follows, it will be convenient to consider a special class of strategies, determined by *policies* (feedback, stationary rules). A *policy* of player MIN is a map:

$$\sigma : S \to \cup_{i\in S} A_i \ , \quad i \mapsto \sigma(i) \in A_i \ .$$

We denote by $\mathscr{S}$ the set of all policies of player MIN. Similarly, a *policy* of MAX is a map:

$$\tau : \cup_{i\in S}(i, A_i) \to \cup_{i\in S,\, a\in A_i} B_{i,a} \ , (i,a) \mapsto \tau(i,a) \in B_{i,a} \ .$$

Note that since the game is in perfect information, MAX observes the action a of MIN, and so the policy of player MAX takes care of this action. We denote by $\mathscr{T}$ the set of all policies of player MAX. It is known that in the discounted game, there exist optimal strategies associated to policies (the action is selected at each step by applying a policy of one player, the policy being the same for all time steps). These policies are obtained by selecting actions achieving the minimum and the maximum in the expression of $T(v)$ in Definition 3.1. See [NS03].

Any choice of policies $(\sigma,\tau) \in \mathscr{S} \times \mathscr{T}$ defines the Markovian matrix $P^{\sigma\tau} \in \mathbb{R}^{n\times n}$ which determines the state trajectory if the two players select their actions according to these policies. i.e., $(P^{\sigma\tau})_{ij} = P_{ij}^{\sigma(i)\tau(i,\sigma(i))}$. Similarly, we define the nonnegative matrix $M^{\sigma\tau}$ with entries $(M^{\sigma\tau})_{ij} = M_{ij}^{\sigma(i)\tau(i,\sigma(i))}$. We denote the cardinality of a finite set $\mathcal{S}$ by $|\mathcal{S}|$. We recall that the size of the input is of order $|S||E|$.

3.2.2 Mean payoff problem

We are now interested in the *undiscounted* case, in which the discount factor γ is identically 1. Then we are considering a two-player perfect information zero-sum Mean-Payoff Stochastic Game (MPSG), where the main quantity of interest is the *mean payoff vector*:

$$\chi(T) := \lim_{k\to\infty} T^k(0)/k \ .$$

The entry $\chi_i(T)$ represents the mean payoff per time unit, if the initial state is i. Here, the mean payoff is defined by considering a family of games in finite horizon k as k tends to infinity. There are alternative approaches, in which the mean payoff is defined as the value of an infinite horizon game [LL69a]. The property of the *uniform value* established in [MN81a] entails that the different natural approaches lead to the same notion of mean payoff.

The analysis of the mean payoff problem is simplified when the following *non-linear eigenproblem* has a solution:

$$\eta e + v = T(v), \qquad \eta \in \mathbb{R},\ v \in \mathbb{R}^n \ , \tag{3.4}$$

where $e := (1 \cdots 1)^\top \in \mathbb{R}^n$ is the unit vector. The scalar η is called the *ergodic constant*, whereas the vector v, which is not unique, is called *bias* or *potential*. When this equation is solvable, we have $\chi(T) = \eta e$, i.e., the mean payoff is independent of the initial state, and it is equal to the ergodic constant. See e.g. [AGH18] for background.

3.3 Contraction rate of order-preserving maps with respect to weighted sup-norms

Let $u \in \mathbb{R}^n$, we write $u \gg 0$ and we say that u is a *positive vector* if for all $i \in [n] := \{1, \cdots, n\}$, $u_i > 0$. Given $u \gg 0$, we define the *weighted sup norm* $\|\cdot\|_u$ by :

$$\|x\|_u = \max_{1 \leqslant i \leqslant n} \frac{x_i}{u_i} = \|u^{-1}x\|_\infty \,, \forall x \in \mathbb{R}^n \ ,$$

where the notation $u^{-1}x := (u_i^{-1}x_i)_{i\in[n]}$ refers to the Hadamard quotient. For x, $y \in \mathbb{R}^n$ we write $x \leqslant y$ if $x_i \leqslant y_i$ for all $i \in [n]$. A function $f : \mathbb{R}^n \to \mathbb{R}^n$ is said to be *order-preserving* if for all x, $y \in \mathbb{R}^n$, if $x \leqslant y$ then $f(x) \leqslant f(y)$.

We next introduce a notion of recession function associated to a non-linear map. Our definition is inspired by the notion of Clarke generalized directional derivative [CLSW98, Ch. 2, S1] of a function f at point z in the direction y

$$f'_z(y) := \limsup_{x \to z,\, s \to 0^+} \frac{f(x + sy) - f(x)}{s} \ . \tag{3.5}$$

We next adapt this idea by considering "variations at infinity" instead of local variations.

Definition 3.3. Given a function $f : \mathbb{R}^n \to \mathbb{R}^n$, we define $\hat{f} : \mathbb{R}^n \to (\mathbb{R} \cup \{+\infty\})^n$ the *Clarke recession function* of f as:

$$\hat{f}(y) = \sup_{s>0,\, x\in\mathbb{R}^n} \frac{f(x + sy) - f(x)}{s}, \quad y \in \mathbb{R}^n \ . \tag{3.6}$$

We chose the name "Clarke recession function" in view of the similarity between (3.6) and (3.5).

The following result is immediate:

Proposition 3.4. *The Clarke recession function is positively homogeneous and convex.*

Theorem 3.5. *Let $f : \mathbb{R}^n \to \mathbb{R}^n$ be an order-preserving function, $u \gg 0$ be a positive vector, and $\lambda \in \mathbb{R}_{\geqslant 0}$. We have $\hat{f}(u) \leqslant \lambda u$ if and only if the function f is λ-contracting in the weighted sup-norm $\|\cdot\|_u$:*

$$\forall x, y \in \mathbb{R}^n, \qquad \|f(x) - f(y)\|_u \leqslant \lambda \|x - y\|_u \ .$$

Proof. Suppose that $\hat{f}(u) = \sup_{x\in\mathbb{R}^n,\, s>0} \frac{f(x+su)-f(x)}{s} \leqslant \lambda u$. Then, for all $x \in \mathbb{R}^n$ and $s > 0$ we have $f(x + su) \leqslant f(x) + s\lambda u$, and by considering $x - su$ instead of x we have also that $f(x - su + su) \leqslant f(x - su) + s\lambda u$, so $f(x) - s\lambda u \leqslant f(x - su)$ for all $x \in \mathbb{R}^n$ and $s > 0$.

Let $x, y \in \mathbb{R}^n$ such that $x \neq y$, we consider $s = \|x - y\|_u > 0$, we have $y - su \leqslant x \leqslant y + su$ and from the monotonicity of f and the previous inequalities, we deduce that $f(y)-s\lambda u \leqslant f(y-su) \leqslant f(x)$ and dually $f(x) \leqslant f(y + su) \leqslant f(y) + s\lambda u$. Then $-s\lambda \leqslant u^{-1}(f(x) - f(y)) \leqslant s\lambda$, and therefore $\|f(x) - f(y)\|_u \leqslant \lambda s = \lambda\|x - y\|_u$.

Now we suppose that f is λ-contracting in the norm $\|\cdot\|_u$. Let $x \in \mathbb{R}^n$ and $s > 0$ then $\|u^{-1}(f(x + su) - f(x))\|_\infty \leqslant \lambda\|u^{-1}(x + su - x)\|_\infty = \lambda s$, so $u^{-1}(f(x + su) - f(x)) \leqslant \lambda s$. Therefore $\frac{f(x+su)-f(x)}{s} \leqslant \lambda u$, and then $\hat{f}(u) \leqslant \lambda u$. □

Following [MPN02, AG13], we define the *Collatz-Wielandt number* of $\hat{f}$ as

$$\mathrm{cw}(\hat{f}) := \inf\{\lambda > 0 \mid \exists u \gg 0; \hat{f}(u) \leqslant \lambda u\} \ .$$

As an immediate consequence of Theorem 3.5, we get:

Corollary 3.6. *If $f : \mathbb{R}^n \to \mathbb{R}^n$ is order-preserving, then:*

$$\mathrm{cw}(\hat{f}) = \inf\{\lambda \in \mathbb{R}_{\geqslant 0} \mid \exists u \gg 0; f \textit{ is } \lambda\textit{-contracting in } \|\cdot\|_u\}.$$

We consider the Shapley operator $T : \mathbb{R}^n \to \mathbb{R}^n$ of Definition 3.1. The following "max-max" operator $T^{\max} : \mathbb{R}^n \to \mathbb{R}^n$ was considered in [AG13]

$$T_i^{\max}(y) = \max_{a\in A_i, b\in B_{i,a}} \{M_i^{ab} y\} \ , \forall i \in S, y \in \mathbb{R}^n \ . \tag{3.7}$$

In contrast to the Clarke recession function $\hat{T}$, $T^{\max}$ generally depends on the choice of the representation of T as a minimax expression. We next show, however, that $T^{\max} = \hat{T}$ if all the terms arising in the minimax expression are "useful" in the following sense.

Definition 3.7. For a given couple of actions (a, b) of the two players, we define the set $C_i^{ab} = \{x \in \mathbb{R}^n \mid T_i(x) = r_i^{ab} + M_i^{ab} x\}$. We say that the couple of actions (a, b) is *useful* if $\mathrm{int}(C_i^{ab}) \neq \emptyset$ for all $i \in [n]$.

In the one player case, checking whether one action is useful reduces to checking whether a polyhedron has a non-empty interior, and this can be done in polynomial time.

Lemma 3.8. *The Clarke recession function of the Shapley operator T satisfies the following inequality:*

$$\hat{T}(y) \leqslant T^{\max}(y) \ , \forall y \in \mathbb{R}^n \ . \tag{3.8}$$

Moreover, the equality holds if all the actions (a, b) are useful.

Proof. Let $i \in S, x, y \in \mathbb{R}^n$ and $s > 0$, we have

$$\begin{aligned} T_i(x + sy) = \min_{a\in A_i} \max_{b\in B_{i,a}} \{r_i^{ab} + M_i^{ab}(x + sy)\} &\leqslant \min_{a\in A_i} \max_{b\in B_{i,a}} \{r_i^{ab} + M_i^{ab} x + s \max_{a\in A_i, b\in B_{i,a}} \{M_i^{ab} y\}\} \\ &\leqslant \min_{a\in A_i} \max_{b\in B_{i,a}} \{r_i^{ab} + M_i^{ab} x\} + s \max_{a\in A_i, b\in B_{i,a}} \{M_i^{ab} y\} \\ &= T_i(x) + s \max_{a\in A_i, b\in B_{i,a}} \{M_i^{ab} y\} \ . \end{aligned}$$

Therefore $\frac{T_i(x+sy)-T_i(x)}{s} \leqslant \max_{a\in A_i, b\in B_{i,a}}\{M_i^{ab}y\}$ which proves the desired inequality. Now we suppose that all the actions (a,b) are useful and we show the opposite inequality. Let $y \in \mathbb{R}^n, i \in S, a \in A_i, b \in B_{i,a}$. Since (a,b) is useful, we can consider $x \in \operatorname{int}(C_i^{ab})$. So there exists $s > 0$ such that $x + sy \in C_i^{ab}$. Then by definition of C_i^{ab}, we have $T_i(x+sy) = M_i^{ab}(x+sy)$ and $T_i(x) = M_i^{ab}x$, then $\frac{T_i(x+sy)-T_i(x)}{s} = M_i^{ab}y$, therefore $\hat{T}_i(y) \geqslant M_i^{ab}y$ and this is for all $i \in S, a \in A_i, b \in B_{i,a}$ which allows to conclude. □

Let f be a continuous positively homogeneous map from $C = \mathbb{R}^n_{\geqslant 0}$ to itself ($f(\lambda v) = \lambda f(v)$ for all $\lambda > 0$ and $v \in C$). The following definitions are taken from [AG13] and [MPN02].

We define the *cone eigenvalue spectral radius* of f which is the supremum of its eigenvalues in the closed convex cone $C = \mathbb{R}^n_{\geqslant 0}$ by $\hat{r}_C(f) := \sup\{\lambda \geqslant 0 \mid \exists v \in C \setminus \{0\}; f(v) = \lambda v\}$.

Theorem 3.9 (Th. 3.1 in [Nus86b])**.** *If f is a continuous, positively homogeneous and order preserving selfmap of $C = \mathbb{R}^n_{\geqslant 0}$ then*

$$\mathrm{cw}_C(f) = \hat{r}_C(f) \ .$$

There is an explicit formula for the Collatz-Wielandt number of $T^{\max}$. Recall that the notation $M^{\sigma\tau}$ refers to the nonnegative matrix associated to a pair of policies (end of Section 3.2.1). We denote by $\rho(\cdot)$ the spectral radius of a matrix.

Theorem 3.10. *We have*

$$\mathrm{cw}(T^{\max}) = \max_{\sigma\in\mathscr{S},\tau\in\mathscr{T}} \rho(M^{\sigma\tau}) \ .$$

Proof. We verify easily that the $T^{\max}$ is a selfmap of $C = \mathbb{R}^n_{\geqslant 0}$, *continuous*, *positively homogeneous* ($T^{\max}(\lambda v) = \lambda T^{\max}(v)$ for all $\lambda \geqslant 0$ and $v \in C$), and *order preserving* ($T^{\max}(u) \leqslant T^{\max}(v)$ for all $u, v \in C$ such that $u \leqslant v$). Therefore, by Theorem 3.9 we have the equality $\mathrm{cw}_C(T^{\max}) = \hat{r}_C(T^{\max})$. We have $T^{\max}(v) = \max_{\sigma\in\mathscr{S},\tau\in\mathscr{T}} M^{\sigma\tau}v$, $\forall v \in C$, then $\hat{r}_C(T^{\max}) = \max_{\sigma\in\mathscr{S},\tau\in\mathscr{T}} \rho(M^{\sigma\tau})$, this is established in Proposition 19 in [AG13] and also in an infinite dimensional context in [AGN11], which yields the result. □

Remark 3.11. Owing to Theorem 3.5 and Lemma 3.8, it suffices to look for a vector $\varphi \gg 0$ such that $T^{\max}(\varphi) \leqslant \lambda\varphi$ for some $\lambda \in [0,1)$, to have that the Shapley operator T is $\lambda-$contracting in the weighted norm $\|\cdot\|_\varphi$.

The following special construction allows us to obtain such a φ by solving a non-linear eigenproblem.

Theorem 3.12 (Th. 7 and proof of Th. 13 in [AG13])**.** *The following assertions are equivalent:*

1. $\max_{\sigma\in\mathscr{S},\tau\in\mathscr{T}} \rho(M^{\sigma\tau}) < 1$;
2. *there exists a unique vector $\varphi \in \mathbb{R}^n_{\geqslant 0}$ such that $\varphi = e + T^{\max}(\varphi)$.*

When these assumptions are satisfied, T is λ-contracting in the weighted norm $\|\cdot\|_\varphi$, with $\lambda := 1 - 1/\|\varphi\|_\infty$.

Proof. If $\max_{\sigma\in\mathscr{S},\tau\in\mathscr{T}} \rho(M^{\sigma\tau}) < 1$ then by Theorem 3.10, we have $\mathrm{cw}_C(T^{\max}) < 1$ then by definition of $\mathrm{cw}_C(T^{\max})$ we deduce the existence of $u \in \operatorname{int}(C) = \mathbb{R}^n_{>0}$ and $\mu \in [0,1)$ such that $T^{\max}(u) \leqslant \mu u$. By Lemma 3.8, we have $\widehat{T^{\max}}(u) \leqslant (T^{\max})^{\max}(u) = T^{\max}(u) \leqslant \mu u$, therefore by Theorem 3.5 we deduce that $T^{\max}$ is $\mu-$contracting under the norm $\|\cdot\|_u$. Then also the self map $v \mapsto e + T^{\max}(v)$ of $C = \mathbb{R}^n_{\geqslant 0}$ is also $\mu-$contracting and therefore there exists a unique vector $\varphi \in \mathbb{R}^n_{\geqslant 0}$ such that

$\varphi = e + T^{\max}(\varphi)$. Then we deduce that $\varphi \geqslant e$ and $T^{\max}(\varphi) = \varphi - e \leqslant \varphi - \frac{1}{\|\varphi\|_\infty}\varphi = \lambda\varphi$ with $\lambda = 1 - 1/\|\varphi\|_\infty$. Therefore by Lemma 3.8, we have $\hat{T}(\varphi) \leqslant T^{\max}(\varphi) \leqslant \lambda\varphi$. Then by Theorem 3.5, T is $\lambda-$contracting under the norm $\|\cdot\|_\varphi$.

Conversely, the existence of such a vector $\varphi \in \mathbb{R}^n_{\geqslant 0}$ implies that $\mathrm{cw}_C(T^{\max}) \leqslant \lambda < 1$, and by Theorem 3.10, we get $\max_{\sigma\in\mathscr{S},\tau\in\mathscr{T}} \rho(M^{\sigma\tau}) = \mathrm{cw}_C(T^{\max}) < 1$. □

3.4 Reduction from a mean payoff problem to a discounted one

We consider here the non-linear eigenproblem (3.4), where T is the Shapley operator in the undiscounted case, and describe a technique introduced in [AG13] to reduce this equation to a fixed point equation of a contracting operator. Recall that (3.4) allows one to solve the mean payoff problem. As noted above, the vector v solution of (3.4) is not unique. In particular, if v is a solution then $v + \alpha e$ also yields a solution for all $\alpha \in \mathbb{R}$. Hence, we shall distinguish a special state $c \in S$ and require $v_c = 0$.

Definition 3.13. For a Markov matrix P and states i, j, we denote:

$$\mathcal{H}_{ij}(P) := \mathbb{E}_P[\inf\{k \geqslant 1 \mid X_k = j\} \mid X_0 = i]$$

the expected first hitting time of state j, for a Markov chain X_k with transition matrix P and initial state i.

For given states i, j, we say that j is accessible from i, under the transition matrix P, if $\mathcal{H}_{ij}(P) < +\infty$, which is equivalent to the existence of a sequence of states $i_1, \cdots, i_k$ such that $P_{ii_1}P_{i_1i_2}\cdots P_{i_kj} > 0$.

The following proposition is straightforward.

Proposition 3.14. *Given a state $c \in S$, we have $\mathcal{H}_{ic}(P) < +\infty$ for all $i \in S$ if and only if P has a unique final (recurrent) class and c belongs to this class.*

A state c with the latter property is called a *renewal* state.

Definition 3.15. For any matrix $P \in \mathbb{R}^{n\times n}$, we denote by $P_{(c)} \in \mathbb{R}^{n\times n}$ the matrix obtained from P by replacing the column c of P with zeros. We denote by P_i the ith row of P, so that $P_i = (P_{ij})_{j\in[n]}$, and we use a similar notation for matrices constructed from P, e.g., $P_{(c)i} = ((P_{(c)})_{ij})_{j\in[n]}$. We define the operator $T_{(c)} : \mathbb{R}^n \mapsto \mathbb{R}^n$ such that

$$T_{(c)}(v) := e + \max_{\sigma\in\mathscr{S},\tau\in\mathscr{T}}[P^{\sigma\tau}_{(c)}v], \quad \forall v \in \mathbb{R}^n. \tag{3.9}$$

Lemma 3.16. *For a Markov matrix P, the vector $(\mathcal{H}_{ic}(P))_{i\in[n]}$ is a fixed point of the operator $v \mapsto e + P_{(c)}v$.*

Proof. For $i \in [n]$, we use Markov property to show that $\mathcal{H}_{ic}(P) = 1 + \mathbb{E}_P[\inf\{k \geqslant 0 \mid X_{k+1} = c\} \mid X_0 = i] = 1 + \sum_{j\in[n],j\neq c}\mathbb{E}_P[\inf\{k \geqslant 1 \mid X_{k+1} = c\} \mid X_1 = j]\mathbb{P}(X_1 = j \mid X_0 = i) = 1 + \sum_{j\in[n],j\neq c} P_{ij}\mathbb{E}_P[\inf\{k \geqslant 1 \mid X_k = c\} \mid X_0 = j] = 1 + P_{(c)}v$. □

Lemma 3.17. *Let $c \in S$ be a given state. The following assertions are equivalent:*

1. *$\forall i \in S, \quad \mathcal{H}_{ic} := \max_{\sigma\in\mathscr{S},\tau\in\mathscr{T}} \mathcal{H}_{ic}(P^{\sigma\tau}) < +\infty$;*
2. *For all couple of policies $(\sigma, \tau) \in \mathscr{S} \times \mathscr{T}$, $P^{\sigma\tau}$ has a unique final class, and the state c is common to each of these final classes;*

3. $\max_{\sigma\in\mathscr{S},\tau\in\mathscr{T}} \rho(P^{\sigma\tau}_{(c)}) < 1$;

4. *There exists a unique vector φ^* solution to the following dynamical equation:*

$$\varphi^* = T_{(c)}(\varphi^*) \ , \tag{3.10}$$

Under these assumptions, we have $\varphi^*_i = \mathcal{H}_{ic}$, for all $i \in S$.

Proof. The equivalence between the first two assertions follows from Proposition 3.14. The equivalence between the assertions 3 and 4 follows from Theorem 3.12. If we denote by $v^{\sigma\tau}$ the fixed point of the operator $v \mapsto e+P^{\sigma\tau}_{(c)}v$, then the fixed point of the operator $T_{(c)}$ is $\varphi^* := \max_{\sigma\tau} v^{\sigma\tau}$, and by Lemma 3.16 we have that $v^{\sigma\tau}_i = \mathcal{H}_{ic}(P^{\sigma\tau})$ for all $i \in S$. Therefore $\varphi^*_i = \mathcal{H}_{ic}$ for all $i \in S$, which proofs the equivalence between the assertions 4 and 1. □

Lemma 3.18. *The map $T_{(c)}$ is order preserving and contracting in the norm $\|\cdot\|_{\varphi^*}$ with the rate $1 - 1/\|\varphi^*\|_\infty$. It follows that if $w \in \mathbb{R}^n_{\geqslant 0}$, verifies $w \geqslant T_{(c)}(w)$ then $w \geqslant \varphi^*$. Similarly, if $w \leqslant T_{(c)}(w)$ then $w \leqslant \varphi^*$.*

Proof. The map $T_{(c)}$ is clearly order preserving, and considering its max-max operator we get

$$T^{\max}_{(c)}(\varphi^*) = T_{(c)}(\varphi^*) - e = \varphi^* - e \leqslant (1 - 1/\|\varphi^*\|_\infty)\varphi^*.$$

Then by Remark 3.11, we deduce that $T_{(c)}$ is contracting in the norm $\|\cdot\|_{\varphi^*}$ with the rate $1 - 1/\|\varphi^*\|_\infty$. Now, if a vector $w \in \mathbb{R}^n_{\geqslant 0}$ satisfies $w \geqslant T_{(c)}(w)$, then by applying k times the order preserving operator $T_{(c)}$, we obtain $w \geqslant (T_{(c)})^k(w)$, and since $T_{(c)}$ is contracting $(T_{(c)})^k(w)$ converges to the unique fixed point φ^* when k goes to infinity, then $w \geqslant \varphi^*$. We deal similarly with the case when $w \leqslant T_{(c)}(w)$. □

In the rest of this section, we make the following assumption.

Assumption A. There exists a state $c \in S$ satisfying the conditions of Lemma 3.17, i.e. the state c is accessible in finite expected time from all states in S and under all policies.

We call such a state c a *deflation state*. We can find such a state if it exists, or certify that there is none, in sub-quadratic time by using directed hypergraphs techniques, details are given in the Section 3.11.

Let $\varphi \in \mathbb{R}^n_{\geqslant 0}$, $\varphi \gg 0$, and $\mathbb{R}^n_c := \{x \in \mathbb{R}^n \mid x_c = 0\}$.

Definition 3.19. For a nonnegative matrix $P \in \mathbb{R}^{n\times n}$, if $\varphi_i \geqslant 1 + P_{(c)i}\varphi, \forall i \in S$, then we denote by $P_{(c,\varphi)}$ the nonnegative matrix obtained from P by replacing the column c by the vector $\varphi_c^{-1}(\varphi - 1 - P_{(c)}\varphi)$.

The following two lemmas are immediate:

Lemma 3.20. *Let $\eta \in \mathbb{R}$, $v \in \mathbb{R}^n$ with $v_c = 0$ and $P \in \mathbb{R}^{n\times n}$. We have $\eta(\varphi-1)+Pv = P_{(c,\varphi)}(\eta\varphi+v)$. In particular $P_{(c,\varphi)}\varphi = \varphi - e$.*

Lemma 3.21. *The map $L_\varphi : (\eta, v) \mapsto w = \eta + \varphi^{-1}v$ from $\mathbb{R} \times \mathbb{R}^n_c$ to $\mathbb{R}^n$, is an isomorphism, with inverse given by $w \mapsto (\eta, v)$ with $\eta = w_c$ and $v = \varphi(w - w_c)$.*

Definition 3.22. For any self-map f of $\mathbb{R}^n$, we denote by $\mathcal{L}_\varphi(f)$ the self-map of $\mathbb{R}^n$, such that for all $w, v \in \mathbb{R}^n$ and $\eta \in \mathbb{R}$ with $v_c = 0$ and $w = \eta + \varphi^{-1}v$, we have

$$\mathcal{L}_\varphi(f)(w) = \varphi^{-1}(\eta(\varphi - 1) + f(v)) \ .$$

Remark 3.23. For a matrix P we have, by Lemma 3.20, $\mathcal{L}_\varphi(P)(w) = P_\varphi w$ where $P_\varphi \in \mathbb{R}^{n\times n}$ is the nonnegative matrix given by $P_{\varphi,ij} := \varphi_i^{-1} P_{(c,\varphi)ij}\varphi_j, (i,j) \in S^2$, so that

$$P_{\varphi,ij} = \begin{cases} \varphi_i^{-1} P_{ij}\varphi_j, \text{ if } j \neq c,\ i \in S \\ 1 - \varphi_i^{-1} - \sum_{k\neq c} \varphi_i^{-1} P_{ik}\varphi_k, \text{ if } j = c,\ i \in S \end{cases} .$$

The construction of the matrix P_φ is inspired by Doob's h-transform, see for example [Dyn69].

We consider the Shapley operator in the undiscounted case

$$T_i(v) = \min_{a\in A_i} \max_{b\in B_{i,a}} \{r_i^{ab} + P_i^{ab} v\}, \quad \forall i \in S, \forall v \in \mathbb{R}^n. \tag{3.11}$$

By Lemma 3.17, we know that there exists a vector $\varphi \in \mathbb{R}^n_{\geqslant 0}$, such that

$$\varphi \geqslant T_{(c)}(\varphi) \ . \tag{3.12}$$

So we can define as above an order-preserving operator $T^\varphi := \mathcal{L}_\varphi(T)$, and we verify easily that

$$T_i^\varphi(w) = \min_{a\in A_i} \max_{b\in B_{i,a}} \{\varphi_i^{-1} r_i^{ab} + P_{\varphi,i}^{ab} w\}, \quad \forall i \in S, \forall w \in \mathbb{R}^n \ . \tag{3.13}$$

Lemma 3.24. *If a vector $\varphi \in \mathbb{R}^n_{\geqslant 0}$ satisfies Equation* (3.12)*, then T^φ is $\lambda_\varphi-$contracting in the sup-norm $\|\cdot\|_\infty$, with $\lambda_\varphi := 1 - 1/\|\varphi\|_\infty$. T^φ can be interpreted as a Shapley operator of a discounted game with discount factors $\leqslant \lambda_\varphi$.*

Proof. The max-max operator associated to the operator T^φ is given by

$$(T^\varphi)_i^{\max}(w) = \max_{a\in A_i, b\in B_{i,a}} \{P_{\varphi,i}^{ab} w\},\ \forall i \in S.$$

We have $P_{\varphi,i}^{ab} e = \sum_{j\in S} P_{\varphi,ij}^{ab} = 1 - \varphi_i^{-1} \leqslant \lambda_\varphi$. We deduce that $(T^\varphi)^{\max} e \leqslant \lambda_\varphi e$. Therefore by Theorem 3.5 and Lemma 3.8 we deduce that T^φ is $\lambda_\varphi-$contracting in the weighted norm $\|\cdot\|_e$ which is the sup-norm $\|\cdot\|_\infty$. □

Theorem 3.25. *Let $\varphi \in \mathbb{R}^n_{\geqslant 0}$ be a vector satisfying Equation* (3.12)*. The non-linear eigenproblem*

$$\eta e + v = T(v) \ , \tag{3.14}$$

where $\eta \in \mathbb{R}$ and $v \in \mathbb{R}^n$ with $v_c = 0$, can be reduced to the fixed point problem:

$$T^\varphi(w) = w \ , \tag{3.15}$$

where $w \in \mathbb{R}^n$ such that $w = \eta + \varphi^{-1} v$. Equation (3.15) *has a unique solution w^*.*

Proof. By replacing $v = \varphi(w - \eta e)$ in Equation (3.14), we get

$$\eta e + \varphi(w - \eta e) = \min_{\sigma\in\mathscr{S}} \max_{\tau\in\mathscr{T}} \{r^{\sigma\tau} + P^{\sigma\tau}\varphi w - \eta P^{\sigma\tau}\varphi\}.$$

Therefore

$$w = \min_{\sigma\in\mathscr{S}} \max_{\tau\in\mathscr{T}} \{\varphi^{-1} r^{\sigma\tau} + w_c(e - \varphi^{-1} - \varphi^{-1} P^{\sigma\tau}\varphi) + \varphi^{-1} P^{\sigma\tau}\varphi w = T^\varphi(w)\}.$$

Finally, Lemma 3.24 ensures the uniqueness of the solution w^* of Equation (3.15). □

We verify easily that for $w \in \mathbb{R}^n$, and $i \in S$ we have

$$T_i^{\varphi}(w) = \min_{a \in A_i} \max_{b \in B_{i,a}} \{\varphi_i^{-1} r_i^{ab} + w_c(1 - \varphi_i^{-1}) + \varphi_i^{-1} P_i^{ab} \varphi(w - w_c e)\}, \quad \forall i \in S. \tag{3.16}$$

Lemma 3.26. *The solution w^* of equation* (3.15) *satisfies $\|w^*\|_\infty \leqslant R$.*

Proof. We have $w^* = T^{\varphi}(w^*) = \min_{\sigma \in \mathscr{S}} \max_{\tau \in \mathscr{T}} [\varphi^{-1} r^{\sigma\tau} + P_{\varphi}^{\sigma\tau} w^*]$, then

$$w^* \leqslant \varphi^{-1} R + \max_{\sigma \in \mathscr{S}, \tau \in \mathscr{T}} [P_{\varphi}^{\sigma\tau} w^*].$$

Therefore $\frac{\varphi w^*}{R} \leqslant e + \max_{\sigma \in \mathscr{S}, \tau \in \mathscr{T}} [\varphi P_{\varphi}^{\sigma\tau} \frac{w^*}{R}]$ and by Remark 3.23 we deduce that

$$\frac{\varphi w^*}{R} \leqslant e + \max_{\sigma \in \mathscr{S}, \tau \in \mathscr{T}} [P_{(c,\varphi)}^{\sigma\tau} \frac{\varphi w^*}{R}] = T^{\max,\varphi}(\frac{\varphi w^*}{R}),$$

where the operator $T^{\max,\varphi}$ is defined by $T^{\max,\varphi}(v) = e + \max_{\sigma \in \mathscr{S}, \tau \in \mathscr{T}} [P_{(c,\varphi)}^{\sigma\tau} v]$. By Lemma 3.20, we notice that $T^{\max,\varphi}(\varphi) = \varphi$. Using Remark 3.11, we can easily see that $T^{\max,\varphi}$ is contracting because its max-max operator satisfies $(T^{\max,\varphi})^{\max}(\varphi) = \varphi - e \leqslant (1 - 1/\|\varphi\|_\infty)\varphi$. So by the property appearing in Lemma 3.18 we deduce that $\frac{\varphi w^*}{R} \leqslant \varphi$, and then $w^* \leqslant R$.

We have $-w^* = \max_{\sigma \in \mathscr{S}} \min_{\tau \in \mathscr{T}} [-\varphi^{-1} r^{\sigma\tau} - P_{\varphi}^{\sigma\tau} w^*] \leqslant \max_{\sigma \in \mathscr{S}, \tau \in \mathscr{T}} [-\varphi^{-1} r^{\sigma\tau} - P_{\varphi}^{\sigma\tau} w^*] \leqslant \varphi^{-1} R + \max_{\sigma \in \mathscr{S}, \tau \in \mathscr{T}} [P_{\varphi}^{\sigma\tau}(-w^*)]$, therefore as above $\frac{-\varphi w^*}{R} \leqslant T^{\max,\varphi}(\frac{-\varphi w^*}{R})$ and then $\frac{-\varphi w^*}{R} \leqslant \varphi$, so $-w^* \leqslant R$. We conclude that $\|w^*\|_\infty \leqslant R$. □

Example 3.27. We give an elementary illustration of the present deflation+h-transform technique. Let $P = \left(\begin{smallmatrix} 0 & 1 \\ 1 & 0 \end{smallmatrix}\right)$, $r \in \mathbb{R}^2$, and consider $T(x) = r + Px$. Let us choose $c = 1$. The first hitting time vector φ^* is such that $\varphi_2^* = 1$ and $\varphi_1^* = 1 + \varphi_2^* = 2$, so $\|\varphi^*\|_\infty = 2$. The operator T^{φ^*} given by (3.16) specializes to $T^{\varphi}(w_1, w_2) = (\frac{r_1}{2} + \frac{w_2}{2}, r_2)$. In accordance with Lemma 3.24, this operator is $1/2-$contracting. The unique fixed point of T^{φ} is $w = ((r_1 + r_2)/2, r_2) = \eta e + \varphi^{-1} v$, from which, by Theorem 3.25, we recover the mean payoff $\eta = (r_1 + r_2)/2$, and $v = (0, (r_2 - r_1)/2)$.

3.5 Deflated Value Iteration

In this section, we will use the deflation technique introduced above to design an algorithm that solves the mean payoff problem when Assumption A is satisfied. To solve the non-linear eigenproblem (3.14), we will find a vector φ satisfying (3.12) by solving (3.10) in an approximate way, and then use φ to define the new operator (3.16) and solve the discounted problem (3.15). This algorithm is presented below in Algorithm 2 and we propose to call it *Deflated Value Iteration* (DVI).

To solve the mean payoff problem of Section 3.2.2, we consider the equation:

$$\eta e + v = T(v) \text{ and } v_c = 0, \ \eta \in \mathbb{R}, \ v \in \mathbb{R}^n \tag{3.17}$$

where T is as in Definition 3.1: $T_i(v) = \min_{a \in A_i} \max_{b \in B_{i,a}} \{r_i^{ab} + \sum_{j \in S} P_{ij}^{ab} v_j\}, \forall i \in S$. Throughout the section, we make Assumption A. We denote by (η^*, v^*) the unique solution of this problem. We know by Lemma 3.17 that there exists a vector $\varphi \in \mathbb{R}^n_{\geqslant 0}$ satisfying (3.12). Theorem 3.25 shows that (3.30) is equivalent to the equation:

$$T^{\varphi}(w) = w, \quad w \in \mathbb{R}^n \ , \tag{3.18}$$

with $\eta = w_c$, $v = \varphi(w - w_c)$, and T^φ is given by (3.16). Therefore, the idea of the DVI algorithm is to compute first a vector φ that satisfies the dynamic inequality (3.12). For that we consider the problem of finding the vector of maximal expected first hitting times of state c, denoted φ^* solution of the dynamic equation (3.10), and we suppose that we know a bound H on it:

$$H \geqslant \|\varphi^*\|_\infty = \max_{i \in S} \mathcal{H}_{ic} \ . \tag{3.19}$$

We define the scalar $\lambda \in [0, 1)$ by

$$\lambda := 1 - 1/H \geqslant 1 - 1/\|\varphi^*\|_\infty \ . \tag{3.20}$$

When a vector φ satisfying (3.12) is found, we construct the operator T^φ as in (3.16). Then we solve the discounted fixed point problem (3.15) by value iteration.

Algorithm 2 Deflated Value Iteration

```
Input: vector u ∈ ℝ^n and M ⩾ 0 such that we have ‖u‖_∞ ⩽ M
φ^0 = 0,
while ‖φ^k − φ^{k−1}‖_∞ ⩾ 1/2 do
    φ^{k+1} = T_(c)(φ^k),
done
φ = 2φ^k,
w^0 = 0,
for k = 0, 1, ··· do
    w^{k+1} = T^φ(w^k),
done
η = w_c^k and v = φ(w^k − w_c^k e),
return (η, v).
```

The following theorem shows the time complexity required by the DVI algorithm to solve the mean payoff problem.

Theorem 3.28. *The Deflated Value Iteration algorithm* (2) *finds a solution* (η, v) *of the mean payoff problem* (3.4) *such that* $|\eta - \eta^*| \leqslant \epsilon$*, and* $\|v - v^*\|_\infty \leqslant \frac{4\epsilon}{1-\lambda}$ *in time complexity:*

$$O\Big(\frac{|S||E|}{1-\lambda} \log\Big(\frac{R}{(1-\lambda)\epsilon}\Big)\Big) \ .$$

Proof. We know from Lemma 3.18, that the operator $T_{(c)}$ is contracting under the weighted norm $\|\cdot\|_{\varphi^*}$ with the rate $\lambda_{\varphi^*} = 1 - 1/\|\varphi^*\|_\infty \leqslant \lambda$, then by recurrence $\|\varphi^k - \varphi^{k-1}\|_{\varphi^*} \leqslant \lambda^{k-1}\|\varphi^1 - \varphi^0\|_{\varphi^*} = \lambda^{k-1}\|e\|^*_\varphi \leqslant \lambda^{k-1}$, because $e \leqslant \varphi^*$. Therefore $\|\varphi^k - \varphi^{k-1}\|_\infty \leqslant \lambda^{k-1}\|\varphi^*\|_\infty \leqslant \lambda^{k-1}/(1-\lambda)$. So we deduce that to have $\|\varphi^k - \varphi^{k-1}\|_\infty \leqslant 1/2$, it suffices that $k - 1 \geqslant \frac{\log(2/(1-\lambda))}{1-\lambda}$. Then the first step of the algorithm needs a time complexity of $O(\frac{|S||E|}{1-\lambda} \log\big(\frac{1}{1-\lambda}\big))$. Now we have $\varphi^k \leqslant e/2 + \varphi^{k-1}$, then $e + \max_{\sigma \in \mathscr{S}, \tau \in \mathscr{T}} P^{\sigma\tau}_{(c)}\varphi^k = T_{(c)}(\varphi^k) \leqslant T_{(c)}(e/2 + \varphi^{k-1}) \leqslant e/2 + T_{(c)}(\varphi^{k-1}) = e/2 + \varphi^k$. Therefore taking $\varphi = 2\varphi^k$ ensures that $T_{(c)}(\varphi) \leqslant \varphi$, and then we can use φ for the deflation as presented in Theorem 3.25.

By a similar reasoning using the λ_φ-contraction of T^φ under the sup-norm $\|\cdot\|_\infty$ and that $\|w^*\|_\infty \leqslant R$ from Lemmas 3.24 and 3.26, we prove that $\|w^k - w^*\|_\infty \leqslant \lambda^k_\varphi \|w^0 - w^*\|_\infty \leqslant R\lambda^k_\varphi$. Therefore to ensure that $\|w^k - w^*\|_\infty \leqslant \epsilon$, it suffices that the number of iterations $k \geqslant \frac{\log(R/\epsilon)}{1-\lambda_\varphi}$. We have $0 = \varphi^0 \leqslant \varphi^*$, then by applying k times the order-preserving operator $T_{(c)}$, we get $\varphi^k \leqslant \varphi^*$, then $\varphi \leqslant 2\varphi^*$.

Then $1/(1-\lambda_\varphi) = \|\varphi\|_\infty \leqslant 2\|\varphi^k\|_\infty \leqslant \frac{2}{1-\lambda}$. Therefore to obtain w^k such that $\|w^k - w^*\|_\infty \leqslant \epsilon$, the second step of the algorithm needs a time complexity $O(\frac{|S||E|}{1-\lambda}\log(\frac{R}{\epsilon}))$. Therefore the complexity of the algorithm is $O(\frac{|S||E|}{1-\lambda}\log(\frac{R}{(1-\lambda)\epsilon}))$.

Finally, we get $|\eta^k - \eta^*| = |w_c^k - w_c^*| \leqslant \epsilon$, and $\|v^k - v^*\|_\infty = \|\varphi(w^k - w_c^k e) - \varphi(w^* - w_c^* e)\|_\infty \leqslant 2\|\varphi\|_\infty \|w^k - w^*\|_\infty \leqslant 4\epsilon/(1-\lambda)$. □

3.6 Numerical results

We recall that we want to solve the equation presented in (3.30):

$$\eta e + v = T(v) \text{ and } v_c = 0,\ \eta \in \mathbb{R},\ v \in \mathbb{R}^n$$

In this section, we will compare the DVI algorithm with the classical relative value iteration (RVI) algorithm. Starting from a vector $v^0 = (0, \cdots, 0)^\top$, RVI computes the following sequence, for $k = 0, \cdots$:

$$v^{k+1} = T(v^k) - T(v^k)_c e, \tag{3.21}$$
$$\eta^{k+1} = T(v^{k+1})_c. \tag{3.22}$$

We note that $v_c^k = 0$. The RVI algorithm converges to the optimal value v^* under the assumption that the Dobrushin ergodicity coefficient

$$\alpha = 1 - \min_{i,j\in[n], a\in A_i, a'\in A_j} \sum_{k\in[n]} \min(P_{ik}^a, P_{jk}^{a'}) \tag{3.23}$$

which bounds the contraction rate of the operator T in the span seminorm $\|x\|_H := \max_{i\in[n]} x_i - \min_{j\in[n]} x_j$, is smaller than 1. For more details one can refer to Section 6.6 of [Put14].

We will compare DVI also with another algorithm, Krasnoselkii-Mann relative value iteration [GS20] (RVI+KM). This algorithm combines a step of relative value iteration (3.24) with a step of Krasnoselkii-Mann damping (3.25). Starting from a vector $v^0 = (0, \cdots, 0)^\top$, it computes the following sequence, for $k = 0, \cdots$:

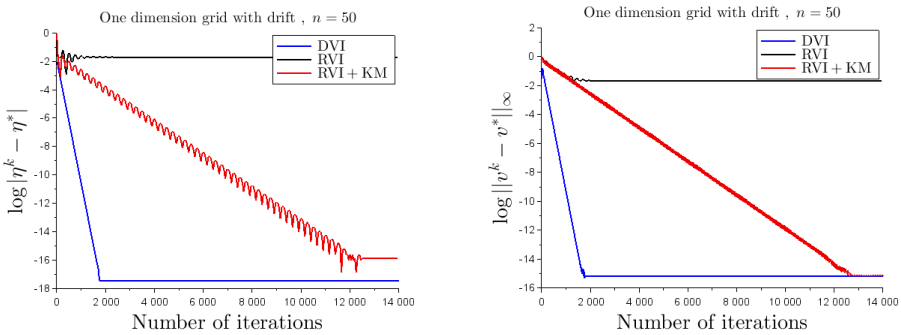
$$\tilde{v}^{k+1} = T(v^k) - T(v^k)_c e, \tag{3.24}$$
$$v^{k+1} = (1-\beta)v^k + \beta\tilde{v}^{k+1}, \tag{3.25}$$
$$\eta^{k+1} = T(v^{k+1})_c. \tag{3.26}$$

where $e = (1, \cdots, 1)^\top \in \mathbb{R}^n$, and $\beta \in (0,1)$ is fixed, $1-\beta$ being interpreted as a *damping parameter*. We note that $v_c^k = 0$. It follows from [GS20, Coro. 13], that v^k does converge to the additive eigenvector v^* of T. Moreover, $\|T(v^k) - v^k\|_H \leqslant 2\|v^*\|_H/\sqrt{\pi\beta(1-\beta)k}$.

In Figure 3.1, we consider a one dimension grid with a transition matrix that represents a cyclic drift. The set of states is $S := \{1, \cdots, n\}$, the reward vector is $r = (0, \cdots, 0, 1)^\top \in \mathbb{R}^n$ and the nonzero entries of the Markov matrix representing this 0-player game are given by:

$$\forall i \in \{1, \cdots, n-1\}, P_{i,i+1} = p,$$
$$\forall i \in \{2, \cdots, n\}, P_{i,i-1} = q,$$
$$P_{n,1} = p,\ P_{1,n} = q\ ,$$

where $p, q > 0$ satisfy $p + q = 1$. We take $p = 0.9$ and $c = n$ as a deflation state for DVI. Figure 3.1 shows that classical relative value iteration does not converge. We note that her $\alpha = 1$. It shows also that deflated value iteration has a better performance than Krasnoselkii-Mann relative value iteration in this case of one dimensional grid with a cyclic drift.



(a) Convergence of the error on the mean payoff η. (b) Convergence of the error on the bias vector v.

Figure (3.1) One dimension grid with drift and $n = 50$ states.

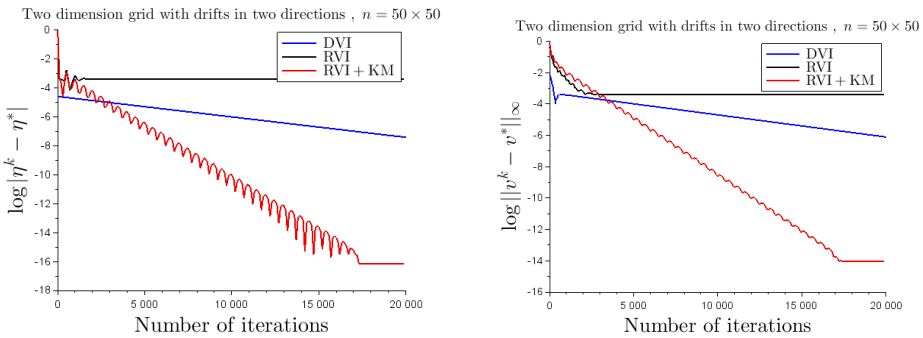
In Figure 3.2, we consider a two dimension grid with a transition matrix that represents a cyclic drift in each dimension. The set of states is $S := \{1, \cdots, l\} \times \{1, \cdots, k\}$, where k and l are integers and the number of states is $n = kl$. The reward vector is given by $r_{i,j} = 0$ for all $i \neq l, j \neq k$ and $r_{l,k} = 1$. The nonzero entries of the Markov matrix representing this 0-player game are given by:

$$\begin{gathered}
\forall i \in \{1, \cdots, l-1\}, \forall j, P_{(i,j),(i+1,j)} = p/2,\ P_{(l,j),(1,j)} = p/2, \\
\forall i \in \{2, \cdots, l\}, \forall j, P_{(i,j),(i-1,j)} = q/2,\ P_{(1,j),(l,j)} = q/2, \\
\forall j \in \{1, \cdots, k-1\}, \forall i, P_{(i,j),(i,j+1)} = p/2,\ P_{(i,k),(i,1)} = p/2, \\
\forall j \in \{2, \cdots, k\}, \forall i, P_{(i,j),(i,j-1)} = q/2,\ P_{(i,1),(i,k)} = q/2,
\end{gathered}$$

where $p, q > 0$ satisfy $p + q = 1$. In this example, we have two cyclic drifts, one horizontal and one vertical. We take $p = 0.9$ and $c = (l, k)$ as a deflation state for DVI. Figure 3.2 shows that relative value iteration does not converge, but Krasnoselkii-Mann relative value iteration converges and has a better performance than deflated value iteration in this case of two dimensional grid with a cyclic drift in each dimension.

In Figure 3.3, we consider a two dimension grid with a transition matrix that represents a cyclic helicoidal drift in addition to a small diffusion. The set of states is $S := \{1, \cdots, l\} \times \{1, \cdots, k\}$, where k and l are integers and the number of states is $n = kl$. The reward vector is given by $r_{i,j} = 0$ for all $i \neq l, j \neq k$ and $r_{l,k} = 1$. The nonzero entries of the Markov matrix representing this 0-player game are given by:

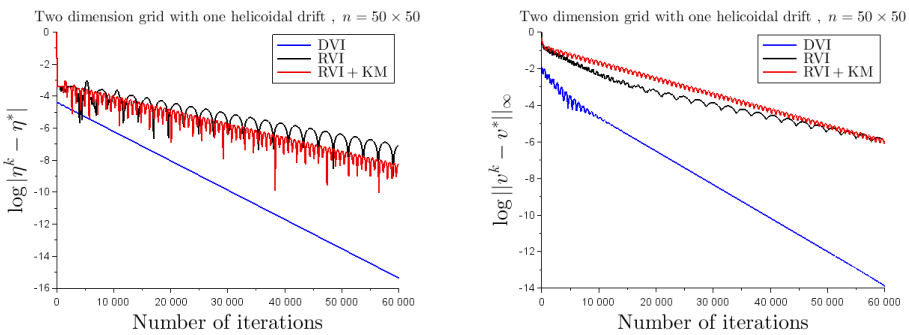
$$\begin{gathered}
\forall j \in \{1, \cdots, k-1\}, \forall i, P_{(i,j),(i,j+1)} = (1+\epsilon)/a, \\
\forall i \in \{1, \cdots, l-1\}, P_{(i,k),(i+1,1)} = (1+\epsilon)/a,\ P_{(l,k),(1,1)} = (1+\epsilon)/a, \\
\forall i \in \{1, \cdots, l-1\}, \forall j, P_{(i,j),(i+1,j)} = (\eta+\epsilon)/a,\ P_{(l,j),(1,j)} = (\eta+\epsilon)/a, \\
\forall i \in \{2, \cdots, l\}, \forall j, P_{(i,j),(i-1,j)} = \epsilon/a,\ P_{(1,j),(l,j)} = \epsilon/a, \\
\forall j \in \{2, \cdots, k\}, \forall i, P_{(i,j),(i,j-1)} = \epsilon/a,\ P_{(i,1),(i,k)} = \epsilon/a,
\end{gathered}$$



(a) Convergence of the error on the mean payoff η. (b) Convergence of the error on the bias vector v.

Figure (3.2) Two dimension grid with drifts in two directions and $n = 50 \times 50$ states.

where $\eta, \epsilon > 0$ are such that $0 < \epsilon < \eta < 1$ and $a = 1 + \eta + 4\epsilon$. In this example, there is a one dimension drift $1/a$ following the cycle $(1,1), \cdots, (1,k), (2,1), \cdots, (l,k), (1,1)$, combined with a smaller cyclic drift η/a and a even smaller diffusion to neighbors ϵ/a. So, the transitions follow an "helicoidal" movement with small perturbation. In Figure 3.3, we take $\eta = 0.05$, $\epsilon = 0.0025$ and $c = (l,k)$ as a deflation state for DVI. Figure 3.3 shows that all three algorithms converge and that deflated value iteration has the best performance in this case of two dimensional grid with a helicoidal drift.



(a) Convergence of the error on the mean payoff η. (b) Convergence of the error on the bias vector v.

Figure (3.3) Two dimension grid with a helicoidal drift and $n = 50 \times 50$ states.

3.7 Extending Sidford et al.'s Variance reduced value iteration method to structured stochastic games

In this section, we will use the deflation technique to split the non-linear eigenproblem (3.14) into two discounted fixed-point problems as in Section 3.5. However to solve these two problems we will use the variance reduction method introduced by Sidford et al. [SWWY18]. We next present a variant of this method that deals with a structured input with two players and a generalized discount, which will allow us to handle both problems (3.10) and (3.15).

We consider a perfect information two-player zero-sum stochastic game with general discount (SG)

as described in Section 3.2, except that we suppose that $P^{\sigma\tau}$ is a sub-Markovian matrix for each couple of policies $(\sigma, \tau) \in \mathscr{S} \times \mathscr{T}$. We suppose the associated Shapley operator can be written as

$$T_i(w) = \min_{a \in A_i} \max_{b \in B_{i,a}} \{\gamma_i^{ab} P_i^{ab} Lw + G_i^{ab}(w)\}, \forall i \in S. \tag{3.27}$$

Here $L \in \mathbb{R}^{n \times n}$ is a sparse operator such that for all $w \in \mathbb{R}^n$, Lw can be computed in $O(|S|)$. For all $i \in S, a \in A_i, b \in B_{i,a}$, G_i^{ab} is a sparse affine operator such that $G_i^{ab}(w)$ can be computed in $O(1)$ for all $w \in \mathbb{R}^n$. For example, by taking $L = \mathrm{Id}$ and $G_i^{ab}(w) = r_i^{ab}, \forall w \in \mathbb{R}^n$ we obtain the Shapley operator of the stochastic game with general discount. The operator L will allow us to handle the deflation (pre-subtraction of $w_c e$) in Equation (3.16).

The problem that we want to solve is:

$$T(w) = w \ . \tag{3.28}$$

In this section, we make the following assumption:

Assumption B. We suppose that

1. T is $\lambda-$contracting under the weighted norm $\|\cdot\|_\psi$, where $\psi \in \mathbb{R}^n$ is a positive vector.
2. The solution w^* of the equation (3.28) verifies $\|w^*\|_\psi \leqslant W$, where $W \geqslant 0$ is a scalar.

We can easily show the following inequalities:

$$\|\psi^{-1}\|_\infty^{-1} \|w\|_\psi \leqslant \|w\|_\infty \leqslant \|\psi\|_\infty \|w\|_\psi, \quad \forall w \in \mathbb{R}^n \ . \tag{3.29}$$

Remark 3.29. In the following, the values $\|\psi^{-1}\|_\infty$ and $\|\psi\|_\infty$ can be replaced by any positive scalars $d_1, d_2 > 0$ such that $\|\psi^{-1}\|_\infty \leqslant d_1$ and $\|\psi\|_\infty \leqslant d_2$.

Remark 3.30. The values $\|\psi^{-1}\|_\infty$ and $\|\psi\|_\infty$ will be used in the following algorithms and in their time complexities. We note that these values can be replaced by any positive scalars $d_1, d_2 > 0$ such that $\|\psi^{-1}\|_\infty \leqslant d_1$ and $\|\psi\|_\infty \leqslant d_2$.

We denote also by $\|\cdot\|_\infty$ the operator norm associated to the sup-norm, so that we have:

$$\|Mw\|_\infty \leqslant \|M\|_\infty \|w\|_\infty, \forall w \in \mathbb{R}^n, \forall M \in \mathbb{R}^{n \times n}.$$

To a given vector $p = (p_j)_{j \in S}$ with $p_j \geqslant 0, \forall j \in S$ and $\sum_{j \in S} p_j \leqslant 1$, we associate the probability vector $\bar{p} = (\bar{p}_j)_{j \in S \cup \{0\}}$ with $\bar{p}_j = p_j, \forall j \in S$ and $\bar{p}_0 = 1 - \sum_{j \in S} p_j$, where 0 is a cemetery state.

For each $i \in S$, $a \in A_i$, $b \in B_{i,a}$, we suppose that we can sample under the probability $\bar{P}_i^{ab}$ associated to the vector P_i^{ab} in time $O(1)$.

We next adapt the algorithms $1 - 6$ presented by Sidford et al. in [SWWY18] to our case with two players. We follow the presentation of [SWWY18], including the decomposition of the algorithm in elementary subroutines. The necessary changes arise from the use of the weighted sup-norm $\|\cdot\|_\psi$ instead of $\|\cdot\|_\infty$, from the sub-Markovian character of the matrices, and from dealing with two players instead of one player. We give the analysis of these modified subroutines in the appendix Section 3.10.

In the following, Algorithm 3 computes an approximation of $P_i^{ab} u$ by sampling under the probability vector $\bar{P}_i^{ab}$. Algorithm 4 computes an approximation of $T(w)$ for $w \in \mathbb{R}^n$, given an initial vector $w_0 \in \mathbb{R}^n$. This algorithm assumes that an approximation of the terms $x_i^{ab} = P_i^{ab} L w_0$, called offsets in [SWWY18], is already known. Then, Algorithm 5 implements a randomized value iteration, using Algorithm 4 at each iteration. To initialize Algorithm 5, the offsets $x_i^{ab} = P_i^{ab} L w_0$ are computed exactly.

Algorithm 3 Approximate transition with cemetery: $\texttt{ApxTransC}(u, M, i, a, b, \epsilon, \delta)$

Input: vector $u \in \mathbb{R}^n$ and $M \geqslant 0$ such that we have $\|u\|_\infty \leqslant M$
Input: State $i \in S$ and actions $a \in A_i$, $b \in B_{i,a}$
Input: Target accuracy $\epsilon > 0$, failure probability $\delta \in (0, 1)$
$u_0 = 0$
$m = \lceil \frac{2M^2}{\epsilon^2} \ln(\frac{2}{\delta}) \rceil$
for $k \in [m]$ **do** choose $i_k \in S \cup \{0\}$ with probabilities $\mathbb{P}(i_k = j) = \bar{P}^{ab}_{ij}$ for $j \in S \cup \{0\}$.
done
return $Y = \frac{1}{m} \sum_{k \in [m]} u_{i_k}$

Algorithm 4 Structured approximate value operator: $\texttt{SApxVal}(w, w_0, x, \epsilon, \delta)$

Input: Current vector $w \in \mathbb{R}^n$ and initial vector $w_0 \in \mathbb{R}^n$.
Input: Precomputed offsets: $x \in \mathbb{R}^E$ with $|x^{ab}_i - P^{ab}_i L w_0| \leqslant \epsilon$ for all $i \in S, a \in A_i, b \in B_{i,a}$.
Input: Target accuracy $\epsilon > 0$, failure probability $\delta \in (0, 1)$
$M = \|L\|_\infty \|w - w_0\|_\infty$
$u = L(w - w_0)$
for $i \in S$ **do**
  for $a \in A_i$ **do**
    for $b \in B_{i,a}$ **do**
      $\tilde{S}^{ab}_i = x^{ab}_i + \texttt{ApxTransC}(u, M, i, a, b, \epsilon, \frac{\delta}{|E|})$
      $\tilde{Q}^{ab}_i = \gamma^{ab}_i \tilde{S}^{ab}_i + G^{ab}_i(w)$
    done
    $\tilde{w}^a_i = \max_{b \in B_{i,a}} \tilde{Q}^{ab}_i$, $\tau(i, a) \in \arg\max b \in B_{i,a}\ \tilde{Q}^{ab}_i$
  done
  $\tilde{w}_i = \min_{a \in A_i} \tilde{w}^a_i$, $\sigma(i) \in \arg\min a \in A_i\ \tilde{w}^a_i$
done
return $(\tilde{w}, \sigma, \tau)$

Algorithm 6 iterates Algorithm 5, using the technique of variance reduction by dividing the error by 2 at every iteration. Algorithm 7 and Algorithm 8 are similar to Algorithm 5 and Algorithm 6, with the difference that the offsets are sampled, instead of being computed exactly.

These subroutines lead to two algorithms, one obtained by calling Algorithm 6 and it runs in quasi-linear time complexity as shown in Theorem 3.31. The second algorithm is obtained by calling Algorithm 8 and it runs in sub-linear time complexity which is presented by Theorem 3.32.

Theorem 3.31 (adaptation of Lem. 4.8 and Lem. 4.9 in [SWWY18])**.** *Algorithm 6 gives with probability* $1 - \delta$ *that* $\|w_k - w^*\|_\psi \leqslant \epsilon_k$ *for all* $k \in [0, K]$*, in particular* $\|w_K - w^*\|_\psi \leqslant \frac{\epsilon}{\|\psi\|_\infty}$ *and then* $\|w_K - w^*\|_\infty \leqslant \epsilon$*, and runs in time*[1]*:*

$$\tilde{O}\Big((|S||E| + \frac{|E|\Gamma^2}{(1-\lambda)^3} \|\psi\|^2_\infty \|\psi^{-1}\|^2_\infty \|L\|^2_\infty) \log(\frac{W}{\epsilon}) \log(\frac{1}{\delta})\Big) \ .$$

Proof. By recurrence we have $\|w_0 - w^*\|_\psi = \|w^*\|_\psi \leqslant W = \epsilon_0$ for $k = 0$. Now we suppose that $\|w_{k-1} - w^*\|_\psi \leqslant \epsilon_{k-1}$ for some $k \in [K]$. By Lemma 3.40, we want to have

$$J \geqslant \lceil \frac{1}{1-\lambda} \log(\frac{\|w_0 - w^*\|_\psi (1-\lambda)}{2\|\psi^{-1}\|_\infty \Gamma \left[\frac{1-\lambda}{4\|\psi^{-1}\|_\infty \Gamma} \epsilon_k\right]}) \rceil = \lceil \frac{1}{1-\lambda} \log(\frac{2\|w_{k-1} - w^*\|_\psi}{\epsilon_k}) \rceil,$$

[1] As in [SWWY18] we use $\tilde{O}$ to hide polylogarithmic factors in the input parameters, i.e. $\tilde{O}(f(x)) = O(f(x) \log(f(x))^{O(1)})$.

Algorithm 5 Structured randomized value iteration: $\texttt{SRandVI}(w_0, J, \epsilon, \delta)$

1: Input: initial vector $w_0 \in \mathbb{R}^n$, number of iterations $J > 0$
2: Input: Target accuracy $\epsilon > 0$, failure probability $\delta \in (0,1)$
3: Compute $x \in \mathbb{R}^E$ such that $x_i^{ab} = P_i^{ab} L w_0$ for all $i \in S$ and $a \in A_i$, $b \in B_{i,a}$.
4: **for** $j \in [J]$ **do** $(w_j, \sigma_j, \tau_j) = \texttt{SApxVal}(w_{j-1}, w_0, x, \epsilon, \frac{\delta}{J})$
5: **done**
6: **return** (w_J, σ_J, τ_J)

Algorithm 6 Structured high precision randomized value iteration:
$\texttt{SHighPrecisionRandVI}(\epsilon, \delta, \lambda, W, \Gamma, \|\psi^{-1}\|_\infty, \|\psi\|_\infty)$

1: Input: Target accuracy $\epsilon > 0$, failure probability $\delta \in (0,1)$
2: Let $K = \lceil \log_2(\frac{\|\psi\|_\infty W}{\epsilon}) \rceil$ and $J = \lceil \frac{1}{1-\lambda} \log(4) \rceil$
3: $w_0 = 0$ and $\epsilon_0 = W$
4: **for** $k \in [K]$ **do**
5: $\quad \epsilon_k = \frac{\epsilon_{k-1}}{2} = \frac{W}{2^k}$
6: $\quad (w_k, \sigma_k, \tau_k) = \texttt{SRandVI}(w_{k-1}, J, \frac{1-\lambda}{4\|\psi^{-1}\|_\infty \Gamma} \epsilon_k, \delta/K)$
7: **done**
8: **return** (w_K, σ_K, τ_K)

which is true since $\|w_{k-1} - w^*\|_\psi \leqslant \epsilon_{k-1} = 2\epsilon_k$ and $J = \lceil \frac{1}{1-\lambda} \log(4) \rceil$. So this ensures that $\|w_k - w^*\|_\psi \leqslant \frac{4\|\psi^{-1}\|_\infty \Gamma [\frac{1-\lambda}{4\|\psi^{-1}\|_\infty \Gamma} \epsilon_k]}{1-\lambda} = \epsilon_k$. In particular we have $\|w_K - w^*\|_\psi \leqslant \epsilon_K = \frac{W}{2^K} \leqslant \frac{\epsilon}{\|\psi\|_\infty}$ since $K = \lceil \log_2(\frac{\|\psi\|_\infty W}{\epsilon}) \rceil$. Therefore by the inequality (3.29), we deduce that $\|w_K - w^*\|_\infty \leqslant \epsilon$. By Lemma 3.41 the kth call of $\texttt{SRandVI}$ in Algorithm 6 takes the time

$$O\Big(|S||E| + J|E|\Big[\frac{\|\psi\|_\infty^2 \|w_{k-1} - w^*\|_\psi^2}{[\frac{1-\lambda}{4\|\psi^{-1}\|_\infty \Gamma} \epsilon_k]^2} + \frac{\Gamma^2 \|\psi\|_\infty^2 \|\psi^{-1}\|_\infty^2}{(1-\lambda)^2}\Big] \|L\|_\infty^2 \log(\frac{|E|J}{\delta/K})\Big) \ .$$

And knowing that $\|w_{k-1} - w^*\|_\psi \leqslant \epsilon_{k-1} = 2\epsilon_k$ and with $K = \lceil \log_2(\frac{\|\psi\|_\infty W}{\epsilon}) \rceil$ iterations, we deduce the result. □

Theorem 3.32 (adaptation of Lem. 4.10 and Lem. 4.12 in [SWWY18])**.** *Algorithm 8 gives with probability $1-\delta$ that $\|w_k - w^*\|_\psi \leqslant \epsilon_k$ for all $k \in [0, K]$, in particular $\|w_K - w^*\|_\psi \leqslant \frac{\epsilon}{\|\psi\|_\infty}$ and then $\|w_K - w^*\|_\infty \leqslant \epsilon$, and runs in time*

$$\tilde{O}\Big(|E|\Gamma^2 \|\psi\|_\infty^2 \|\psi^{-1}\|_\infty^2 \Big[\frac{\|\psi\|_\infty^2 W^2}{(1-\lambda)^2 \epsilon^2} + \frac{1}{(1-\lambda)^3}\Big] \|L\|_\infty^2 \log(\frac{1}{\delta})\Big) \ .$$

Proof. The proof of the first part is similar to Theorem 3.31.

By Lemma 3.42 the kth iteration in Algorithm 8 takes

$$O\Big(|E|\Big(\frac{\|\psi\|_\infty^2 W^2}{[\frac{1-\lambda}{4\|\psi^{-1}\|_\infty \Gamma} \epsilon_k]^2} + J\Big[\frac{\|\psi\|_\infty^2 \|w_{k-1} - w^*\|_\psi^2}{[\frac{1-\lambda}{4\|\psi^{-1}\|_\infty \Gamma} \epsilon_k]^2} + \frac{\Gamma^2 \|\psi\|_\infty^2 \|\psi^{-1}\|_\infty^2}{(1-\lambda)^2}\Big]\Big) \|L\|_\infty^2 \log(\frac{|E|J}{\delta})\Big)$$

We have $\|w_{k-1} - w^*\|_\psi \leqslant \epsilon_{k-1} = 2\epsilon_k$, then the kth iteration takes

$$\tilde{O}\Big(|E|\Gamma^2 \|\psi\|_\infty^2 \|\psi^{-1}\|_\infty^2 \Big[\frac{W^2}{(1-\lambda)^2 \epsilon_k^2} + \frac{1}{(1-\lambda)^3}\Big] \|L\|_\infty^2 \log(\frac{1}{\delta})\Big) \ .$$

We have $K = \lceil \log_2(\frac{\|\psi\|_\infty W}{\epsilon}) \rceil$ iterations, and $\frac{\epsilon}{2\|\psi\|_\infty} \leqslant \epsilon_K = \frac{W}{2^K} \leqslant \frac{\epsilon}{\|\psi\|_\infty}$, then we deduce the result. □

Algorithm 7 Structured sampled randomized value iteration: $\texttt{SSampledRandVI}(w_0, J, \epsilon, \delta)$

1: Input: initial vector $w_0 \in \mathbb{R}^n$, number of iterations $J > 0$
2: Input: Target accuracy $\epsilon > 0$, failure probability $\delta \in (0,1)$
3: Sample to obtain approximate offsets: $\tilde{x} \in \mathbb{R}^E$ such that with probability $1 - \frac{\delta}{2}$, $|\tilde{x}_i^{ab} - P_i^{ab} L w_0| \leqslant \epsilon$ for all $i \in S$ and $a \in A_i$, $b \in B_{i,a}$:
$\tilde{x}_i^{ab} = \texttt{ApxTransC}(w_0, \|L\|_\infty \|w_0\|_\infty, i, a, b, \epsilon, \frac{\delta}{2|E|})$
4: **for** $j \in [J]$ **do** $(w_j, \sigma_j, \tau_j) = \texttt{SApxVal}(w_{j-1}, w_0, \tilde{x}, \epsilon, \frac{\delta}{2J})$
5: **done**
6: **return** (w_J, σ_J, τ_J)

Algorithm 8 Structured sublinear randomized value iteration: $\texttt{SSublinearRandVI}(\epsilon, \delta, \lambda, W, \Gamma, \|\psi^{-1}\|_\infty, \|\psi\|_\infty)$

1: Input: Target accuracy $\epsilon > 0$, failure probability $\delta \in (0,1)$
2: Let $K = \lceil \log_2(\frac{\|\psi\|_\infty W}{\epsilon}) \rceil$ and $J = \lceil \frac{1}{1-\lambda} \log(4) \rceil$
3: $w_0 = 0$ and $\epsilon_0 = W$
4: **for** $k \in [K]$ **do**
5: $\epsilon_k = \frac{\epsilon_{k-1}}{2} = \frac{W}{2^k}$
6: $(w_k, \sigma_k, \tau_k) = \texttt{SSampledRandVI}(w_{k-1}, J, \frac{(1-\lambda)\epsilon_k}{4\|\psi^{-1}\|_\infty \Gamma}, \frac{\delta}{K})$
7: **done**
8: **return** (w_K, σ_K, τ_K)

3.8 Variance reduced deflated value iteration for mean payoff problems

In this section we consider the same framework introduced in Section 3.5. So our aim is to solve the mean payoff problem of Section 3.2.2, by considering the equation:

$$\eta e + v = T(v) \text{ and } v_c = 0,\ \eta \in \mathbb{R},\ v \in \mathbb{R}^n \ , \tag{3.30}$$

and reducing it, by Theorem 3.25, to the equation:

$$T^\varphi(w) = w, \quad w \in \mathbb{R}^n \ , \tag{3.31}$$

with $\eta = w_c$, $v = \varphi(w - w_c)$, and T^φ is given by (3.16), where $\varphi \in \mathbb{R}^n_{\geqslant 0}$ is a vector satisfying (3.12).

3.8.1 Computing an h-transform of the mean payoff problem

Here we want to find a vector $\varphi \in \mathbb{R}^n_{\geqslant 0}$ satisfying (3.12). First, we consider the problem of finding the vector of maximal expected first hitting times of state c, denoted φ^* as in (3.10),

We remark that the component φ_c^* can be computed in time $O(|E|)$ from the other components since $\varphi_c^* = 1 + \max_{a \in A_i, b \in B_{i,a}} [\sum_{j \in S, j \neq c} P_{ij}^{ab} \varphi_j^*]$. By considering $w^* = (\varphi_i^*)_{i \in S \setminus \{c\}} \in \mathbb{R}^{n-1}$ and the matrices $\tilde{P}^{\sigma\tau} \in \mathbb{R}^{(n-1)\times(n-1)}$ defined from $P^{\sigma\tau}$ by removing the c row and the c column, the problem becomes

$$w^* = T^m(w^*) \ , \tag{3.32}$$

where the operator T^m is such that

$$T_i^m(w) = 1 + \max_{a,b} [\tilde{P}_i^{ab} w], \quad \forall i \in S \setminus \{c\}, \forall w \in \mathbb{R}^{n-1}$$

The operator T^m is a particular case of the operator T in (3.28): for $(\sigma, \tau) \in \mathscr{S} \times \mathscr{T}$, $P^{\sigma\tau} = \tilde{P}^{\sigma\tau}$ are sub-Markovian. For $(i, a, b) \in E$, $\gamma_i^{ab} = 1$ and then $\Gamma = 1$, $G_i^{ab}(w) = 1$ for all $w \in \mathbb{R}^{n-1}$ and L is the identity then $\|L\|_\infty = 1$.

From Theorem 3.12 and Equation (3.20), we know that the operator T^m is $\mu-$contracting in the sup-norm $\|\cdot\|_{w^*}$, with $\mu := 1 - 1/\|w^*\|_\infty \leqslant 1 - 1/\|\varphi^*\|_\infty \leqslant \lambda$, so here we take $\psi = w^*$, and λ as contraction rate. We have $\|w^{*-1}\|_\infty \leqslant 1$ and $\|w^*\|_\infty \leqslant \|\varphi^*\|_\infty \leqslant \frac{1}{1-\lambda}$, then according to Remark 3.29 we can take in Algorithm 6 and Algorithm 8 the scalars 1 and $\frac{1}{1-\lambda}$ instead of $\|w^{*-1}\|_\infty$ and $\|w^*\|_\infty$ respectively. We have $\|w^*\|_{w^*} = 1$, then we take $W = 1$.

We use Algorithm 6 and Algorithm 8 to find an $\epsilon-$approximation of w^* in near linear and sublinear time respectively. Then we can deduce φ' an $\epsilon-$approximation of φ^*. By taking $\epsilon = \frac{1}{4}$ and considering $\varphi = 2\varphi'$, we deduce the following Theorem 3.33 and Theorem 3.34. In both theorems $\lambda_\varphi := 1 - 1/\|\varphi\|_\infty$ satisfies $\frac{1}{1-\lambda_\varphi} \leqslant \frac{2}{1-\lambda} + \frac{1}{2}$.

Theorem 3.33. *By calling Algorithm 6, we can, with probability $1 - \delta$, find $\varphi \in \mathbb{R}^n_{\geqslant 0}$ such that $\varphi \geqslant T_{(c)}(\varphi)$, in time*

$$\tilde{O}\Big(\Big(|S||E| + \frac{|E|}{(1-\lambda)^5}\Big)\log(\frac{1}{\delta})\Big) \ .$$

Proof. By applying Theorem 3.31, and setting $w = w_K$, we get the run time and the inequality $\|w - w^*\|_\infty \leqslant \epsilon$ with probability $1 - \delta$. We consider $\varphi' \in \mathbb{R}^n$ such that $\varphi'_i = w_i, \forall i \in S \setminus \{c\}$ and $\varphi'_c = 1 + \max_{a,b}[\sum_{j \in S, j \neq c} P_{ij}^{ab}\varphi'_j]$, we deduce easily that $\|\varphi' - \varphi^*\|_\infty \leqslant \epsilon$. We have $\|T_{(c)}(\varphi') - \varphi'\|_\infty \leqslant \|T_{(c)}(\varphi') - \varphi^*\|_\infty + \|\varphi^* - \varphi'\|_\infty \leqslant (\lambda + 1)\|\varphi' - \varphi^*\|_\infty \leqslant 2\epsilon$. Since $\epsilon = 1/4$, we get that $e + \max_{\sigma \in \mathscr{S}, \tau \in \mathscr{T}} P_{(c)}^{\sigma\tau}\varphi' = T_{(c)}(\varphi') \leqslant e/2 + \varphi'$, therefore $\varphi = 2\varphi'$ satisfies $T_{(c)}(\varphi) \leqslant \varphi$. We have $\|\varphi'\|_\infty \leqslant \|\varphi^*\|_\infty + \epsilon$, then we get that $\frac{1}{1-\lambda_\varphi} = \|\varphi\|_\infty \leqslant 2\|\varphi^*\|_\infty + \frac{1}{2} \leqslant \frac{2}{1-\lambda} + \frac{1}{2} = O(\frac{1}{1-\lambda})$. □

Theorem 3.34. *By calling Algorithm 8, we can, with probability $1 - \delta$, find $\varphi \in \mathbb{R}^n_{\geqslant 0}$ such that $\varphi \geqslant T_{(c)}(\varphi)$, in time*

$$\tilde{O}\Big(\frac{|E|}{(1-\lambda)^6}\log(\frac{1}{\delta})\Big) \ .$$

Proof. As for Theorem 3.33 but by applying Theorem 3.32. □

3.8.2 Solving the mean payoff problem

We suppose that we have identified a vector φ satisfying (3.12). Now we consider the equation (3.31).

As in Section 3.7, we can write T^φ in the general form given by Equation (3.27), where $\gamma_i^{ab} = \varphi_i^{-1}$ for all $i \in S, a \in A_i, b \in B_{i,a}$ and then $\Gamma = \max_{(i,a,b) \in E} \gamma_i^{ab} \leqslant 1$, $L : w \mapsto \varphi(w - w_c e)$ is a sparse linear operator and we have $\|L\|_\infty \leqslant 2\|\varphi\|_\infty = \frac{2}{1-\lambda_\varphi}$ and $G_i^{ab} : w \mapsto \varphi_i^{-1} r_i^{ab} + w_c(1 - \varphi_i^{-1})$ is also a sparse affine operator for all $i \in S, a \in A_i, b \in B_{i,a}$.

By Lemma 3.24, T^φ is a $\lambda_\varphi-$contraction in the sup-norm $\|\cdot\|_\infty$, where $\lambda_\varphi = 1 - 1/\|\varphi\|_\infty$. We take here $\psi = e$, which means $\|\cdot\|_\psi = \|\cdot\|_\infty$ and $\|\psi^{-1}\|_\infty = \|\psi\|_\infty = 1$.

By Lemma 3.26, we have $\|w^*\|_\infty \leqslant R$, where w^* is the solution of Equation (3.31). So we take $W = R$.

The following theorems give an approximation for w^* and then also for v^* and η^* in nearly linear time with Theorem 3.35 (based on Theorem 3.33 and Theorem 3.31), and in sublinear time with Theorem 3.36 (based on Theorem 3.34 and Theorem 3.32). The time complexities considered include the times needed to find φ.

Theorem 3.35. *With probability $1-\delta$, we find φ satisfying* (3.12) *and the call of Algorithm 6, with the parameters $(\epsilon, \frac{\delta}{2}, \lambda_\varphi, R, 1, 1, 1)$, returns $w \in \mathbb{R}^n$ such that $\|w - w^*\|_\infty \leqslant \epsilon$. Therefore we obtain $\eta = w_c$ and $v = \varphi(w - w_c e)$ such that $|\eta - \eta^*| \leqslant \epsilon$ and $\|v - v^*\|_\infty \leqslant \frac{5\epsilon}{1-\lambda}$. The run time needed is*

$$\tilde{O}\Big(\Big(|S||E| + \frac{|E|}{(1-\lambda)^5}\Big)\log(\frac{R}{\epsilon})\log(\frac{1}{\delta})\Big) \ .$$

Proof. By Theorem 3.33 we can find $\varphi \in \mathbb{R}^n_{\geqslant 0}$ such that with probability $1-\frac{\delta}{2}$, $\varphi \geqslant T_{(c)}(\varphi)$, in time $\tilde{O}((|S||E| + \frac{|E|}{(1-\lambda)^5})\log(\frac{1}{\delta}))$. By applying Theorem 3.31, the call of Algorithm 6, with the parameters $(\epsilon, \frac{\delta}{2}, \lambda_\varphi, R, 1, 1, 1)$, returns w such that with probability $1-\frac{\delta}{2}$, we have $\|w - w^*\|_\infty \leqslant \epsilon$, and it runs in time $\tilde{O}((|S||E| + \frac{|E|}{(1-\lambda)^5})\log(\frac{R}{\epsilon})\log(\frac{1}{\delta}))$ which leads the desired complexity. Finally, we get $\eta = w_c$ and $v = \varphi(w - w_c e)$ that satisfy $|\eta - \eta^*| \leqslant \epsilon$ and $\|v - v^*\|_\infty \leqslant \frac{5\epsilon}{1-\lambda}$ because $\|\varphi\|_\infty \leqslant \frac{2}{1-\lambda} + \frac{1}{2}$. □

Theorem 3.36. *With probability $1-\delta$, we find φ satisfying* (3.12) *and the call of Algorithm 8, with the parameters $(\epsilon, \frac{\delta}{2}, \lambda_\varphi, R, 1, 1, 1)$ returns $w \in \mathbb{R}^n$ such that $\|w - w^*\|_\infty \leqslant \epsilon$. Therefore we obtain $\eta = w_c$ and $v = \varphi(w - w_c e)$ such that $|\eta - \eta^*| \leqslant \epsilon$ and $\|v - v^*\|_\infty \leqslant \frac{5\epsilon}{1-\lambda}$. The run time needed is*

$$\tilde{O}\Big(|E|\Big[\frac{R^2}{(1-\lambda)^4\epsilon^2} + \frac{1}{(1-\lambda)^6}\Big]\log(\frac{1}{\delta})\Big) \ .$$

Proof. As for Theorem 3.35 but by applying Theorem 3.34 and Theorem 3.32. □

3.9 Comparison with alternative approaches

3.9.1 A cyclic example: no contraction or mixing

The convergence proof of the sampled relative value iteration in [GJG15] requires the Dobrushin ergodicity coefficient α, presented in (3.23), to be smaller than 1. Consider the 0-player instance with a cyclic matrix P, presented in Example 3.27. Here, $\alpha = 1$, and actually, relative value iteration does not converge. Moreover, the mixing time used in the bound of [Wan17] is infinite. However, as shown in Example 3.27, the deflation+h-transform methods reduces to a fixed point problem with a contraction rate of $1/2$.

3.9.2 An example with small hitting times

Let us first consider a $0-$player problem with state space $[n]$, $T(v) = r + Qv, \forall v \in \mathbb{R}^n$, where $r \in \mathbb{R}^n$ fixed and the probability transition matrix $Q \in \mathbb{R}^{n\times n}$, such that

$$Q_{i,1} = Q_{i,i+1} = 1/2, i \in [n-1], \quad Q_{n,1} = 1 \ .$$

We can easily prove that the expected first return time to state n is $\mathcal{T}_{cc} = \Omega(2^n)$. If we denote by ν the stationary distribution of Q, it follows that $\nu_n = O(2^{-n})$. In [Wan17], one supposes the stationary distribution of Q satisfies $\frac{1}{\sqrt{\tau}n}e \leqslant \nu \leqslant \frac{\sqrt{\tau}}{n}e$, so, $\tau = \Omega(2^{2n}/n^2)$. The complexity bound of [Wan17] is exponential in this example, since it includes a τ^2 factor.

By using our technique, we will first choose $c = 1$ and we verify easily that the first hitting time vector φ^* satisfies $\|\varphi^*\|_\infty \leqslant 2$ (more precisely $\varphi^*_i = 2 - \frac{1}{2^{n-i}}, \forall i \in [n]$). Then the new operator T^{φ^*} is $1/2-$contracting which leads to fast convergence. In particular, Theorem 3.35 gives a time complexity

$\tilde{O}(n^2 \log(\frac{R}{\epsilon}) \log(\frac{1}{\delta}))$, and Theorem 3.36 gives a time complexity $\tilde{O}(n\frac{R^2}{\epsilon^2} \log(\frac{1}{\delta}))$, where R is an upper bound on the payments, ϵ is the target accuracy and δ is the failure probability.

In this 0-player example, $\alpha = 1/2$, so we could use relative value iteration. This is no longer the case if we consider the following 1-player variant. By identifying $n+1$ and 1, consider the stochastic matrix

$$Q'_{i+1,2} = Q'_{i+1,i+2} = \frac{1}{2}, i \in [n-1], \quad Q'_{1,2} = 1 \ .$$

let $r' \in \mathbb{R}^n$ be a vector of payments, and consider the Bellman operator $T(x) = \max(r + Qx, r' + Q'x)$, so that there are two actions in every state. Now, $\alpha = 1$ and the convergence of the relative value iteration [GJG15] is not guaranteed. However, we observe that for all policies, the probability to reach the set of states $\{1, 2\}$ in one step is at least $1/2$. Moreover, for all actions, the probability of the transition $1 \to 2$ is at least $1/2$. It follows that by choosing $c = 2$, the vector of maximal hitting times φ^* satisfies $\|\varphi^*\|_\infty \leqslant 4$ (more precisely $\varphi^*_1 = 2$ and $\varphi^*_i = 4 - \frac{1}{2^{n-i}}, \forall i \in \{2, \cdots, n\}$). Hence, the deflation+h-transform method still has a sublinear behavior on this example.

3.10 Analysis of Algorithms 3 to 8

Algorithm 3 using sub-Markovian matrices is equivalent to algorithm 1 in [SWWY18], so we have the same next lemma based on Hoeffding's inequality.

Lemma 3.37 (Lem. 4.2 in [SWWY18])**.** *Algorithm 3 runs in time $O(M^2\epsilon^{-2}\log(\frac{1}{\delta}))$, and it outputs Y such that $|Y - P_i^{ab}u| \leqslant \epsilon$ with probability $1-\delta$.*

Lemma 3.38 (adaptation of Lem. 4.3 in [SWWY18])**.** *With probability $1-\delta$, Algorithm 4 returns $\tilde{w}$ such that $\|\tilde{w} - T(w)\|_\infty \leqslant 2\Gamma\epsilon$, and then*

$$\|\tilde{w} - T(w)\|_\psi \leqslant 2\|\psi^{-1}\|_\infty \Gamma\epsilon \ , \tag{3.33}$$

and it runs in time:

$$O\left(|E|\lceil \|w - w_0\|_\infty^2 \|L\|_\infty^2 \epsilon^{-2} \log\left(\frac{|E|}{\delta}\right)\rceil\right) \ .$$

Proof. The proof of Lemma 4.3 in [SWWY18] is still true in our case with two players and state-actions dependent discount factor γ_i^{ab} and leads to $\|\tilde{w} - T(w)\|_\infty \leqslant 2\Gamma\epsilon$ with probability $1-\delta$. By Equation (3.29), we get the inequality $\|\tilde{w} - T(w)\|_\psi \leqslant 2\|\psi^{-1}\|_\infty\Gamma\epsilon$, and we obtain the desired complexity by noticing that in Algorithm 4, we have $\|u\|_\infty \leqslant \|w - w_0\|_\infty^2 \|L\|_\infty$. □

Lemma 3.39. *If $w, w' \in \mathbb{R}^n$ satisfy $\|w' - T(w)\|_\psi \leqslant \alpha$ then $\|w' - w^*\|_\psi \leqslant \alpha + \lambda\|w - w^*\|_\psi$.*

Proof. $\|w' - w^*\|_\psi = \|w' - T(w^*)\|_\psi \leqslant \|w' - T(w)\|_\psi + \|T(w) - T(w^*)\|_\psi \leqslant \alpha + \lambda\|w - w^*\|_\psi$. □

The following lemma gives an estimation of the error of the randomized value iteration proposed in Algorithm 5 and in Algorithm 5.

Lemma 3.40 (adaptation of Lem. 4.5 in [SWWY18])**.** *The sequence $(w_j)_{j\in[J]}$ generated by Algorithm 5 (the same by Algorithm 5) satisfies with probability $1-\delta$, that for all $j \in [J]$:*

$$\|w_j - w^*\|_\psi \leqslant \frac{2\|\psi^{-1}\|_\infty\Gamma\epsilon}{1-\lambda} + \exp(-j(1-\lambda))\|w_0 - w^*\|_\psi \tag{3.34}$$

and if $J \geqslant \lceil \frac{1}{1-\lambda}\log(\frac{\|w_0 - w^*\|_\psi(1-\lambda)}{2\|\psi^{-1}\|_\infty\Gamma\epsilon})\rceil$ then $\|w_J - w^*\|_\psi \leqslant \frac{4\|\psi^{-1}\|_\infty\Gamma\epsilon}{1-\lambda}$.

Proof. The proof follows the lines of the proof of Lem. 4.5 in [SWWY18], but by using Equation (3.33) and the λ-contraction of T under the weighted sup norm $\|\cdot\|_\psi$. □

The following Lemmas 3.41 and 3.42 show the time complexities of Algorithms 5 and 7 respectively.

Lemma 3.41 (adaptation of Lem. 4.6 in [SWWY18])**.** *Algorithm 5 runs in time*

$$O\Big(|S\|E| + J|E|\Big[\frac{\|\psi\|_\infty^2\|w_0 - w^*\|_\psi^2}{\epsilon^2} + \frac{\Gamma^2\|\psi\|_\infty^2\|\psi^{-1}\|_\infty^2}{(1-\lambda)^2}\Big]\|L\|_\infty^2 \log\big(\frac{|E|J}{\delta}\big)\Big) \ .$$

Proof. The proof follows the lines of the proof of Lem. 4.6 in [SWWY18], but by using Equation (3.34). □

Lemma 3.42 (adaptation of Lem. 4.11 in [SWWY18])**.** *Algorithm 7 runs in time*

$$O\Big(|E|\Big(\frac{\|\psi\|_\infty^2 W^2}{\epsilon^2} + J\Big[\frac{\|\psi\|_\infty^2\|w_0 - w^*\|_\psi^2}{\epsilon^2} + \frac{\Gamma^2\|\psi\|_\infty^2\|\psi^{-1}\|_\infty^2}{(1-\lambda)^2}\Big]\Big)\|L\|_\infty^2 \log\big(\frac{|E|J}{\delta}\big)\Big) \ .$$

Proof. The sampling of $\tilde{x}$ takes the time $O(|E|\|w_0\|_\infty^2\|L\|_\infty^2\epsilon^{-2}\log(\frac{E|L}{\delta}))$. We have $\|w_0\|_\psi^2 \leqslant (\|w_0 - w^*\|_\psi + \|w^*\|_\psi)^2 \leqslant 2\|w_0 - w^*\|_\psi^2 + 2\|w^*\|_\psi^2 \leqslant 2\|w_0 - w^*\|_\psi^2 + 2W^2$, then $\|w_0\|_\infty^2 \leqslant 2\|\psi\|_\infty^2\|w_0 - w^*\|_\psi^2 + 2\|\psi\|_\infty^2 W^2$. The time taken by the J calls of `SApxVal` was already computed in Lemma 3.41, then we deduce the result. □

3.11 Certificate of existence of a deflation state using directed hypergraphs

In this section, we will show that certifying if a deflation state c exists and finding one in the case of existence can be done in a sub-quadratic time complexity. The analysis will be based on directed hypergraphs (see for reference [GLPN93, AGH15b]).

A *directed hypergraph* is a pair $(\mathcal{V}, \mathcal{A})$, where $\mathcal{V}$ is a set of vertices, and $\mathcal{A}$ is a set of hyperarcs. A *hyperarc* a is a pair (T, H) of non-empty subsets of $\mathcal{V}$. T is called the *tail* of a and H is called the *head* of a. A *hyperpath* from x to y in the directed hypergraph is a sequence of $k \geqslant 1$ hyperarcs $(T_1, H_1), \cdots, (T_k, H_k) \in \mathcal{A}$ satisfying $T_i \subset \cup_{j=0}^{i-1} H_j$ for all $i = 1, \cdots, k+1$, with the conventions $H_0 = \{x\}$ and $T_{k+1} = \{y\}$. The vertex y is said to be reachable from x if $x = y$ or there exists a hyperpath from x to y. We denote by $\mathrm{Reach}(x)$ the set of all vertices reachable from the vertex x.

To the game described in Section 3.2.1, we associate a directed hypergraph $\mathcal{H} = (\mathcal{V}, \mathcal{A})$ defined as follows. The set of vertices is $\mathcal{V} = S \cup E$, and the set of hyperarcs is $\mathcal{A} = \{(j, (i, a, b)) \mid i, j \in S, a \in A_i, b \in B_{i,a}$ such that $P_{ij}^{ab} > 0\} \cup \{(E_i, i) \mid i \in S\}$, where $E_i = \{(i, a, b) \mid a \in A_i, b \in B_{i,a}\}$ for $i \in S$.

Proposition 3.43. *Let $c \in S$. The state c is accessible from all the other states in S under all policies, i.e. c is a deflation state, if and only if* $\mathrm{Reach}(c) \supset S$ *in the directed hypergraph $\mathcal{H}$.*

Proof. To prove this proposition we need only to prove that for each $i \in S$, $i \in \mathrm{Reach}(c)$ if and only if c is accessible from i under all policies. Let $i \in S$ and σ, τ two policies of players MIN and MAX respectively, we suppose that $i \in \mathrm{Reach}(c)$, let $(T_1, H_1), \cdots, (T_k, H_k)$ be the shortest hyperpath from c to i in the hypergraph $\mathcal{H}$. By definition of a hyperpath, we have $T_1 = \{c\}$ and since this hyperpath is the shortest, we have $i \in H_k$. By definition of the hypergraph $\mathcal{H}$, we have necessarily $H_k = \{i\}$ and $T_k = E_i$. We know that $(i, \sigma(i), \tau(i, \sigma(i))) \in E_i$, then by definition of a hyperpath there is

$1 \leqslant l < k$ such that $(i, \sigma(i), \tau(i, \sigma(i))) \in H_l$. Then, by definition of the hypergraph $\mathcal{H}$, we have $H_l = \{(i, \sigma(i), \tau(i, \sigma(i)))\}$ and $T_l = \{j\}$ for some $j \in S$ that satisfies $P_{ij}^{\sigma(i),\tau(i,\sigma(i))}$. Therefore, j is accessible from i under the policies σ, τ and j is reachable from c in the hypergraph $\mathcal{H}$ with a smaller hyperpath $(T_1, H_1), \cdots, (T_l, H_l)$. Then, by recurrence we deduce that c is accessible from i under all policies of the two players.

Conversely let $i \in S \setminus \{c\}$, we suppose that the state c is accessible from i under all policies and we want to prove that $i \in \mathrm{Reach}(c)$. We define the sets $S_0 = \{c\}$, $S_1 = \{s \notin S_0 \mid \forall a \in A_s, b \in B_{s,a}, P_{sc}^{ab} > 0\}$ which is the set of states accessible from c under all policies in one step, $S_2 = \{i \notin S_0 \cup S_1 \mid \forall a \in A_s, b \in B_{s,a}, \exists j \in S_1, P_{sj}^{ab} > 0\}$ which is the set of states accessible from c under all policies in two steps, $\cdots$, $S_k = \{s \notin \bigcup_{l=0}^{k-1} S_l \mid \forall a \in A_s, b \in B_{s,a}, \exists j \in S_{k-1}, P_{sj}^{ab} > 0\}$ which is the set of states accessible from c under all policies in k steps, where $k \leqslant n$ is the biggest integer such that S_k is not empty. We notice that the set of states accessible from c under all policies is exactly the union $\bigcup_{l=0}^{k} S_l$. Then, knowing that i is accessible from c under all policies, there exists $1 \leqslant l \leqslant k$ such that $i \in S_l$. Then, $\forall e \in E_i, \exists j_e \in \bigcup_{p=0}^{l-1} S_p$, such that (j_e, e) is an hyperarc in $\mathcal{H}$. Therefore, by recurrence over $0 \leqslant p \leqslant l$, we deduce that there is a hyperpath from c to i, i.e. $i \in \mathrm{Reach}(c)$. □

Proposition 3.44. *Finding a deflation state can be done in the sub-quadratic time complexity* $O(|S|^2|E|)$.

Proof. For each state $i \in S$, we can compute the set $\mathrm{Reach}(i)$ of the vertices reachable from i. This is known to be solvable in linear time $O(|\mathcal{A}|)$ (see for instance [GLPN93]). We have $|\mathcal{A}| \leqslant |S|(|E| + 1)$ by definition of the hypergraph $\mathcal{H}$, then computing $\mathrm{Reach}(i)$ can be done in $O(|S||E|)$. Now, we just need to check if $\mathrm{Reach}(i) \supset S$ for each $i \in S$. Therefore, checking the existence of a deflation state and finding all the deflation states can be done in the sub-quadratic time complexity $O(|S|^2|E|)$. □

Part II

Tropical best approximation

CHAPTER 4

Tropical linear regression and mean payoff games: or, how to measure the distance to equilibria

Abstract

In this chapter, we study a tropical linear regression problem consisting in finding the best approximation of a set of points by a tropical hyperplane. We establish a strong duality theorem, showing that the value of this problem coincides with the maximal radius of a Hilbert's ball included in a tropical polyhedron. We also show that this regression problem is polynomial-time equivalent to mean payoff games. We illustrate our results by solving an inverse problem from auction theory. In this setting, a tropical hyperplane represents the set of equilibrium prices. Tropical linear regression allows us to quantify the distance of a market to the set of equilibria, and infer secret preferences of a decision maker. This chapter is based on the preprint [AGQS21].

4.1 Introduction

4.1.1 The tropical linear regression problem

A *tropical hyperplane* in the n-dimensional tropical vector space $(\mathbb{R} \cup \{-\infty\})^n$ is a set of vectors of the form

$$\mathcal{H}_a = \{x \in (\mathbb{R} \cup \{-\infty\})^n, \qquad \max_{1 \leqslant i \leqslant n} a_i + x_i \text{ is achieved at least twice}\} \ . \tag{4.1}$$

Such a hyperplane is parametrized by the vector $a = (a_1, \ldots, a_n) \in (\mathbb{R} \cup \{-\infty\})^n$, which is required to be non-identically $-\infty$.

Tropical hyperplanes are among the most basic objects in tropical geometry. They are images by the valuation of hyperplanes over non-archimedean fields, and so, they are the simplest examples of tropical linear spaces [SS04, FR15] and tropical hypersurfaces [EKL06]. Tropical hyperplanes arise in tropical convexity [CGQ04, DS04], since closed tropical convex sets can be described as intersections of tropical half-spaces. A further motivation arises from the study of pricing problems: tropical hypersurfaces have been used in [BK19] to represent the influence of prices on the decision of agents buying bundles of elementary products. The "unit demand" case (bundles of cardinality one) is modeled by tropical hyperplanes.

In this chapter, we address the following tropical analogue of the linear regression problem. Given a finite set of points $\mathcal{V} \subset (\mathbb{R} \cup \{-\infty\})^n$, we look for the best approximation of these points by a tropical hyperplane. Of course, the notion of "best approximation" depends on the metric. A canonical choice in tropical geometry is the (additive version of) Hilbert's projective metric. Its restriction to $\mathbb{R}^n$ is induced by the so called *Hilbert's seminorm* or *Hopf oscillation*

$$\|x\|_H := \max_{i \in [n]} x_i - \min_{i \in [n]} x_i \ .$$

It is a projective metric, in the sense that the distance between two points is zero if and only if these two points differ by an additive constant. Hence, we formulate the *tropical linear regression problem* as the following optimization problem:

$$\operatorname*{Min}_{a} \max_{v \in \mathcal{V}} \min_{x \in \mathcal{H}_a \cap \mathbb{R}^n} \|v - x\|_H \tag{4.2}$$

where the minimum is taken over the space of parameters of tropical hyperplanes. For simplicity, we assume for the moment that the vectors $v \in \mathcal{V}$ have finite entries, this assumption will be relaxed in the body of the article.

Equation (4.2) is a non-convex optimization problem, which is of a disjunctive nature since a tropical hyperplane is a union of convex cones.

The tropical linear regression problem (4.2) is not only of theoretical interest. We shall see that it allows one to quantify the "distance to equilibrium" of a market model, and to infer hidden preferences of a decision maker.

4.1.2 Results

We show that tropical linear regression is tractable, theoretically, and to some extent, computationally. Our main result is a strong duality theorem, Theorem 4.23, showing that the infimum of the distance of the set of points $\mathcal{V}$ to a tropical hyperplane coincides with the supremum of the radii of Hilbert's balls included in the tropical convex cone generated by the elements of $\mathcal{V}$. This provides optimality certificates

which can be interpreted geometrically as collections of n "witness" points among the elements of $\mathcal{V}$. Our approach also entails that tropical linear regression is polynomial-time equivalent to solving mean payoff games. The latter games, originally studied in [EM79, GKK88], are among the problems in the complexity class NP $\cap$ co-NP [ZP96] for which no polynomial time algorithm is known. However, several effective methods are available [GKK88, ZP96, BV07, DG06]. In particular, policy iteration allows to solve large scale instances [Cha09], even if it is generally super-polynomial [Fri09a]. Thus, the present results lead to a practical solution of the tropical linear regression problem.

We subsequently study variants of the tropical linear regression problem, involving in particular the *signed* notion of tropical hyperplane, obtained by requiring the maximum in (4.1) to be achieved by two indices i, j belonging to prescribed disjoint subsets I, J of $[n]$. We also establish a strong duality theorem in this setting, and provide reductions to mean payoff games for these variants.

We finally illustrate tropical linear regression by an application to an auction model. We consider a market governed by an invitation to tender procedure. We suppose that a decision maker selects repeatedly bids made by firms, based on the bid prices, which are ultimately made public (after the decision is taken), and also on other criteria (assessments of the technical quality of each firm or of environmental impact) or influence factors (like bribes). This is a variant of the classical "first-price sealed-bid auction" [Kri02], with a bias induced by the secret preference. Here, we define the market to be at equilibria if for each invitation, there are at least two best offers. Hence, in the simplest model (unit demand), the set of equilibria prices can be represented by a tropical hyperplane. We distinguish two versions of this problem, one in which only the prices are public, and the other, in which the identities of the winners of the successive invitations are also known. In both cases, we show that solving a tropical linear regression problem allows an observer to quantify the distance of such a market to equilibrium, and also to infer secret preference factors. This solves, in the special case of unit-demand, an inverse problem, consisting in identifying the agent preferences and utilities in auction models, like the one of [BK19]. This might be of interest to a regulation authority wishing to quantify anomalies, or to a bidder, who, seeing the history of the market, would wish to determine how much he should have bidded to win a given invitation or to get the best price for an invitation that he won, thus avoiding the "winner's curse".

4.1.3 Related work, and discussion

Several "best approximation" problems have been studied in tropical geometry. The simplest one consists in finding the nearest point in a (closed) tropical module, in the sense of Hilbert's metric. The solution is given by the tropical projection [CGQ04], see also [AGNS11]. The best approximation in the space of ultrametrics, which is a fundamental example of tropical module in view of its application to phylogenetics, has been thoroughly studied [CF00, LSTY17, Ber20]. Another important special case is the best approximation of a point by a tropical linear space [Ard04, JSY07]. In contrast with the regression problem studied here, these problems concern the approximation of a *single* point.

It is a general principle that regression (best approximation) is somehow dual to separation. Hence, tropical linear regression should be compared with the tropical support vector machines (SVM) introduced in [GJ08], and further studied in [TWY20]. Whereas the input of the tropical SVM problem (a configuration of points in dimension $n-1$ partitioned in n color classes) is the same as the one of the version "with types" of the tropical linear regression problem, we explain in Remark 4.45 why both problems differ in essential ways.

A different problem of tropical regression consists in finding a vector x minimizing the sup-norm $\|y - Ax\|_\infty$ where y is a vector of observations, and A is a known matrix acting tropically on x. This can be solved in (strongly) polynomial time, again by a tropical projection [But10]. See also [CF00] for a

general version of this result. Tropical linear regression problems of this nature have been studied in the context of learning [MCT21]. The sparse version is of practical interest; it arose in the approximation of solutions of Hamilton-Jacobi PDE, where it was shown to be equivalent to a non-metric infinite dimensional facility location problem [GMQ11]. The finite dimension version which is NP-hard is studied in [TM19].

A different tropical regression problem, with a L_1-type error term (instead of sup-norm here), has been solved in [YZZ19, Theorem 4], in the special case of a configuration of n points in dimension $n-1$. The value is given by a *tropical volume* [DGJ17], instead of an inner radius.

Tropical geometry has been applied to economics in [BK19], see also [TY19], and [DKM01] for early results in this direction. Our modeling of agent's responses to prices is inspired by [BK19]. Auction models taking into account bribery have been studied in particular in [CLMV05, BP07, Rac13].

We build on the results of [AGG12], showing the equivalence between tropical linear programming and mean payoff games. Further reductions and equivalences, concerning in particular the problem of the emptiness of tropical linear prevarieties, were given in [GP13]. The relation between the mean payoff of a game and the inner radius of a Shapley operator was first observed in [Sko18, AGKS18], where it was applied to define a condition number and derive complexity results for games. The inner radius of tropical polyhedra defined by n generators in dimension $n-1$ was initially characterized in [Ser07], as a tropical eigenvalue.

Open problems related to the present work are discussed in the concluding section.

4.1.4 Organization

In Section 4.2, we recall the needed results concerning tropical algebra, mean payoff games, and non-linear Perron-Frobenius theory. In Section 4.3, we show that computing the inner radius of a tropical polyhedron given by generators is equivalent to solving a mean payoff game. Section 4.4 contains our main results, including Theorem 4.23, the strong duality theorem for tropical linear regression. Several variants of the tropical linear regression problem are dealt with in Section 4.5. In Section 4.6, we explain how to solve tropical linear regression problems in practice, using mean payoff games algorithms. In Section 4.7, we give an application to an auction problem. The appendix provides sufficient conditions for the existence of finite eigenvectors of a class of Shapley operators. These conditions are helpful when dealing with regression problems for configurations of points with $-\infty$ coordinates.

4.2 Preliminaries

4.2.1 Tropical cones

The max-plus semifield $\mathbb{R}_{\max}$ is the set of real numbers, completed by $-\infty$ and equipped with the addition $(a,b) \mapsto \max(a,b)$ and the multiplication $(a,b) \mapsto a \odot b := a+b$. The name "tropical" will be used in the sequel as a synonym of "max-plus". We shall occasionally use variants of this semifield. These include *the min-plus semifield* $\mathbb{R}_{\min}$, which is the set $\mathbb{R} \cup \{+\infty\}$, equipped with the addition $(a,b) \mapsto \min(a,b)$ and the multiplication $(a,b) \mapsto a \odot b := a+b$. This semifield is isomorphic to $\mathbb{R}_{\max}$. These also include the subsemifield $\mathbb{Z}_{\max} \subset \mathbb{R}_{\max}$, with ground set $\mathbb{Z} \cup \{-\infty\}$. We refer the reader to [BCOQ92, But10, MS15a] for background on tropical algebra.

For any integer n, we set $[n] := \{1,\dots,n\}$. For all $x,y \in (\mathbb{R}_{\max})^n$, $A \in (\mathbb{R}_{\max})^{n\times m}$, and $\lambda \in \mathbb{R}_{\max}$, $\lambda + x \in (\mathbb{R}_{\max})^n$ denotes the vector with entries $\lambda + x_i$, for $i \in [n]$, $\lambda + A \in (\mathbb{R}_{\max})^{n\times m}$ denotes the matrix with entries $\lambda + A_{ij}$, for $i \in [n], j \in [m]$, $x \vee y = \sup(x,y)$ denotes the vector with entries $\max(x_i, y_i)$, for $i \in [n]$, and $x \wedge y = \inf(x,y)$ denotes the vector with entries $\min(x_i, y_i)$, for

$i \in [n]$. The set $(\mathbb{R}_{\max})^n$ equipped with the addition $(x, y) \mapsto x \vee y$ and the action $(\lambda, x) \mapsto \lambda + x$ of $\mathbb{R}_{\max}$ is a tropical module, i.e. a module over the semifield $\mathbb{R}_{\max}$.

A subset $\mathcal{C}$ of $(\mathbb{R}_{\max})^n$ is a *tropical (convex) cone* or equivalently a *tropical submodule* of $(\mathbb{R}_{\max})^n$ if it satisfies $x, y \in \mathcal{C}$ and $\lambda \in \mathbb{R}_{\max}$ implies $\lambda + x \in \mathcal{C}$ and $x \vee y \in \mathcal{C}$. We endow $(\mathbb{R}_{\max})^n$ with the topology defined by the metric $\delta(x, y) = \max_{i\in[n]} |e^{x_i} - e^{y_i}|$. It induces the usual topology in $\mathbb{R}^n$. For any given subset $\mathcal{V}$ of $(\mathbb{R}_{\max})^n$, we denote by $\mathrm{Sp}(\mathcal{V})$ the tropical submodule of $(\mathbb{R}_{\max})^n$ generated by $\mathcal{V}$, that is the minimal tropical submodule of $(\mathbb{R}_{\max})^n$ containing $\mathcal{V}$. A tropical polyhedral cone $\mathcal{C}$ is a tropical cone which is finitely generated, that is such that there exists a finite subset $\mathcal{V}$ such that $\mathcal{C} = \mathrm{Sp}(\mathcal{V})$. For any given matrix V, we also denote by $\mathrm{Col}(V)$ the column space of V, that is the tropical polyhedral cone generated by the columns of V, and we denote by $\mathrm{Row}(V)$ the row space of V, that is the tropical polyhedral cone generated by the rows of V.

A tropical polyhedral cone can also be defined externally by a system of tropical linear inequalities of the form

$$\max_{j\in[n]}(A_{ij} + x_j) \leqslant \max_{j\in[n]}(B_{ij} + x_j), \qquad i \in [m] \ , \tag{4.3}$$

where A_{ij}, B_{ij} belong to $\mathbb{R}_{\max}$, see [GK11]. Then, A and B will be thought of as $m \times n$ matrices with entries in $\mathbb{R}_{\max}$.

Let $A \in (\mathbb{R}_{\max})^{m\times n}$ and $x \in (\mathbb{R}_{\max})^n$. We denote by Ax the vector in $(\mathbb{R}_{\max})^m$ with entries $(Ax)_i = \max_{j\in[n]}(A_{ij} + x_j)$, for $i \in [m]$. To a matrix $A \in (\mathbb{R}_{\max})^{m\times n}$, we associate the operator $A^\sharp : (\mathbb{R}_{\min})^m \to (\mathbb{R}_{\min})^n$, given by:

$$\forall y \in (\mathbb{R}_{\min})^m, \forall j \in [n], (A^\sharp y)_j = \min_{i\in[m]}(-A_{ij} + y_i) \ ,$$

with the convention $-\infty + \infty = +\infty$. The operator $A^\sharp$ is called the *adjoint* of A and we can easily check that it satisfies the following property:

$$\forall x \in (\mathbb{R}_{\max})^n, \forall y \in (\mathbb{R}_{\min})^m, Ax \leqslant y \Leftrightarrow x \leqslant A^\sharp y \ .$$

We define the identity matrix $\mathrm{I} \in (\mathbb{R}_{\max})^{n\times n}$ by $\forall i \in [n], \mathrm{I}_{ii} = 0$, and $\forall i, j \in [n], i \neq j, \mathrm{I}_{ij} = -\infty$.

A scalar μ is a *tropical eigenvalue* of a matrix $M \in (\mathbb{R}_{\max})^{n\times n}$ if there exists a vector $u \in (\mathbb{R}_{\max})^n$, not identically $-\infty$, such that $Mu = \mu + u$ in the tropical sense. The eigenvalue is known to be unique when the digraph of M is strongly connected, then it coincides with the maximum weight-to-length ratio of the circuits of the digraph of M. We denote it by $\lambda(M)$. See [BCOQ92, But10] for more information.

4.2.2 Mean payoff games

We consider zero-sum deterministic games, with perfect information, defined as follows. There are two players, "Max" and "Min" (the maximizer and the minimizer), who will move a token on a weighted digraph. We assume this digraph is finite and bipartite: the node set is the disjoint union of two non-empty sets $S^{\max}$ and $S^{\min}$, and the arc set $\mathcal{A}$ is included in $(S^{\max} \times S^{\min}) \cup (S^{\min} \times S^{\max})$. The set of states of the game is the set of nodes of the digraph. We associate a real weight w_{rs} to each arc (r, s).

The two players alternate their actions. When the token is in node $i \in S^{\min}$, Player Min must choose an arc (i, j) in the digraph, meaning he moves the token to node j, and pays w_{ij} to player Max. When Player Min has no possible action, that is, when there are no arcs of the form (i, j) in the digraph, the game terminates, and Player Max receives $+\infty$. Similarly, when the token is in node $j \in S^{\max}$, Player Max must choose an arc (j, i) in the digraph. Then he moves the token from node j to node i, and

receives w_{ji} from Player Min. When Player Max has no possible action, that is there are no arcs of the form (j, i) in the digraph, the game terminates, and Player Max receives $-\infty$.

We measure the time in turns, i.e., a time step consists of two half-turns (a move made by Player Min followed by a move made by Player Max). We consider the following game in horizon k: starting from an initial state $\bar{\imath} \in S^{\min}$. the two players make k moves each, unless the game terminates before. So, if the game does not terminate before time k, the *history* of the game is described by a sequence of nodes $\bar{\imath} = i_0, j_1, i_1, \ldots, j_k, i_k$, belonging alternatively to $S^{\min}$ and $S^{\max}$, and the total payment received by Player Max is given by

$$R_{\bar{\imath}}^k = w_{i_0 j_1} + w_{j_1 i_1} + w_{i_1 j_2} + \cdots + w_{j_k i_k} \ .$$

If the game terminates before time k, we set $R_{\bar{\imath}}^k = \pm\infty$ depending on the player who had no available action. The following assumption requires Player Min to have at least one available action in every state:

Assumption C. For all $i \in S^{\min}$, there exists $j \in S^{\max}$ such that (i, j) is an arc of the digraph of the zero-sum deterministic game.

In this way, we always have $R_{\bar{\imath}}^k \in \mathbb{R} \cup \{-\infty\}$. We shall also consider the dual assumption.

Assumption D. For all $j \in S^{\max}$, there exists $i \in S^{\min}$ such that (j, i) is an arc of the digraph of the zero-sum deterministic game.

In most works on mean payoff games, both assumptions are required to hold, which entails in particular that $R_{\bar{\imath}}^k$ is finite. Here, we shall occasionally relax Assumption D, but always require Assumption C, so that $R_{\bar{\imath}}^k \in \mathbb{R} \cup \{-\infty\}$. This leads to an unpleasant symmetry breaking. However, we shall see that this generality will be sometimes needed to handle the application to tropical linear regression. Indeed, from a tropical perspective, $-\infty$ is the zero element, hence a meaningful value.

A strategy of a player is a map which associates to the history of the game an action of this player. Assuming that Player Min plays according to strategy σ, and that Player Max plays according to strategy τ, we shall write $R_{\bar{\imath}}^k = R_{\bar{\imath}}^k(\sigma, \tau)$ to indicate the dependence on these strategies. It follows from standard dynamic programming arguments that the game in horizon k starting from node $\bar{\imath}$ has a value $v_{\bar{\imath}}^k$ and that Players Min and Max have optimal strategies σ^* and τ^*, respectively, see e.g. [MSZ15a, Th. IV.3.2]. This means that the payment function has the following saddle point property:

$$R_{\bar{\imath}}^k(\sigma, \tau^*) \leqslant v_{\bar{\imath}}^k = R_{\bar{\imath}}^k(\sigma^*, \tau^*) \leqslant R_{\bar{\imath}}^k(\sigma^*, \tau)$$

for all strategies σ, τ. Moreover, the *value vector* $v^k := (v_i^k)_{i \in S^{\min}}$ is determined by the following dynamic programming equation

$$v^k = T(v^{k-1}), \qquad v^0 = 0$$

where $T : (\mathbb{R} \cup \{-\infty\})^n \to (\mathbb{R} \cup \{-\infty\})^n$ is the *Shapley operator*, defined, for $i \in S^{\min}$, by

$$T_i(x) = \min_{j,\ (i,j) \in \mathcal{A}} (w_{ij} + \max_{l,\ (j,l) \in \mathcal{A}} (w_{jl} + x_l)) \ . \tag{4.4}$$

Owing to Assumption C, the above minimum is never taken over an empty set, whereas the above maximum is never taken over an empty set when Assumption D is made. By convention, the maximum of an empty set is $-\infty$. When both assumptions hold, T sends $\mathbb{R}^n$ to $\mathbb{R}^n$.

We are interested in the limit

$$\chi(T) := \lim_{k \to \infty} T^k(0)/k = \lim_{k \to \infty} v^k/k \ .$$

Thus, $\chi_i(T)$ yields the limit of the mean payoff per time unit, for the game starting from the initial state i, when the horizon tends to infinity. It follows from [Koh80] that the limit does exist, and that $\chi(T) \in \mathbb{R}^n$, when Assumption C and Assumption D hold (and more generally, when T is a piecewise affine self-map of $\mathbb{R}^n$ that is non-expansive in some norm, see [AGG12] and Section 4.2.3 below for details). Alternatively, under the same assumptions, $\chi_i(T)$ can be characterized as the value of an infinite mean-payoff game, in which player Max wishes to maximize the liminf of the average payment received per time unit, whereas player Min wishes to minimize the liminf of the same quantity — this is the approach originally described by Ehrenfeucht and Mycielski [EM79]. It follows from a general result of Mertens and Neyman, on the existence of the so called *uniform value* [MN81b], that this approach leads to the same notion of mean payoff. Hence, we shall refer to $\chi_i(T)$ as the (asymptotic) mean payoff starting from node i.

More generally, the limit $\chi(T) \in (\mathbb{R} \cup \{-\infty\})^n$ does exist as soon as Assumption C is satisfied. To see this, observe that we can always construct in polynomial time an equivalent game satisfying also Assumption D. Indeed, let us delete any node i of $S^{\min}$ in which Player Min has at least one action $(i,j) \in \mathcal{A}$ such that Player Max has no available action in state j. After at most $|S^{\min}|$ of such deletions, we arrive at a new game, played on a bipartite subdigraph of the original graph, induced by a subset of nodes belonging to Player Min $S'^{\max} \subset S^{\min}$. Note that $S'^{\max}$ may be empty. It is immediate that this subdigraph satisfies both Assumption C and D. So, for $i \in S^{\min}$, the existence of $\lim_k v_i^k/k$ follows from the result already established, whereas for $i \in S^{\min} \setminus S'^{\max}$, we have $v_i^k = -\infty$ for k large enough, implying $\lim_k v_i^k/k = -\infty$.

A (stationary) *policy* of Player Min is a map $\sigma : S^{\min} \to S^{\max}$ such that $(i, \sigma(i)) \in \mathcal{A}$ for all $i \in S^{\min}$. Such a policy determines a one-player game, in which Player Min always selects moves $i \to \sigma(i)$. This one-player game corresponds to the Shapley operator T^σ, defined by

$$T_i^\sigma(x) = w_{i\sigma(i)} + \max_{l,\, (\sigma(i),l)\in\mathcal{A}} (w_{\sigma(i)l} + x_l) \ .$$

Similarly, a policy of Player Max is a map $\tau : S^{\max} \to S^{\min}$ such that $(j, \tau(j)) \in \mathcal{A}$ for all $j \in S^{\max}$. It determines a one-player game, with Shapley operator ${}^\tau T$ defined by

$${}^\tau T_i(x) = \min_{j,\, (i,j)\in\mathcal{A}} (w_{ij} + w_{j\tau(j)} + x_{\tau(j)}) \ .$$

A result of Liggett and Lippman [LL69b] entails that each player has optimal strategies in a mean payoff game, which are obtained by applying a stationary policy. This entails in particular that

$$\chi(T) = \min_\sigma \chi(T^\sigma) = \max_\tau \chi({}^\tau T) \ ,$$

under Assumptions C and D. The mean payoff $\chi_i(T)$ is known to coincide with the weight-to-length ratio of a circuit of the bipartite digraph of the game, the "length" being measured as the number of full turns, i.e., as the number of Min nodes of the circuit (one half of the ordinary length). In particular, if the payments w_{rs} are integers, the mean payoff is a rational number p/q, where p, q are integers, and q is a positive integer bounded by the maximal length of a circuit of the bipartite digraph of the game (measuring the length as the number of nodes of Min that are visited) and $|p/q|$ is bounded by $2|\max_{rs} w_{rs}|$.

Now we formalize the following problem.

Problem 1 (Mean payoff games)**.** Input: A finite bipartite directed graph with integer weights, satisfying Assumptions C and D, together with an initial node $\imath$. Question: Is the mean payoff $\chi_\imath(T)$ starting from node $\imath$ nonnegative?

As discussed in the introduction, Problem 1 is a fundamental problem in algorithmic game theory [GKK88]. It belongs to the class NP $\cap$ coNP [ZP96], no polynomial time algorithm is known.

It will be useful to keep in mind several equivalent versions of this problem.

As a first variant, one may ask whether $\chi_\imath(T)$ is positive, instead of non-negative. This variant is equivalent to the negated version of Problem 1: considering $\tilde{T}(x) = -T(-x)$, i.e., the Shapley operator of the game in which all payments are negated, we have that $\chi(\tilde{T}) = -\chi(T)$, and so, $\chi_\imath(T) > 0$ iff $\chi_\imath(\tilde{T}) \leqslant 0$.

As observed above, the variant of mean payoff games in which Assumption D is relaxed reduces to the variant in which this assumption holds by a preprocessing, so there is no restriction on requiring Assumption D in Problem 1.

Another variant consists in computing $\chi_\imath(T)$, instead of deciding whether $\chi_\imath(T)$ is nonnegative. This problem of computation polytime Turing-reduces to Problem 1 by binary search. Indeed, given a rational number $\alpha = p/q$, we can consider the modified game with integer weights $w^\alpha_{rs} = 2q(w_{rs} - \alpha/2)$, which corresponds to replacing the Shapley operator T by $T^\alpha := 2q(-\alpha + T)$. Thus $\chi_\imath(T^\alpha) \geqslant 0$ iff $\chi_\imath(T) \geqslant \alpha$. Then, since the mean payoff $\chi_\imath(T)$ is a rational number whose absolute value is bounded by $2\max_{rs}|w_{rs}|$ and whose denominator is bounded by $|S^{\min}|$, we can compute $\chi_\imath(T)$ by a dichotomy argument, calling at each step an oracle solving Problem 1 for a modified game with weights w^α_{rs}.

4.2.3 Perron-Frobenius tools

We now recall some tools from Perron-Frobenius theory, in relation with mean payoff games. We refer the reader to [AGG12] for more information.

We denote by $\bot$ the vector of $(\mathbb{R}_{\max})^n$ identically equal to $-\infty$. We consider the *Hilbert's projective metric*, defined for vectors $x, y \in (\mathbb{R}_{\max})^n$ where at least one of them is not equal to $\bot$, by

$$d(x,y) = \inf\{\lambda - \mu \mid \lambda, \mu \in \mathbb{R},\ \mu + y_i \leqslant x_i \leqslant \lambda + y_i\ \forall i \in [n]\} \in \mathbb{R}_{\geqslant 0} \cup \{+\infty\}\ .$$

In addition, we set $d(\bot, \bot) := 0$.

The *support* of a vector $x \in (\mathbb{R}_{\max})^n$ is defined by $\operatorname{supp} x := \{i \in [n] \mid x_i \neq -\infty\}$. Each subset $I \subset [n]$ yields a *part* P_I of $(\mathbb{R}_{\max})^n$, consisting of vectors with support I.

Observe that $d(x,y)$ is finite if and only if x and y belong to the same part P_I. Moreover, if $I \neq \emptyset$,

$$d(x,y) = \max_{i \in I}(x_i - y_i) - \min_{i \in I}(x_i - y_i)\ .$$

We denote by $\mathbb{P}(\mathbb{R}_{\max})^n$ the *tropical projective space*, i.e., the quotient of the set of non-identically $-\infty$ vectors of $(\mathbb{R}_{\max})^n$ by the equivalence relation $\sim$ which identifies tropically proportional vectors. We shall abuse notation and denote by the same symbol a vector and its equivalence class. Similarly, we shall think of a part P_I with $I \neq \emptyset$ as a subset of the tropical projective space.

Observe that $d(x,y)$ vanishes if and only x and y represent the same point of the tropical projective space, so that d yields a well defined metric on each part of the tropical projective space. We denote by $B(a,r)$ the closed ball centered at $a \in \mathbb{R}^n$ with radius r under Hilbert's projective metric.

It will be convenient to consider an abstract version of the concrete Shapley operators used so far. We call (abstract) *Shapley operator* a map $T : (\mathbb{R}_{\max})^n \to (\mathbb{R}_{\max})^n$ that is order preserving, continuous, and such that $T(\alpha + x) = \alpha + T(x)$ for all $\alpha \in \mathbb{R}_{\max}$ and $x \in (\mathbb{R}_{\max})^n$. Observe that the operator T defined by (4.4), with $S^{\min} = [n]$, is a special case of abstract Shapley operator, as soon as Assumption C holds. We shall often consider situations in which an abstract Shapley operator restricts to a map $\mathbb{R}^n \to \mathbb{R}^n$, we will still use the term *Shapley operator* for the restricted map.

We are interested in the non-linear spectral problem for T, consisting in finding a vector $u \in (\mathbb{R}_{\max})^n$, non-identically $-\infty$, and a scalar $\lambda \in \mathbb{R}_{\max}$ such that $T(u) = \lambda + u$. The *spectral radius* of T is defined as

$$\rho(T) = \sup\{\lambda \in \mathbb{R} \cup \{-\infty\} \mid \exists u \in (\mathbb{R} \cup \{-\infty\})^n, u \neq \bot,\ T(u) = \lambda + u\} \ . \tag{4.5}$$

Variants of this spectral radius are given by the Collatz-Wielandt number cw,

$$\mathrm{cw}(T) = \inf\{\lambda \in \mathbb{R} \mid \exists u \in \mathbb{R}^n,\ T(u) \leqslant \lambda + u\} \ . \tag{4.6}$$

and by the dual Collatz-Wielandt number

$$\mathrm{cw}'(T) = \sup\{\lambda \in \mathbb{R} \cup \{-\infty\} \mid \exists u \in (\mathbb{R} \cup \{-\infty\})^n, u \neq \bot,\ T(u) \geqslant \lambda + u\} \ . \tag{4.7}$$

For all $x \in (\mathbb{R}_{\max})^n$, we define $\operatorname{top} x := \max_{i\in[n]} x_i$. We shall also consider

$$\overline{\chi}(T) := \lim_k \operatorname{top}(T^k(0))/k = \inf_{k\geqslant 1} \operatorname{top}(T^k(0))/k \ .$$

The existence of the limit and the fact it coincides with the infimum follow from the subadditivity property $\operatorname{top}(T^{k+l}(0)) \leqslant \operatorname{top}(T^k(0)) + \operatorname{top}(T^l(0))$. Of course, when the limit $\chi(T) = \lim_k T^k(0)/k$ exists, we have $\overline{\chi}(T) = \operatorname{top} \chi(T) = \max_{i\in[n]} \chi_i(T)$. Then, $\overline{\chi}(T)$ may be interpreted as the value of a modified mean payoff game, in which Player Max chooses first the initial state $i \in [n]$, and then, the games starts from this state as described in Section 4.2.2. Thus, in the sequel, we shall refer to $\overline{\chi}(T)$ as the *upper mean payoff* associated to the operator T.

The following result, which follows from [AGG12], provides several spectral characterizations of this upper mean payoff. We say that a map F from $\mathbb{R}^n$ to $(\mathbb{R} \cup \{-\infty\})^n$ is *piecewise affine* if we can cover $\mathbb{R}^n$ by finitely many polyhedra in such a way that each coordinate map F_i is either affine, or identically $-\infty$, on each of these polyhedra.

Theorem 4.1. *Let $T : (\mathbb{R}_{\max})^n \to (\mathbb{R}_{\max})^n$ be a Shapley operator. Then,*

$$\mathrm{cw}'(T) = \rho(T) = \overline{\chi}(T) = \mathrm{cw}(T) \ , \tag{4.8}$$

and the suprema in (4.5) *and* (4.7) *are always achieved.*

Moreover, if the restriction of T to $\mathbb{R}^n$ is piecewise affine, and if $\rho(T) \neq -\infty$, then the infimum in (4.6) *is also achieved.*

Before giving the details of the derivation of Theorem 4.1 from [AGG12], we need to recall a result of Kohlberg. An *invariant half-line* of a Shapley operator $T : \mathbb{R}^n \to \mathbb{R}^n$ is a pair $(u, \eta) \in \mathbb{R}^n \times \mathbb{R}^n$ such that

$$T(u + s\eta) = u + (s + 1)\eta, \qquad \forall s \geqslant 0 \ .$$

Recall that a self-map of $\mathbb{R}^n$ is *non-expansive* for a fixed norm $\|\cdot\|$ if $\|T(x) - T(y)\| \leqslant \|x - y\|$. Observe that a Shapley operator that preserves $\mathbb{R}^n$ is automatically non-expansive in the sup-norm (see e.g. [GG04]).

Theorem 4.2 ([Koh80])**.** *A piecewise affine map $T : \mathbb{R}^n \to \mathbb{R}^n$ that is nonexpansive in some norm admits an invariant half-line.*

If a Shapley operator $T : \mathbb{R}^n \to \mathbb{R}^n$ has an invariant half-line (u, η), it is immediate, using the fact that T is nonexpansive in the sup-norm, that $\chi(T) = \lim_k T^k(0)/k = \lim_k T^k(u)/k = \lim_k (u + k\eta)/k = \eta$. Thus, the invariant half-line determines the mean payoff vector.

Proof of Theorem 4.1. The equalities in (4.8) are established in [AGG12, Lemma 2.8], where they are derived from a theorem of Nussbaum concerning continuous, order preserving and positively homogeneous self-maps of the orthant, see [Nus86a, Theorem 3.1] and also [GG04, Prop. 1]. If $T(u) = \rho(T) + u$ with $u \neq \bot$, we also have $T(v) = \rho(T) + v$, where $v := u - \operatorname{top} u$ is such that $\operatorname{top} v = 0$. Using the compactness of $\{v \in (\mathbb{R}_{\max})^n \mid \operatorname{top} v = 0\}$ and the continuity of T on this set, we deduce that the supremum in (4.5) is always achieved. A similar argument shows that the supremum in (4.7) is also achieved.

Consider now $F(x) = T(x) \vee (\operatorname{cw}(T) + x)$, which sends $\mathbb{R}^n$ to $\mathbb{R}^n$, and which is piecewise affine because the action spaces are finite. It is immediate that $\operatorname{cw}(F) = \operatorname{cw}(T)$. Let us take an invariant half-line (u, η) of F. Then, it follows from $F^k(u) = u + k\eta$, and from the nonexpansiveness of F in the sup-norm that $\bar{\chi}(T) = \lim_{k\to\infty} \operatorname{top} F^k(0)/k = \lim_{k\to\infty} \operatorname{top} F^k(u)/k = \operatorname{top} \eta$. Moreover, $F(u) = u + \eta \leqslant u + \operatorname{cw}(F)$, and so, $T(u) \leqslant u + \operatorname{cw}(T)$. □

Proposition 4.3. *A piecewise affine Shapley operator $T : \mathbb{R}^n \to \mathbb{R}^n$ admits a finite eigenvector if and only if the mean payoff $\chi_i(T)$ is independent of the choice of the initial state $i \in [n]$.*

Proof. By Theorem 4.2, T has an invariant half-line (u, η) and $\chi(T) = \eta$. So, if $\chi_i(T) = \lambda$ for all i, we have $T(u) = \lambda + u$, showing that u is a finite eigenvector of T. Conversely, if $T(u) = \lambda + u$ for some $u \in \mathbb{R}^n$, then, using the nonexpansiveness of T, $\chi(T) = \lim_k T^k(0)/k = \lim_k T^k(u)/k = \lim_k (u + k\lambda)/k = (\lambda, \dots, \lambda)$. □

4.3 Inner radius of a tropical polyhedron defined by generators

For any subset $\mathcal{W}$ of $(\mathbb{R}_{\max})^n$, we define the *inner radius* of $\mathcal{W}$, denoted $\text{in-rad}(\mathcal{W})$, as the supremum of the radii of Hilbert's balls centered at a point in $\mathbb{R}^n$ and included in $\operatorname{Sp}(\mathcal{W})$. More generally, for all non-empty subsets $I \subset [n]$, we define the *relative inner radius* of $\mathcal{W}$, denoted by $\text{in-rad}_I(\mathcal{W})$, as the supremum of the radii of Hilbert's balls centered at a point in the part P_I of $(\mathbb{R}_{\max})^n$ and included in $\operatorname{Sp}(\mathcal{W})$. Thus, in particular, $\text{in-rad}_{[n]}(\mathcal{W}) = \text{in-rad}(\mathcal{W})$. Observe that the relative inner radius depends only on the image of $\mathcal{W} \cap P_I$ in the tropical projective space $\mathbb{P}(\mathbb{R}_{\max})^n$.

In [AGKS18], it is shown that computing the inner radius of a tropical polyhedral cone given by an external description $P = \{x \in (\mathbb{R}_{\max})^n \mid Ax \leqslant Bx\}$ reduces to computing the Collatz-Wielandt number $\operatorname{cw}(T)$ of a Shapley operator.

In this chapter, we consider the somehow dual situation in which the tropical polyhedral cone is given by an internal description,

$$\operatorname{Col}(V) = \{Vx \mid x \in (\mathbb{R}_{\max})^p\} \ ,$$

where V is a $n \times p$ matrix with entries in the tropical semifield $\mathbb{R}_{\max}$, rather by an external description. Recall that the size of an external description of a tropical polyhedral cone can be exponential in the size of an internal description, and vice versa [AGK11]. This leads us to consider the following problem.

Problem 2. Input: a matrix $V \in \mathbb{Z}_{\max}^{n\times p}$. Goal: Compute the inner radius of $\operatorname{Col}(V)$.

We shall make the following assumption.

Assumption E. The matrix V has no identically $-\infty$ row and no identically $-\infty$ column.

This is not restrictive. Indeed, let $I \subset [n]$ (resp. $J \subset [p]$) denote the set of indices of non-identically $-\infty$ rows (resp. columns) of V and V' denote the $I \times J$ submatrix of V. For $K \subset [n]$, if K is not included in I, we have $\text{in-rad}_K(V) = -\infty$, whereas if $K = I$, then $\text{in-rad}_K(V) = \text{in-rad}(V')$. More generally, the relative inner radii of V for $K \subset I$ coincide with the ones of V' (up to permutations of rows of V).

In [AGG12], the tropical linear independence of the columns of the matrix V was studied by means of a specific Shapley operator, which will also play a key role in our approach. We set $E = \{(i,k) \in [n] \times [p] \mid V_{ik} \neq -\infty\}$. Consider the operator $T : (\mathbb{R}_{\max})^n \to (\mathbb{R}_{\max})^n$, defined by

$$T_i(x) = \inf_{k \in [p], (i,k) \in E} \Big[-V_{ik} + \max_{j \in [n], j \neq i} (V_{jk} + x_j) \Big] \ . \tag{4.9}$$

Owing to Assumption E, the latter infimum is never taken over an empty family, so the operator does send $(\mathbb{R}_{\max})^n$ to $(\mathbb{R}_{\max})^n$. We shall sometimes write T_V instead of T to emphasize the dependence on V. Observe that T is exactly the Shapley operator of a mean payoff game defined in Section 4.2.2: the set of nodes belonging to Player Min is $S^{\min} := [n]$, the set of nodes belonging to Player Max is $S^{\max} := E$, with the set of allowed moves

$$\mathcal{A} = \{(i,(i,k)) \mid i \in [n], k \in [p], (i,k) \in E\} \cup \{((i,k),j) \mid (i,k) \in E, (j,k) \in E,\ i \neq j\} \ . \tag{4.10}$$

The payment associated with the arc $(i,(i,k))$ is $w_{i,(i,k)} = -V_{ik}$, whereas the payment associated with $((i,k),j)$ is $w_{(i,k),j} = V_{jk}$.

Remark 4.4. In this game, the mean payoff $\chi_\imath(T)$ starting from any state $\imath$ is always nonpositive. Indeed, Player Min can always play a "tit for tat" policy, moving to state (j,k) from state j, and thus, paying $-V_{jk}$ to Max, if the last move of Max was $(i,k) \to j$, so that Min paid V_{jk}. In this way, Min can cancel the last payment he made, which guarantees a nonpositive mean payoff.

Given a vector $a \in (\mathbb{R}_{\max})^n$, $a \neq \bot$, we define the *tropical hyperplane*:

$$\mathcal{H}_a := \{x \in (\mathbb{R}_{\max})^n \mid \max_{i \in [n]} (a_i + x_i) \text{ achieved at least twice}\} \ .$$

Observe that $\mathcal{H}_a$ depends only on the point in the tropical projective space represented by a. Moreover, $\mathcal{H}_a$ is stable under the additive action of scalars, so that $\mathcal{H}_a$ can be identified with the subset of the tropical projective space consisting of the equivalence classes of non-identically $-\infty$ vectors of $\mathcal{H}_a$.

For a finite vector $a \in \mathbb{R}^n$, the tropical hyperplane $\mathcal{H}_a$ divides $(\mathbb{R}_{\max})^n$ into n sectors $(S_i(a))_{i \in [n]}$, defined by

$$S_i(a) := \{x \in (\mathbb{R}_{\max})^n \mid \forall j \in [n],\ x_i + a_i \geqslant x_j + a_j\} \ . \tag{4.11}$$

The vector $-a$, which is unique up to an additive constant, is called the *apex* of $\mathcal{H}_a$. Indeed, the set $\mathcal{H}_a \cap \mathbb{R}^n$ modulo the scalar additions is the support of a polyhedral complex and $-a \in \mathbb{R}^n$ is the unique vertex (cell of dimension 0) of this complex. Then, we shall say that $\mathcal{H}_a$ has a *finite apex*. See Figure 4.1 for an illustration.

The following result shows that verifying whether there is a tropical hyperplane containing a given collection of vectors reduces to solving a mean payoff game.

Proposition 4.5. *[AGG12, Corollary 4.8] For $a \in (\mathbb{R}_{\max})^n$ such that $a \neq \bot$, suppose that $V \in (\mathbb{R}_{\max})^{n \times p}$ satisfies Assumption E, and let T be defined as above. Then, the following assertions are equivalent:*

1. $a \leqslant T(a)$;

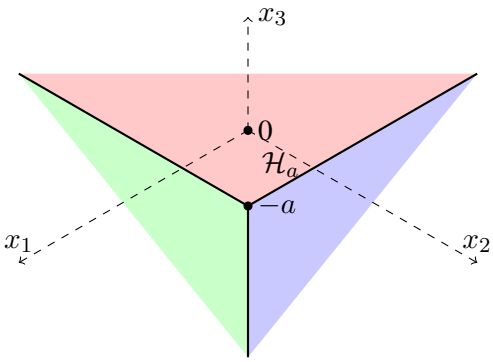


Figure (4.1) The hyperplane $\mathcal{H}_a$ with finite apex $a = (0, 0, 1)^\top$ and the sectors that $\mathcal{H}_a$ defines in the projective space $\mathbb{P}(\mathbb{R}_{\max})^3$.

2. *The column space* $\mathrm{Col}(V)$ *is included in* $\mathcal{H}_a$.

Corollary 4.6. *The columns of* V *are contained in a tropical hyperplane iff* $\rho(T)$ *is nonnegative.*

Proof. This follows from the equality $\rho(T) = \mathrm{cw}'(T)$ in Theorem 4.1 and the fact the supremum is achieved in (4.7). □

Theorem 4.7. *Let* $T = T_V$ *be the Shapley operator associated to the matrix* $V \in (\mathbb{R}_{\max})^{n\times p}$ *defined in* (4.9). *Then,* $\rho(T) \leqslant 0$. *Moreover,*

$$-\rho(T) = \text{in-rad}(\mathrm{Col}(V)).$$

If $\rho(T)$ *is finite, a maximal Hilbert's ball included in* $\mathrm{Col}(V) \cap \mathbb{R}^n$ *is given by* $B(-a, -\rho(T))$ *where* a *is any vector in* $\mathbb{R}^n$ *such that* $T(a) \leqslant \rho(T) + a$.

We will deduce Theorem 4.7 from the following lemma:

Lemma 4.8. *For all* $\lambda \in [-\infty, 0]$ *and* $a \in \mathbb{R}^n$,

$$B(-a, -\lambda) \subset \mathrm{Col}(V) \iff T(a) \leqslant \lambda + a$$

Proof. Suppose first that λ is finite. Then, considering (4.9), we see that $T(a) \leqslant \lambda + a$ is equivalent to

$$\forall i \in [n],\ \exists k \in [p],\ \forall j \in [n],\ j \neq i, \quad -\lambda - a_i + a_j \leqslant V_{ik} - V_{jk}\ . \tag{4.12}$$

Let $x \in \mathbb{R}^n$, we have $x \in B(-a, -\lambda)$ if and only if

$$\forall i \in [n],\ \forall j \in [n], \quad x_i - x_j \leqslant -\lambda - a_i + a_j\ . \tag{4.13}$$

Moreover, the basic properties of residuation entail that $VV^\sharp \leqslant \mathrm{I}$, where $V^\sharp x$ is the maximal element y such that $Vy \leqslant x$. It follows that $x \in \mathrm{Col}(V)$ if and only if $x = VV^\sharp x$, or equivalently, $x \leqslant VV^\sharp x$. The latter property can be rewritten as $x_i \leqslant \max_{k\in[p]}\{V_{ik} + \min_{j\in[n]}(-V_{jk} + x_j)\}$, for all $i \in [n]$, which is equivalent to

$$\forall i \in [n],\ \exists k \in [p],\ \forall j \in [n], \quad x_i - x_j \leqslant V_{ik} - V_{jk}\ . \tag{4.14}$$

We can see that if (4.12) and (4.13) are true then (4.14) follows, which shows the "$\Leftarrow$" direction of the lemma.

Now, we suppose that $B(-a,-\lambda) \subset \mathrm{Col}(V)$. For a given $i \in [n]$, we consider the vector $x^{(i)} \in \mathbb{R}^n$ given by $x_i^{(i)} = -\lambda - a_i$ and $x_j^{(i)} = -a_j$ for all $j \neq i$. Since $\lambda \leqslant 0$, we have $x^{(i)} \in B(-a,-\lambda)$, then $x^{(i)} \in \mathrm{Col}(V)$. Therefore by (4.14), there exists $k \in [p]$ such that $\forall j \in [n],\ x_i^{(i)} - x_j^{(i)} \leqslant V_{ik} - V_{jk}$. Moreover, we have $\forall j \in [n], j \neq i, x_i^{(i)} - x_j^{(i)} = -\lambda - a_i + a_j$. Finally this yields (4.12), which proves that $T(a) \leqslant \lambda + a$.

We finally show that the conclusion of the lemma is still true when $\lambda = -\infty$. This follows from $B(-a,+\infty) = \cup_{\mu\in(-\infty,0)} B(-a,-\mu)$ and $-\infty + a = \inf_{\mu\in(-\infty,0)} \mu + a$. □

Proof of Theorem 4.7. If $B(-a,-\lambda) \subset \mathrm{Col}(V)$ for some finite a, with $\lambda \leqslant 0$, by Lemma 4.8, we see that $T(a) \leqslant \lambda + a$, and we deduce from the Collatz-Wielandt property (Theorem 4.1) that $\rho(T) \leqslant \lambda$, and so, the radius of the ball, $-\lambda$, is bounded above by $-\rho(T)$.

Moreover, it follows from Assumption E that $\mathrm{Col}(V)$ has a finite vector a; indeed, we can take for a the supremum of the columns of V. Then, $B(-a,0) \subset \mathrm{Col}(V)$, and by the previous observation, $0 \leqslant -\rho(T)$.

If $\rho(T) = -\infty$, then using the expression of the Collatz-Wielandt number of T, we get that for all finite $\lambda \leqslant 0$, there exists a finite vector $a \in \mathbb{R}^n$ such that $T(a) \leqslant \lambda + a$. By Lemma 4.8, this implies that $B(-a,-\lambda) \subset \mathrm{Col}(V)$, and so $\text{in-rad}(\mathrm{Col}(V)) \geqslant -\lambda$. Since this holds for all $\lambda \leqslant 0$, we deduce that $\text{in-rad}(\mathrm{Col}(V)) = +\infty = -\rho(T)$ is the supremum of the radius of a Hilbert's ball included in $\mathrm{Col}(V) \cap \mathbb{R}^n$.

Finally, if $\rho(T)$ is finite, since the infimum is attained in the expression of the Collatz-Wielandt number of T (see Theorem 4.1), there exists a finite vector $a \in \mathbb{R}^n$ such that $T(a) \leqslant \rho(T) + a$. By Lemma 4.8, this entails that $B(-a,-\rho(T)) \subset \mathrm{Col}(V)$.

This shows that $-\rho(T)$ is the maximum radius of a Hilbert's ball included in $\mathrm{Col}(V) \cap \mathbb{R}^n$. □

Remark 4.9. One can give an alternative, less direct proof, of Theorem 4.7 by deriving it from Theorem 16 of [AGKS18]. The latter result shows that if T is a Shapley operator which satisfies the technical condition (T must be "diagonal free"), then, the supremum of the radii of Hilbert's balls included in $S(T) := \{x \in \mathbb{R}^n \mid x \leqslant T(x)\}$ coincides with $\sup\{\mu \in \mathbb{R} \mid \exists v \in \mathbb{R}^n, \mu + v \leqslant T(v)\}$. The initial part of the proof of Lemma 4.8, up to (4.14), entails that $\mathrm{Col}(V)$ is precisely the set of vectors x such that $x \leqslant -T(-x)$.

The following is an immediate corollary of Theorem 4.7

Corollary 4.10. *The set* $\mathrm{Col}(V) \cap \mathbb{R}^n$ *is of empty interior if and only* $\rho(T) = 0$. □

By combining Corollary 4.10 and Corollary 4.6, we recover the following known result, established in [DSS05] (when the entries of the matrix V are finite).

Corollary 4.11 (Compare with Th. 4.2 of [DSS05])**.** *The set* $\mathrm{Col}(V) \cap \mathbb{R}^n$ *is of empty interior if and only if* $\mathrm{Col}(V)$ *is included in a tropical hyperplane.* □

The following additional corollary implies that we can check in polynomial time whether the inner radius of $\mathrm{Col}(V)$ is finite.

Corollary 4.12. *The following assertions are equivalent:*

1. *The inner radius of* $\mathrm{Col}(V)$ *is infinite;*
2. *There is no part of* $(\mathbb{R}_{\max})^n$ *that is left invariant by the operator* T*;*
3. $T^n(0)$ *is the vector identically equal to* $-\infty$*;*

4. $\rho(T) = -\infty$.

Proof. (3)⇒(1): Suppose that $T^n(0)$ is equal to $\perp$, the identically $-\infty$ vector. Let us take $u \in (\mathbb{R}_{\max})^n$, not identically $-\infty$, such that $T(u) = \rho(T) + u$. Then, there is a constant $\alpha \in \mathbb{R}$ such that $u_i \leqslant \alpha$, for all $i \in [n]$, and so $n\rho(T) + u = T^n(u) \leqslant T^n(0) + \alpha$ is the identically $-\infty$ vector. It follows that $\rho(T) = -\infty$. Then, by Theorem 4.7, the inner radius of $\mathrm{Col}(V)$ is infinite.

(1)⇒(2): Let I be a non-empty subset of $[n]$, and suppose that the part P_I consisting of vectors of $(\mathbb{R}_{\max})^n$ of support I is left invariant by T. Let u be the vector in this part such that $u_i = 0$ for all $i \in I$. Since $T(u) \in P_I$, there exists a real number α such that $T(u) \geqslant \alpha + u$. Hence, $\rho(T) = \mathrm{cw}'(T) \geqslant \alpha > -\infty$, and, by Theorem 4.7, $\text{in-rad}(\mathrm{Col}(V)) = -\rho(T) < +\infty$.

(2)⇒(3): Consider the map $\pi : (\mathbb{R}_{\max})^n \to \mathcal{P}([n])$, which sends a vector u to its support, $\pi(u) = \{i \in [n] \mid u_i \neq -\infty\}$, and consider the equivalence relation $\ker \pi$ on $(\mathbb{R}_{\max})^n$, such that $(x, y) \in \ker \pi$ iff $\pi(x) = \pi(y)$. The quotient set $(\mathbb{R}_{\max})^n / \ker \pi$ can be identified to $\mathcal{P}([n])$, and the order on $(\mathbb{R}_{\max})^n$ induces an order on $(\mathbb{R}_{\max})^n / \ker \pi$, corresponding to the inclusion order on $\mathcal{P}([n])$. The elements if $(\mathbb{R}_{\max})^n / \ker \pi$ are precisely the parts of $(\mathbb{R}_{\max})^n$, together with the singleton consisting of the identically $-\infty$ vector. Let $\perp = \pi^{-1}(\emptyset)$ denote this singleton, and let $\top = \pi^{-1}([n])$. Observe that $\top$ is the maximal element of $(\mathbb{R}_{\max})^n / \ker \pi$, and that $\perp$ is its minimal element.

Since the operator T is order preserving and commutes with the addition of a constant, it induces a map T_π from $(\mathbb{R}_{\max})^n / \ker \pi$ to itself, which is still order preserving. Moreover, the fixed points of T_π distinct from $\perp$ are precisely the parts of $(\mathbb{R}_{\max})^n$ that are invariant by T. We have, $T_\pi(\top) \leqslant \top$, from which we deduce that $((T_\pi)^k(\top))_{k \geqslant 0}$ is a nonincreasing sequence. If $(T_\pi)^k(\top) = (T_\pi)^{k+1}(\top) \neq \perp$, for some k, then $(T_\pi)^k(\top)$ would be an invariant part of T, contradicting the assumption. It follows that the sequence $(T_\pi^k(\top))_{k \geqslant 0}$ strictly decreases until it reaches $\perp$. Since the maximal cardinality of a chain in the lattice $\mathcal{P}(n)$ is $n + 1$, it follows that $(T_\pi)^n(\top) = \perp$. Hence, $T^n(0)$ is the identically $-\infty$ vector.

Finally, the equivalence between (1) and (4) follows from Theorem 4.7. □

Recall that a vector u in a tropical cone $\mathcal{V} \subset (\mathbb{R}_{\max})^n$ is *extreme* [GK07, BSS07] if $u = v \vee w$ with $v, w \in \mathcal{V}$ implies that $u = v$ or $u = w$. An *extreme direction* of $\mathcal{V}$ is of the form $\mathbb{R}_{\max} + u$, for some extreme vector of $\mathcal{V}$, i.e., it consists of the tropical scalar multiples of u. We say that a tropical cone in $(\mathbb{R}_{\max})^n$ is *simplicial* if it has precisely n extreme directions.

Proposition 4.13. *If a Hilbert's ball of positive radius is included in* $\mathrm{Col}(V)$*, then it is also included in a simplicial tropical cone generated by some* n *columns of* V*.*

Proof. For all maps $\sigma : [n] \to [p]$, such that $(i, \sigma(i)) \in E$, we consider the Shapley operator of the one-player game obtained when player MIN selects the action $k = \sigma(i)$ in state i, that is,

$$T^\sigma : (\mathbb{R}_{\max})^n \to (\mathbb{R}_{\max})^n, \qquad T_i^\sigma(x) = -V_{i\sigma(i)} + \max_{j \in [n], j \neq i} (V_{j\sigma(i)} + x_j) \ .$$

If $B(-a, -\lambda) \subset \mathrm{Col}(V)$, then, by Lemma 4.8, $T(a) \leqslant \lambda + a$. So, by choosing $k = \sigma(i)$ that achieves the minimum in the expression of $T(a)$ in (4.9), we get $T^\sigma(a) \leqslant \lambda + a$. Let $J := \sigma([n])$, so that $|J| \leqslant n$. Since $(i, \sigma(i)) \in E$ holds for all $i \in [n]$, the submatrix $V[J]$ of V, obtained by keeping the columns in J, cannot have a $-\infty$ row. Hence Lemma 4.8 can be applied to $V[J]$. We deduce that $B(-a, -\lambda) \subset \mathrm{Col}(V[J])$. Up to eliminating elements of J, we may assume that the set J is minimal to generate $\mathrm{Col}(V[J])$.

Let u^j denote the jth column of V. Then, every u^j must be extreme in $\mathrm{Col}(V[J])$. Indeed, suppose that $u^j = v \vee w$ with $v, w \in \mathcal{V}$ with $u^j \neq v$ and $u^j \neq w$. Then, we can write $v = \vee_{k \in J}(\lambda_k + u^k)$ and $w = \vee_{k \in J}(\mu_k + u^k)$, for some $\lambda_k, \mu_k \in \mathbb{R}_{\max}$. Moreover, we must have $\lambda_j < 0$, otherwise, $v \geqslant u^j$, and since $v \leqslant v \vee w = u^j$, $v = u^j$, a contradiction. A similar result holds for w. Since

$u^j = v \vee w = \vee_{k\in J}((\lambda_k \vee \mu_k) + u^k)$, and $\lambda_j \vee \mu_j < 0$, we deduce that $u^j = \vee_{k\in J\setminus\{j\}}((\lambda_k \vee \mu_k) + u^k)$ is generated by the columns $\{u^k \mid k \in J \setminus \{j\}\}$, contradicting the minimality of J. It follows that every column of $V[J]$ is extreme in $\mathrm{Col}(V[J])$.

To show that $\mathrm{Col}(V[J])$ is simplicial, it remains to check that $|J| \geqslant n$. It is known that a collection of at most $n-1$ vectors in $(\mathbb{R}_{\max})^n$ is included in a tropical hyperplane – this follows for instance from a tropical analogue of the Radon theorem, see e.g. [But03a] or [AGG09, Coro. 6.13]; or this can be deduced from the characterization of the tropical rank [DSS05, IR09, AGG12]. So if $|J| < n$, then $\mathrm{Col}(V[J])$ is of empty interior, contradicting $B(-a, -\lambda) \subset \mathrm{Col}(V[J])$. □

We get as a corollary the following result.

Corollary 4.14. *We have*

$$\text{in-rad}(\mathrm{Col}(V)) = \max_J \text{in-rad}(\mathrm{Col}(V[J])) \tag{4.15}$$

where the maximum is taken over all subsets $J \subset [p]$ of cardinality n. Moreover, if the inner radius is positive, the maximum is achieved by J such that $\mathrm{Col}(V[J])$ is simplicial.

By convention, if $p < n$, the maximum in (4.15) is zero.

Proof. The inequality $\geqslant$ in (4.15) is trivial. If $\text{in-rad}(\mathrm{Col}(V)) = 0$, the equality trivially holds in (4.15). If $\text{in-rad}(\mathrm{Col}(V)) > 0$, then for all $0 \leqslant \lambda < \text{in-rad}(\mathrm{Col}(V))$, there exists a Hilbert's ball of radius λ included in $\mathrm{Col}(V)$. By Proposition 4.13, this ball is also included in a simplicial tropical cone generated by columns of V, which means that there exists $J \subset [p]$ of cardinality n such that $\lambda \leqslant \text{in-rad}(\mathrm{Col}(V[J])) \leqslant \max_J \text{in-rad}(\mathrm{Col}(V[J]))$. Since this holds for all $0 \leqslant \lambda < \text{in-rad}(\mathrm{Col}(V))$, we deduce the inequality $\text{in-rad}(\mathrm{Col}(V)) \leqslant \max_J \text{in-rad}(\mathrm{Col}(V[J]))$ and so the equality. □

Corollary 4.15. *Computing the inner-radius of a tropical polyhedron (Problem 2) is polynomial-time Turing equivalent to mean payoff games (Problem 1).*

Proof. We observed immediately after stating Problem 1 that the problem of computing $\chi_i(T)$, where T is the Shapley operator of a deterministic mean payoff game, satisfying Assumption C, polynomially Turing-reduces to mean payoff games. By Theorem 4.7, the opposite of the inner-radius is equal to $\rho(T)$. Since, $\rho(T) = \max_{i\in[n]} \chi_i(T)$, computing the inner-radius polynomially Turing-reduces to mean payoff games.

Conversely, Corollary 3.11 of [GP13] shows in particular that mean payoff games (Problem 1) polynomially Turing-reduces to checking whether a collection of vectors $v^1, \ldots, v^p$ of $(\mathbb{Z}_{\max})^n$ are included in a tropical hyperplane. By Corollary 4.11 and Corollary 4.10, the latter problem is equivalent to checking whether the inner-radius of a tropical polyhedral cone $\mathrm{Col}(V)$ vanishes. □

Corollary 4.16. *Computing the center of a Hilbert's ball of maximal radius included in $\mathrm{Col}(V)$, where $V \in \mathbb{Z}_{\max}^{n\times p}$, polynomially Turing-reduces to mean payoff games.*

Proof. We first compute the maximal radius, $-\rho(T)$, which has been noted above, polynomially Turing-reduces to mean payoff games. We can also obtain by the same type of reduction an optimal policy σ of Player Min, which satisfies $\rho(T) = \rho(T^\sigma)$. Indeed, for each move of player Min $i \to j$, we can consider a modified Shapley operator $T^{(i,j)}$ corresponding to the game in which player Min makes the move $i \to j$ when in node i (i.e., this player has no choice in node i), and all the other allowed moves are unchanged. By checking whether $\rho(T^{(i,j)}) = \rho(T)$, we can verify if the move $i \to j$ belongs to an optimal policy of Player Min. By repeatedly restricting the freedom of moves of Player Min, we arrive,

after a polynomial number of evaluation of $\rho(\cdot)$, at such an optimal policy σ. We showed that the center of an optimal Hilbert's ball is of the form $-u$ where $u \in \mathbb{R}^n$ and $T(u) \leqslant \rho(T) + u$. Since $T \leqslant T^\sigma$, and $\rho(T) = \rho(T^\sigma)$, it suffices to construct a vector in $\mathbb{R}^n$ such that $T^\sigma(u) \leqslant \rho(T^\sigma) + u$. Considering the tropically linear map $B := -\rho(T^\sigma) + T^\sigma$, we see this is equivalent to $Bu \leqslant u$. A standard result of tropical spectral theory shows that one can compute such a vector u by solving a shortest path problem. Actually, a tropical generating family of the set of such vectors u is the set of columns of the so called "metric closure" or "Kleene star" B^* of the matrix B, defined as the tropical sum of the tropical powers of B, see e.g. [BCOQ92, Th. 3.101] and [But10, Sect. 4.4]. Moreover, the tropical sum u of the columns of B^* is a finite vector. In this way, we constructed u such that $Bu \leqslant u$, and so $T(u) \leqslant \rho(T) + u$. $\square$

Remark 4.17. A subset $J \subset [p]$ satisfying $\text{in-rad}(\text{Col}(V)) = \max_J \text{in-rad}(\text{Col}(V[J]))$ can be computed by using any mean payoff game algorithm that returns, together with the mean payoff $\overline{\chi}(T)$, a vector $u \in \mathbb{R}^n$ such that $T(u) \leqslant \overline{\chi}(T) + u$. Indeed, we saw in the proof of Proposition 4.13 that, taking any policy σ such that $T(u) = T^\sigma(u)$, and setting $J := \sigma([n])$, we have $B(-u, -\overline{\chi}(T)) \subset \text{Col}(V[J])$.

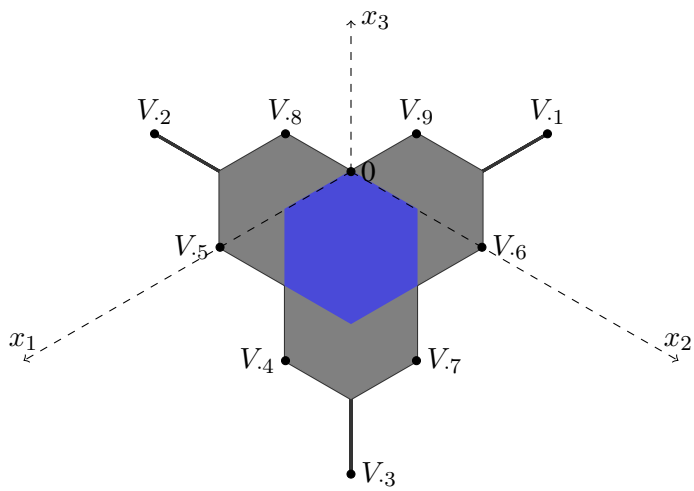


Figure (4.2) Example of an inner ball of the column space $\text{Col}(V)$ in the projective space $\mathbb{P}(\mathbb{R}_{\max})^3$, where $V = \begin{pmatrix} -3 & 0 & 0 & 1 & 1 & -1 & 0 & 0 & -1 \\ 0 & -3 & 0 & 0 & -1 & 1 & 1 & -1 & 0 \\ -1 & -1 & -4 & -2 & -1 & -1 & -2 & 0 & 0 \end{pmatrix}$.

We can verify easily that $\lambda = -1$ and $a = (0, 0, 1)^\top$ satisfy $T(a) = \lambda + a$. Moreover, a policy σ such that $T^\sigma(a) = T(a)$ is given by $\sigma(1) = 4$, $\sigma(2) = 6$ and $\sigma(3) = 8$. Therefore, by Theorem 4.7 the maximal radius of a Hilbert's ball included in $\text{Col}(V)$ is $-\lambda = 1$. Moreover, a maximal Hilbert's ball is given by $B(a, 1)$, and $B(a, 1)$ is included in the simplicial cone $\text{Col}(V[J])$ where $J = \{4, 6, 8\} = \sigma([3])$. This Hilbert's ball, together with the simplicial cone $\text{Col}(V[J])$, are shown in Figure 4.2. Observe that the set J such that $\text{in-rad}(\text{Col}(V[J])) = \text{in-rad}(\text{Col}(V))$ is not unique, indeed, every $J' = \{i, j, k\}$ with $i \in \{4, 5\}$, $j \in \{6, 7\}$ and $k \in \{8, 9\}$ is a candidate.

4.4 The strong duality theorem for tropical linear regression

In this section we will study the best approximation of a set of points in the tropical projective space by a tropical hyperplane. We will show that the best error of approximation is equal to the inner radius of

the tropical module generated by this set of points.

Let $\mathcal{V} = \{v^{(1)}, \cdots, v^{(p)}\} \subset \mathbb{P}(\mathbb{R}_{\max})^n$ be a finite subset of the tropical projective space. Since we mainly focus on $\mathrm{Sp}(\mathcal{V})$, by abusing notions, we denote by $V \in (\mathbb{R}_{\max})^{n\times p}$ the matrix whose columns are given by some representatives of $v^{(1)}, \cdots, v^{(p)}$. Note that $\mathrm{Sp}(\mathcal{V}) = \mathrm{Col}(V)$, which does not depend on the choice of the representatives of $v^{(1)}, \cdots, v^{(p)}$. In the following, we use the notation $r^{\mathrm{in}}_{\mathcal{V}} = \text{in-rad}(\mathrm{Sp}(\mathcal{V}))$.

We introduce a one-sided Hausdorff distance from a set $A \subset \mathbb{P}(\mathbb{R}_{\max})^n$ to a set $B \subset \mathbb{P}(\mathbb{R}_{\max})^n$ with respect to the Hilbert's projective metric, which we shall call the *Hilbert's distance* from A to B:

$$\mathrm{dist}_H(A, B) := \sup_{a\in A} \mathrm{dist}_H(a, B) \ , \quad \text{with } \mathrm{dist}_H(a, B) := \inf_{b\in B} d(a, b) \ . \tag{4.16}$$

Note that $\mathrm{dist}_H(A, B) = 0$ if and only if for each part P_I of the tropical projective space, $A \cap P_I$ is included in the closure of $B \cap P_I$ with respect to the relative topology of P_I.

We are interested in the following *tropical linear regression problem*, consisting of finding a best hyperplane approximation of the set $\mathcal{V}$ in Hilbert's distance:

$$\inf_{a\in\mathbb{P}(\mathbb{R}_{\max})^n} \mathrm{dist}_H(\mathcal{V}, \mathcal{H}_a) \ . \tag{4.17}$$

Observe that if there is an index $i \in [n]$ such that $v_i^{(1)} = \cdots = v_i^{(p)} = -\infty$, then the tropical linear regression problem is trivially solved by setting $a_i = 0$ and $a_j = -\infty$ for $j \neq i$. Hence, in the sequel, we shall assume that the matrix V satisfies Assumption E. In particular, considering the operator T defined in Equation (4.9), we know from Theorem 4.7 that the inner radius of $\mathrm{Sp}(\mathcal{V})$ is $-\rho(T)$.

The following lemma gives a simple formula for the Hilbert's distance from a point to a hyperplane.

Lemma 4.18. *For $x, a \in \mathbb{P}(\mathbb{R}_{\max})^n$, let $i^* \in \arg\max_{i\in[n]}(x_i + a_i)$. Then the Hilbert's distance from the point x to the hyperplane $\mathcal{H}_a$ is*

$$\mathrm{dist}_H(x, \mathcal{H}_a) = x_{i^*} + a_{i^*} - \max_{i\in[n], i\neq i^*}(x_i + a_i) \ , \tag{4.18}$$

where we use the convention $(-\infty) - (-\infty) = 0$.

Proof. If $\max_{i\in[n]}(x_i + a_i) = -\infty$, then $x \in \mathcal{H}_a$ and Equation (4.18) holds with the convention $(-\infty) - (-\infty) = 0$. If $\max_{i\in[n]}(x_i + a_i) \neq -\infty$ and the maximum in the expression is attained twice, then $x \in \mathcal{H}_a$ and Equation (4.18) holds.

Now we focus on the case $\max_{j\in[n], j\neq i^*}(x_j + a_j) < x_{i^*} + a_{i^*} \in \mathbb{R}$, which implies $x_{i^*} \in \mathbb{R}$ and $a_{i^*} \in \mathbb{R}$. We split the argument into the following two cases.

Case 1: $\max_{j\in[n], j\neq i^*}(x_j + a_j) \in \mathbb{R}$. Then $\delta := x_{i^*} + a_{i^*} - \max_{j\in[n], j\neq i^*}(x_i + a_i) > 0$. Consider the point $\tilde{x}$ given by

$$\begin{cases} \tilde{x}_{i^*} = x_{i^*} - \delta & , \\ \tilde{x}_j = x_j & , \text{ for } j \in [n], j \neq i^* \end{cases}$$

Then $\tilde{x} \in \mathcal{H}_a$ and $d(x, \tilde{x}) = \delta$, implying $\mathrm{dist}_H(x, \mathcal{H}_a) \leqslant \delta$. Now, let $x' \in \mathcal{H}_a$, then the maximum in $\max_{j\in[n]}(x'_j + a_j)$ is achieved at least twice. So there exists $i \neq i^*$, such that $\max_{j\in[n]}(x'_j + a_j) = x'_i + a_i$. Since $i \neq i^*$, we have $\delta \leqslant x_{i^*} + a_{i^*} - (x_i + a_i)$, then $x'_i - x_i \geqslant x'_i + a_i - (x_{i^*} + a_{i^*}) + \delta$. Since $x'_{i^*} + a_{i^*} \leqslant \max_{k\in[n]}(x'_k + a_k) = x'_i + a_i$, then $x'_{i^*} - x_{i^*} \leqslant x'_i + a_i - (x_{i^*} + a_{i^*})$. Therefore $d(x, x') \geqslant (x'_i - x_i) - (x'_{i^*} - x_{i^*}) \geqslant \delta$, which proves $\mathrm{dist}_H(x, \mathcal{H}_a) \geqslant \delta$.

Case 2: $\max_{j\in[n], j\neq i^*}(x_j + a_j) = -\infty$. For $x' \in \mathcal{H}_a$, there exists $i \neq i^*$, such that $\max_{j\in[n]}(x'_j + a_j) = x'_i + a_i$. If $x'_i + a_i = -\infty$, then $x'_{i^*} - a_{i^*} = -\infty$. Since $a_{i^*} \in \mathbb{R}$, then $x'_{i^*} = -\infty$. Thus

the fact that $x_{i^*} \in \mathbb{R}$ forces $d(x, x') = +\infty$, i.e., $\mathrm{dist}_H(x, \mathcal{H}_a) = +\infty$ and Equation (4.18) holds. Now if $x'_i + a_i \in \mathbb{R}$, then $x'_i \in \mathbb{R}$ and $a_i \in \mathbb{R}$. Since the assumption $\max_{j\in[n], j\neq i^*}(x_j + a_j) = -\infty$ and $i \neq i^*$ gives us $x_i + a_i = -\infty$, we have $x_i = -\infty$, which leads to $d(x, x') = +\infty$. Therefore $\mathrm{dist}_H(x, \mathcal{H}_a) = +\infty$ and Equation (4.18) holds. □

The next lemma shows that the distance from a Hilbert's ball to any tropical hyperplane is bounded below by the radius of this ball.

Lemma 4.19. *For $a, b \in \mathbb{P}(\mathbb{R}_{\max})^n$, suppose that the supports of a and b are not disjoint. Then, for all $r \geqslant 0$, we have*

$$\mathrm{dist}_H(B(a, r), \mathcal{H}_b) \geqslant r \ . \tag{4.19}$$

Proof. Let $i^* \in \arg\max_{i\in[n]}(a_i + b_i)$. Since the supports of a, b are not disjoint, we have $a_{i^*} + b_{i^*} > -\infty$. Define $x \in (\mathbb{R}_{\max})^n$ by $x_{i^*} = r + a_{i^*}$ and $x_i = a_i$ for all $i \neq i^*$. Then $x \in B(a, r)$, and for all $i \neq i^*$, $x_{i^*} + b_{i^*} = r + a_{i^*} + b_{i^*} \geqslant r + a_i + b_i = r + x_i + b_i$. So by Lemma 4.18 we deduce that $\mathrm{dist}_H(x, \mathcal{H}_b) \geqslant r$, which implies that $\mathrm{dist}_H(B(a, r), \mathcal{H}_b) \geqslant r$. □

Lemma 4.20. *Suppose that $\mathcal{W}$ is a tropical cone in $(\mathbb{R}_{\max})^n$. Then,*

$$\mathrm{dist}_H(\mathrm{Sp}(\mathcal{V}), \mathcal{W}) = \mathrm{dist}_H(\mathcal{V}, \mathcal{W}) \ . \tag{4.20}$$

Proof. Consider an element $x \in \mathrm{Sp}(\mathcal{V})$, so that there exists a finite subset of points $(v^{(j)})_{j\in J}$ of $\mathcal{V}$ and $(\alpha_j)_{j\in J} \in \mathbb{R}^J$, satisfying $x = \vee_{j\in J}(\alpha_j + v^{(j)})$. Take $\lambda > \mathrm{dist}_H(\mathcal{V}, \mathcal{W})$. Then, for any $j \in J$, there exists $w^{(j)} \in \mathcal{W}$ such that $d(v^{(j)}, w^{(j)}) \leqslant \lambda$, and so, there are real numbers γ^j, β^j such that $\gamma^j + w^{(j)} \leqslant v^{(j)} \leqslant \beta^j + w^{(j)}$, and $\beta^j - \gamma^j \leqslant \lambda$. After replacing $w^{(j)}$ by $\gamma^j + w^{(j)} \in \mathcal{W}$, we may assume that $\gamma^j = 0$. Then, $\vee_{j\in J}\alpha_j + w^{(j)} \leqslant x \leqslant \lambda + \vee_{j\in J}\alpha_j + w^{(j)}$, which entails that $\mathrm{dist}_H(x, \mathcal{W}) \leqslant \lambda$. Since this holds for all $\lambda > \mathrm{dist}_H(\mathcal{V}, \mathcal{W})$, we deduce that $\mathrm{dist}_H(x, \mathcal{W}) \leqslant \mathrm{dist}_H(\mathcal{V}, \mathcal{W})$, and so, $\mathrm{dist}_H(\mathrm{Sp}(\mathcal{V}), \mathcal{W}) \leqslant \mathrm{dist}_H(\mathcal{V}, \mathcal{W})$.

The other inequality follows from $\mathcal{V} \subset \mathrm{Sp}(\mathcal{V})$. □

The next lemma shows that the distance from the set $\mathcal{V}$ to any tropical hyperplane is always greater than or equal to the radius of any Hilbert's ball included in the module $\mathrm{Sp}(\mathcal{V})$.

Lemma 4.21 (Weak duality)**.** *We have the following inequality*

$$r^{\mathrm{in}}_{\mathcal{V}} = \sup\{r \geqslant 0 \mid \exists a \in \mathbb{R}^n, B(a, r) \subset \mathrm{Sp}(\mathcal{V})\} \leqslant \inf_{b\in\mathbb{P}(\mathbb{R}_{\max})^n} \mathrm{dist}_H(\mathcal{V}, \mathcal{H}_b) \ . \tag{4.21}$$

Proof. Let $a \in \mathbb{R}^n$ and $r \geqslant 0$ such that $B(a, r) \subset \mathrm{Sp}(\mathcal{V})$, and let $b \in \mathbb{P}(\mathbb{R}_{\max})^n$. Since the supports of a and b are not disjoint, by Lemma 4.19, we have $r \leqslant \mathrm{dist}_H(B(a, r), \mathcal{H}_b)$. Since $B(a, r) \subset \mathrm{Sp}(\mathcal{V})$, then $\mathrm{dist}_H(B(a, r), \mathcal{H}_b) \leqslant \mathrm{dist}_H(\mathrm{Sp}(\mathcal{V}), \mathcal{H}_b)$. Therefore, by using Lemma 4.20, we conclude that $r \leqslant \mathrm{dist}_H(\mathcal{V}, \mathcal{H}_b)$. □

Lemma 4.22. *For all $\lambda \in [-\infty, 0]$ and $b \in \mathbb{P}(\mathbb{R}_{\max})^n$, we have*

$$T(b) \geqslant \lambda + b \Leftrightarrow \mathrm{dist}_H(\mathcal{V}, \mathcal{H}_b) \leqslant -\lambda \ . \tag{4.22}$$

Proof. The equivalence is trivial if $\lambda = -\infty$, so, we suppose that $\lambda \in (-\infty, 0]$. Suppose in addition that $T(b) \geqslant \lambda + b$, i.e., for any $i \in [n]$,

$$\min_{k\in[p],(i,k)\in E}[-V_{ik} + \max_{j\in[n],j\neq i}(V_{jk} + b_j)] \geqslant \lambda + b_i.$$

Then for any $i \in [n]$ and any $k \in [p]$,

$$V_{ik} + b_i \leqslant \max_{j\in[n], j\neq i} (V_{jk} + b_j) - \lambda.$$

For each $k \in [p]$, by taking $i \in \arg\max_{j\in[n]}(V_{jk} + b_j)$ and using Lemma 4.18, we deduce that the distance from the column $V_{\cdot k} = v^{(k)}$ to the hyperplane $\mathcal{H}_b$ is $\leqslant -\lambda$, which implies $\mathrm{dist}_H(\mathcal{V}, \mathcal{H}_b) \leqslant -\lambda$.

Now we suppose that $\mathrm{dist}_H(\mathcal{V}, \mathcal{H}_b) \leqslant -\lambda$. For $k \in [p]$ and $i \in [n]$, if $i \notin \arg\max_{j\in[n]}(V_{jk} + b_j)$, then

$$V_{ik} + b_i \leqslant \max_{j\in[n], j\neq i} (V_{jk} + b_j) \leqslant \max_{j\in[n], j\neq i} (V_{jk} + b_j) - \lambda.$$

Otherwise, if $i \in \arg\max_{j\in[n]}(V_{jk}+b_j)$, then knowing that $\mathrm{dist}_H(v^{(k)}, \mathcal{H}_b) \leqslant -\lambda$ and using Lemma 4.18, we get $V_{ik} + b_i \leqslant \max_{j\in[n], j\neq i}(V_{jk} + b_j) - \lambda$. Therefore, we deduce that for all $i \in [n]$ and $k \in [p]$,

$$V_{ik} + b_i \leqslant \max_{j\in[n], j\neq i} (V_{jk} + b_j) - \lambda.$$

Thus for all $i \in [n]$,

$$\min_{k\in[p], (i,k)\in E} [-V_{ik} + \max_{j\in[n], j\neq i} (V_{jk} + b_j)] \geqslant \lambda + b_i,$$

namely $T(b) \geqslant \lambda + b$. □

The following theorem presents a strong duality result between a best tropical hyperplane approximation of a set $\mathcal{V}$ of points and the largest inner balls that its module $\mathrm{Sp}(\mathcal{V})$ contains.

Theorem 4.23 (Strong duality)**.** *We have*

$$\min_{b\in\mathbb{P}(\mathbb{R}_{\max})^n} \mathrm{dist}_H(\mathcal{V}, \mathcal{H}_b) = r^{\mathrm{in}}_{\mathcal{V}} = \sup\{r \geqslant 0 \mid \exists a \in \mathbb{R}^n, B(a, r) \subset \mathrm{Sp}(\mathcal{V})\}. \tag{4.23}$$

The minimum is achieved by any vector $b \in \mathbb{P}(\mathbb{R}_{\max})^n$ such that $T(b) \geqslant \rho(T) + b$. Moreover, if $r^{\mathrm{in}}_{\mathcal{V}}$ is finite, the supremum is achieved by a ball $B(-c, r^{\mathrm{in}}_{\mathcal{V}})$ where $c \in \mathbb{R}^n$ is any vector such that $T(c) \leqslant \rho(T) + c$.

Proof. Theorem 4.7 entails that $r^{\mathrm{in}}_{\mathcal{V}} = -\rho(T)$ and that the last assertion of the theorem holds. Moreover, the existence of a vector $b \in \mathbb{P}(\mathbb{R}_{\max})^n$ such $T(b) \geqslant \rho(T) + b$ follows from Theorem 4.1. Then, by Lemma 4.22, we have $\mathrm{dist}_H(\mathcal{V}, \mathcal{H}_b) \leqslant r^{\mathrm{in}}_{\mathcal{V}}$, which combined with the weak duality property (4.21) implies that the equality holds in (4.21), and that b such that $T(b) \geqslant \rho(T) + b$ achieves the minimum in (4.23). □

The following lemma allows us to bound from below the value of the tropical linear regression problem by looking at points in the sectors of a hyperplane $\mathcal{H}_a$.

Lemma 4.24. *If $a \in \mathbb{R}^n$ and $r \in [0, +\infty]$ are such that*

$$\forall i \in [n], \exists k \in [p], \quad v^{(k)} \in S_i(a) \text{ and } \mathrm{dist}_H(v^{(k)}, \mathcal{H}_a) \geqslant r \ ,$$

then $B(-a, r) \subset \mathrm{Col}(V)$ and $\min_{b\in\mathbb{P}(\mathbb{R}_{\max})^n} \mathrm{dist}_H(\mathcal{V}, \mathcal{H}_b) \geqslant r$.

Proof. If $r = +\infty$, then for any $i \in [n]$, there is some $\sigma_i \in [p]$ such that $v^{(\sigma_i)} \in S_i(a)$ and $\mathrm{dist}_H(v^{(\sigma_i)}, \mathcal{H}_a) = +\infty$. Since a is finite, for any $i \in [n]$, we have $v_i^{(\sigma_i)} \in \mathbb{R}$ and $v_j^{(\sigma(i))} = -\infty$ for any $j \neq i$. We deduce that $\mathrm{Sp}(v^{(\sigma_1)}, \cdots, v^{(\sigma_n)}) = \mathbb{R}_{\max}$. Then $\mathrm{Col}(V) = \mathbb{R}_{\max}$, and so $B(-a, +\infty) = \mathbb{R}_{\max} \subset \mathrm{Col}(V)$.

Now we consider $r \in [0, +\infty)$. For $i \in [n]$, by the assumption of this lemma, there exists $k \in [p]$ such that $v^{(k)} \in S_i(a)$ and $\mathrm{dist}_H(v^{(k)}, \mathcal{H}_a) \geqslant r$. Hence, by using (4.11) and Lemma 4.18, we deduce that the column $V_{\cdot k} = v^{(k)}$ satisfies $V_{ik} + a_i \geqslant r + \max_{j \neq i}(V_{jk} + a_j)$. Therefore, we have $-V_{ik} + \max_{j \neq i}(V_{jk} + a_j) \leqslant -r + a_i$, which implies for any $i \in [n]$,

$$T_i(a) = \inf_{l \in [p], (i,l) \in E} \Big[-V_{il} + \max_{j \in [n], j \neq i} (V_{jl} + a_j) \Big] \leqslant -r + a_i,$$

i.e., $T(a) \leqslant -r + a$. Therefore, by Lemma 4.8, we deduce that $B(-a, r) \subset \mathrm{Col}(V)$.

Finally by Theorem 4.23, we have

$$\min_{b \in \mathbb{P}(\mathbb{R}_{\max})^n} \mathrm{dist}_H(\mathcal{V}, \mathcal{H}_b) = r_{\mathcal{V}}^{\mathrm{in}} = \sup\{s \geqslant 0 \mid \exists x \in \mathbb{R}^n, B(x, s) \subseteq \mathrm{Sp}(\mathcal{V})\} \geqslant r.$$

□

Given a hyperplane $\mathcal{H}_b$, we call *witness point* of $\mathcal{H}_b$ any point p in $\mathcal{V}$ such that the distance from p to the hyperplane $\mathcal{H}_b$ equals the distance from the set $\mathcal{V}$ to this hyperplane.

Theorem 4.25 (Optimality certificates)**.** *Let $a \in \mathbb{R}^n$, then the following assertions are equivalent:*

1. $T(a) = \rho(T) + a$;
2. *The hyperplane $\mathcal{H}_a$ admits a witness point in each sector, meaning that $\forall i \in [n], \exists k \in [p],\ v^{(k)} \in S_i(a)$ and $\mathrm{dist}_H(v^{(k)}, \mathcal{H}_a) = \mathrm{dist}_H(\mathcal{V}, \mathcal{H}_a)$.*

Moreover, if these assertions hold, then, $\rho(T) = -\mathrm{dist}_H(\mathcal{V}, \mathcal{H}_a)$, $\mathcal{H}_a$ is an optimal solution of the tropical linear regression problem, and $B(-a, \mathrm{dist}_H(\mathcal{V}, \mathcal{H}_a))$ is a Hilbert's ball of maximal radius included in $\mathrm{Sp}(\mathcal{V})$.

Proof. If $a \in \mathbb{R}^n$ satisfies $T(a) = \rho(T) + a = -r_{\mathcal{V}}^{\mathrm{in}} + a$, then by Theorem 4.23, $\mathcal{H}_a$ is optimal in Equation (4.23), i.e., $\mathrm{dist}_H(\mathcal{V}, \mathcal{H}_a) = r_{\mathcal{V}}^{\mathrm{in}}$, and for all $i \in [n]$ we have

$$\min_{k \in [p], (i,k) \in E} [-V_{ik} + \max_{j \in [n], j \neq i} (V_{jk} + a_j)] = -r_{\mathcal{V}}^{\mathrm{in}} + a_i.$$

Then for all $i \in [n]$, there exists $k \in [p]$ such that

$$-V_{ik} + \max_{j \in [n], j \neq i} (V_{jk} + a_j) = -r_{\mathcal{V}}^{\mathrm{in}} + a_i,$$

i.e., $V_{ik} + a_i = r_{\mathcal{V}}^{\mathrm{in}} + \max_{j \in [n], j \neq i}(V_{jk} + a_j)$. This implies that $v^{(k)} = V_{\cdot k} \in S_i(a)$, and also by Lemma 4.18, that $\mathrm{dist}_H(v^{(k)}, \mathcal{H}_a) = r_{\mathcal{V}}^{\mathrm{in}} = \mathrm{dist}_H(\mathcal{V}, \mathcal{H}_a)$.

Now, we suppose that we have assertion (2). By Lemma 4.24, we have

$$\min_{b \in \mathbb{P}(\mathbb{R}_{\max})^n} \mathrm{dist}_H(\mathcal{V}, \mathcal{H}_b) \geqslant \mathrm{dist}_H(\mathcal{V}, \mathcal{H}_a),$$

which means that $\mathcal{H}_a$ achieves the minimum in (4.23), so that $\mathrm{dist}_H(\mathcal{V}, \mathcal{H}_a) = r_{\mathcal{V}}^{\mathrm{in}}$. Hence, $\forall k \in [p], \mathrm{dist}_H(v^{(k)}, \mathcal{H}_a) \leqslant r_{\mathcal{V}}^{\mathrm{in}}$, so that by Lemma 4.18 we have $\forall k \in [p], \forall i \in [n]$,

$$V_{ik} + a_i \leqslant r_{\mathcal{V}}^{\mathrm{in}} + \max_{j \in [n], j \neq i} (V_{jk} + a_j).$$

Therefore, we obtain

$$\forall i \in [n], \forall k \in [p], (i,k) \in E; -r^{\text{in}}_{\mathcal{V}} + a_i \leqslant -V_{ik} + \max_{j\in[n], j\neq i}(V_{jk} + a_j) \ . \tag{4.24}$$

Assertion (2) also implies $\forall i \in [n], \exists k \in [p], V_{ik} + a_i \geqslant \max_{j\in[n],j\neq i}(V_{jk} + a_j)$ and $\operatorname{dist}_H(v^{(k)}, \mathcal{H}_a) = r^{\text{in}}_{\mathcal{V}}$, with $V_{ik} \neq -\infty$ because $a \in \mathbb{R}^n$ and $V_{\cdot k} \neq \bot$. This means, by Lemma 4.18, that $\forall i \in [n], \exists k \in [p], (i,k) \in E, V_{ik} + a_i = r^{\text{in}}_{\mathcal{V}} + \max_{j\in[n],j\neq i}(V_{jk} + a_j)$. Then

$$\forall i \in [n], \exists k \in [p], (i,k) \in E; -r^{\text{in}}_{\mathcal{V}} + a_i = -V_{ik} + \max_{j\in[n], j\neq i}(V_{jk} + a_j) \ . \tag{4.25}$$

From (4.24) and (4.25), we conclude that

$$\forall i \in [n], -r^{\text{in}}_{\mathcal{V}} + a_i = \min_{k\in[p],(i,k)\in E}[-V_{ik} + \max_{j\in[n],j\neq i}(V_{jk} + a_j)] = T_i(a).$$

Therefore $T(a) = \rho(T) + a$.

The final part of the theorem follows from Theorem 4.23. □

Remark 4.26. When $T(a) = \rho(T)+a$ and $a \in \mathbb{R}^n$, Theorem 4.25 and Theorem 4.23 entail the following remarkable property: there is an optimal Hilbert's ball whose center coincides with the apex of an optimal regression hyperplane. This property is illustrated in Figure 4.4 below.

Remark 4.27. The situation in which $T(a) = \bot$ holds for some finite vector a (or equivalently, for all finite vectors a) is degenerate. Indeed, we observe from the proof of Theorem 4.25 that $T(a) = \bot$ for some finite vector a if and only if, for all $i \in [n]$, there is a vector $v^{(k)}$ such that $v^{(k)}_i$ is finite and all $v^{(k)}_j$ with $j \neq i$ are $-\infty$. Then, V contains a $n \times n$ diagonal submatrix, and so, $\operatorname{Col}(V) = \operatorname{Sp}(\mathcal{V}) = (\mathbb{R}_{\max})^n$.

We next exhibit a situation in which the existence of a finite eigenvector, required to apply Theorem 4.25, is guaranteed.

Proposition 4.28. *Suppose that all the vectors $v \in \mathcal{V}$ have finite entries. Then, the operator T has a finite eigenvector a.*

Proof. Theorems 9 and 13 of [GG04] imply that an order preserving and additively homogeneous map $T : \mathbb{R}^n \to \mathbb{R}^n$ has a finite eigenvector if the recession function $\hat{T}(x) := \lim_{s\to\infty} s^{-1}T(sx)$ has only fixed points on the diagonal. When the matrix V is finite, considering $T := T_V$, we have $\hat{T}_i(x) = \max_{j\in[n],j\neq i} x_j$, for all $i \in [n]$, so the latter condition is trivially satisfied. This entails that there exists a vector $a \in \mathbb{R}^n$ such that $T(a) = \rho(T) + a$. □

A more general condition, involving the notion of dominions, is given in Section 4.8.

The following proposition shows that we can determine witness points from a policy $\sigma : [n] \mapsto [p]$, that satisfies $T(a) = T^\sigma(a)$ where a is a finite eigenvector of the operator T. For an illustration of this lemma see Figure 4.2.

Proposition 4.29. *Let $a \in \mathbb{R}^n$ such that $T(a) = -r^{\text{in}}_{\mathcal{V}} + a$, and $\sigma : [n] \mapsto [p]$ a map, such that $\forall i \in [n], (i, \sigma(i)) \in E$. We have $T(a) = T^\sigma(a)$ if and only if for all $i \in [n]$, $V_{\cdot\sigma(i)}$ is a witness point of $\mathcal{H}_a$ that belongs to the sector $S_i(a)$.*

Proof. If $T(a) = T^\sigma(a)$, then $T^\sigma(a) = -r^{\text{in}}_{\mathcal{V}} + a$. Therefore, we have for all $i \in [n]$, $-V_{i\sigma(i)} + \max_{j\neq i}(V_{j\sigma(i)} + a_j) = -r^{\text{in}}_{\mathcal{V}} + a_i$, i.e. $V_{i\sigma(i)} + a_i = r^{\text{in}}_{\mathcal{V}} + \max_{j\neq i}(V_{j\sigma(i)} + a_j)$, which means that $V_{\cdot\sigma(i)} \in S_i(a)$ and, by Lemma 4.18, that $\operatorname{dist}_H(V_{\cdot\sigma(i)}, \mathcal{H}_a) = r^{\text{in}}_{\mathcal{V}}$, i.e for all $i \in [n]$, $V_{\cdot\sigma(i)}$ is a witness point in the sector $S_i(a)$.

Conversely, if for all $i \in [n]$, $V_{\cdot\sigma(i)}$ is a witness point in the sector $S_i(a)$. Let $i \in [n]$, we have then $V_{i\sigma(i)} + a_i = r^{\text{in}}_{\mathcal{V}} + \max_{j\neq i}(V_{j\sigma(i)} + a_j)$, i.e. $-V_{i\sigma(i)} + \max_{j\neq i}(V_{j\sigma(i)} + a_j) = -r^{\text{in}}_{\mathcal{V}} + a_i$. We know that for all $k \in [p]$, $\text{dist}_H(V_{\cdot k}, \mathcal{H}_a) \leqslant r^{\text{in}}_{\mathcal{V}}$, then by Lemma 4.18, $V_{ik} + a_i \leqslant r^{\text{in}}_{\mathcal{V}} + \max_{j\neq i}(V_{jk} + a_j)$, i.e. $-V_{ik} + \max_{j\neq i}(V_{jk} + a_j) \geqslant -r^{\text{in}}_{\mathcal{V}} + a_i$. Therefore, $T_i(a) = \inf_{k\in[p],(i,k)\in E}\left[-V_{ik} + \max_{j\in[n],j\neq i}(V_{jk} + a_j)\right] = -V_{i\sigma(i)} + \max_{j\neq i}(V_{j\sigma(i)} + a_j) = T^{\sigma}_i(a)$. □

We now formalize the tropical linear regression problem:

Problem 3 (Tropical linear regression)**.** Input: a finite set of vectors $\mathcal{V} \subset \mathbb{Z}^n_{\max}$. Goal: compute the infimum of the one-sided Hausdorff distance of $\mathcal{V}$ to a tropical hyperplane, i.e., the value of the optimization problem (4.17).

Corollary 4.30. *The tropical linear regression problem (Problem 3) is polynomial time Turing-equivalent to mean payoff games (Problem 1).*

Proof. This follows from the strong duality theorem (Theorem 4.23) and Corollary 4.15. □

Corollary 4.31. *Computing an optimal regression hyperplane $\mathcal{H}_a$ in* (4.17)*, given a finite set of vectors $\mathcal{V} \subset \mathbb{Z}^n_{\max}$, polynomially Turing-reduces to mean payoff games.*

Proof. By Theorem 4.23, we need to find a vector a such that $T(a) \geqslant \rho(T) + a$. Arguing as in the proof of Corollary 4.16, but exchanging the roles of Player Max and Min, we end up with an optimal policy τ of Player Max. Then, it suffices to find a vector $a \in (\mathbb{R}\cup\{-\infty\})^n$, $a \neq \bot$, such that ${}^{\tau}T(a) \geqslant \rho({}^{\tau}T) + a$. Still arguing as in the proof of Corollary 4.16, we are reduced to a problem of tropical (min-plus instead max-plus) spectral theory, which again reduces to a shortest path problem. □

In Figure 4.3, we consider the same matrix V as in Figure 4.2. The Figure 4.3 shows the witness points in each of the sectors defined by the hyperplane $\mathcal{H}_a$ where $a = (0, 0, 1)^\top$ satisfies $T(a) = \lambda + a$ with $\lambda = -1$. In this example, we have two witness points in each sector: $V_{\cdot 4}$ and $V_{\cdot 5}$ are the witness points in the sector $S_1(a)$ (in green), $V_{\cdot 6}$ and $V_{\cdot 7}$ are the witness points in the sector $S_2(a)$ (in blue) and $V_{\cdot 8}$ and $V_{\cdot 9}$ are the witness points in the sector $S_3(a)$ (in red).

In Figure 4.4, we consider the following matrix $U \in \mathbb{R}^{3\times 4}$:

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

The operator associated to U is the following map $T : (\mathbb{R}_{\max})^n \to (\mathbb{R}_{\max})^n$:

$$T\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} \min[1 + \max(x_2, x_3), \max(-1 + x_2, x_3), -1 + \max(x_2, -2 + x_3), \max(1 + x_2, -2 + x_3)] \\ \min[\max(-1 + x_1, x_3), 1 + \max(x_1, x_3), \max(1 + x_1, -2 + x_3), -1 + \max(x_1, -2 + x_3)] \\ \min[\max(-1 + x_1, x_2), 1 + \max(x_1, -1 + x_2), 2 + \max(1 + x_1, x_2), 2 + \max(x_1, 1 + x_2)] \end{pmatrix}.$$

We verify easily that $\lambda = -1$ and $a = (0, 0, 1)^\top$ satisfy $T(a) = \lambda + a$, so that the inner radius of $\text{Col}(U)$ is $r^{\text{in}}_{\mathcal{U}} = 1$. In this example, other hyperplanes like $\mathcal{H}_b$ and $\mathcal{H}_c$, with $b = (0, 0, -1)^\top$ and $c = (0, 0, -\infty)^\top$, are also optimal solutions of the tropical linear regression problem, but $\mathcal{H}_a$ is the only hyperplane such that a is a finite eigenvector of the operator T and, hence, that satisfies also $B(-a, 1) \subset \text{Col}(U)$.

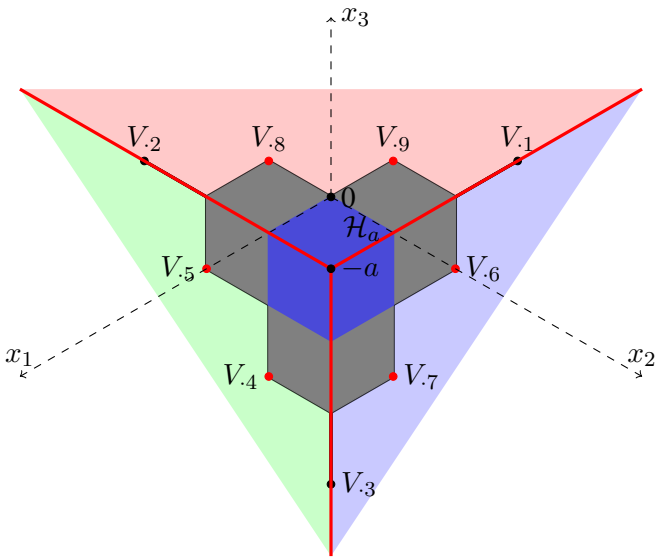


Figure (4.3) The inner ball of a column space $\mathrm{Col}(V)$ and the linear regression of the columns of V.

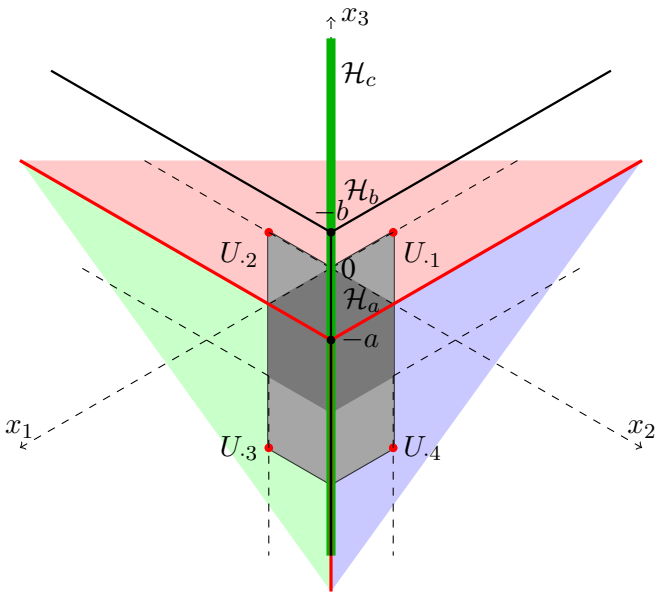


Figure (4.4) A column space $\mathrm{Col}(U)$ (light and dark gray regions) with multiple hyperplanes that are optimal solutions of the tropical linear regression problem, and multiple inner balls of maximal radius, but a unique optimal hyperplane with witness points in each sector, corresponding to the finite eigenvector $a = (0, 0, 1)^\top$ of T and to the inner ball in dark gray.

4.5 Tropical linear regression with sign or type patterns

Here, we study several variants of the tropical linear regression problem, which can also be solved by the present technique of reduction to a mean payoff game. The second of these variants (with "types") will arise in the economic application of Section 4.7.

4.5.1 Tropical linear regression with signs

Given $I, J \subset [n]$ such that $I, J \neq \emptyset$, $I \cup J = [n]$ and $I \cap J = \emptyset$ and $a \in \mathbb{P}(\mathbb{R}_{\max})^n$, we define the *signed tropical hyperplane* of type (I, J):

$$\mathcal{H}_a^{IJ} := \{x \in (\mathbb{R}_{\max})^n \mid \max_{i \in I}(a_i + x_i) = \max_{j \in J}(a_j + x_j)\} \tag{4.26}$$

Given a set $\mathcal{V} \subset (\mathbb{R}_{\max})^n$, of cardinality $|\mathcal{V}| = p$, the *signed tropical linear regression problem* of type (I, J) consists in finding the best approximation of $\mathcal{V}$ by a signed hyperplane of type (I, J):

$$\min_{a \in \mathbb{P}(\mathbb{R}_{\max})^n} \operatorname{dist}_H(\mathcal{V}, \mathcal{H}_a^{IJ}) \ . \tag{4.27}$$

Let M be a closed tropical cone of $(\mathbb{R}_{\max})^n$ and $x \in (\mathbb{R}_{\max})^n$. The projection $P_M(x)$ of the point x onto M [CGQ04] is defined by:

$$P_M(x) := \max\{z \in M \mid z \leqslant x\} \ . \tag{4.28}$$

The Hilbert's distance from x to M is achievd by the projection $P_M(x)$.

Theorem 4.32 ([CGQ04])**.** *Given a closed tropical semimodule $M \subset (\mathbb{R}_{\max})^n$ and $x \in (\mathbb{R}_{\max})^n$, we have:*

$$\operatorname{dist}_H(x, M) = d(x, P_M(x)) \ .$$

The following lemma identifies the projection of a point $x \in \mathbb{P}(\mathbb{R}_{\max})^n$ onto a signed tropical hyperplane $\mathcal{H}_a^{IJ}$.

Lemma 4.33. *Let $x, a \in \mathbb{P}(\mathbb{R}_{\max})^n$ and $K = \operatorname{supp} a$. The projection $P_{\mathcal{H}_a^{IJ}}(x)$ of x onto $\mathcal{H}_a^{IJ}$ is given by:*

$$[P_{\mathcal{H}_a^{IJ}}(x)]_l = \begin{cases} x_l & \text{, for } l \in K^c \\ \min\{x_l, -a_l + \max_{j \in J}(a_j + x_j)\} & \text{, for } l \in I \cap K \\ \min\{x_l, -a_l + \max_{i \in I}(a_i + x_i)\} & \text{, for } l \in J \cap K \end{cases} \tag{4.29}$$

where K^c denotes the complementary of K in $[n]$.

Proof. Denote the right hand side vector of (4.29) by $\tilde{x}$. From (4.28), we have $P_{\mathcal{H}_a^{IJ}}(x) = \max\{z \in \mathcal{H}_a^{IJ} \mid z \leqslant x\}$. Let $z \in \mathcal{H}_a^{IJ}$ such that $z \leqslant x$. We will prove that $z \leqslant \tilde{x}$. Let $l \in I$, if $l \in I \cap K^c$, we have right away that $z_l \leqslant x_l = \tilde{x}_l$. Now if $l \in I \cap K$, knowing that $z \in \mathcal{H}_a^{IJ}$ and using (4.26), we have $a_l + z_l \leqslant \max_{i \in I}(a_i + z_i) = \max_{j \in J}(a_j + z_j) \leqslant \max_{j \in J}(a_j + x_j)$. Then, $z_l \leqslant -a_l + \max_{j \in J}(a_j + x_j)$. We know also that $z_l \leqslant x_l$, then $z_l \leqslant \tilde{x}_l$. Similarly the inequality $z_l \leqslant \tilde{x}_l$ can also be proved for all $l \in J$. Therefore, for all $z \in \mathcal{H}_a^{IJ}$, if $z \leqslant x$ then $z \leqslant \tilde{x}$. Using (4.28), it suffices now to prove that $\tilde{x} \in \mathcal{H}_a^{IJ}$. Indeed, $\max_{i \in I}(a_i + \tilde{x}_i) = \max_{i \in I \cap K}(a_i + \tilde{x}_i) = \max_{i \in I \cap K}\{\min(a_i + x_i, \max_{j \in J}(a_j + x_j))\} = \min\{\max_{i \in I \cap K}(a_i + x_i), \max_{j \in J}(a_j + x_j)\} = \min\{\max_{i \in I}(a_i + x_i), \max_{j \in J}(a_j + x_j)\}$, and by symmetry we deduce that $\max_{j \in J}(a_j + \tilde{x}_j)$ is also equal to the same quantity, and so $\tilde{x} \in \mathcal{H}_a^{IJ}$. $\square$

Remark 4.34. The formula of Lemma 4.33 may be compared with formula for the projection of a point onto a tropical half-space $\{x \in (\mathbb{R}_{\max})^n \mid \max_{i\in I}(a_i + x_i) \leqslant \max_{j\in J}(a_j + x_j)\}$, see [AGNS11, Th. 5.1].

Proposition 4.35. *Let $x, a \in \mathbb{P}(\mathbb{R}_{\max})^n$. The Hilbert's distance of the point x to the signed hyperplane $\mathcal{H}_a^{IJ}$ is:*

$$\mathrm{dist}_H(x, \mathcal{H}_a^{IJ}) = |\max_{i\in I}(x_i + a_i) - \max_{j\in J}(x_j + a_j)| \ , \tag{4.30}$$

if at least one of these maxima is finite, and $\mathrm{dist}_H(x, \mathcal{H}_a^{IJ}) = 0$ *otherwise.*

Proof. From Theorem 4.32, we have $\mathrm{dist}_H(x, \mathcal{H}_a^{IJ}) = d(x, \tilde{x})$ with $\tilde{x} = P_{\mathcal{H}_a^{IJ}}(x)$. Let $K = \mathrm{supp}\, a$ and $O = \mathrm{supp}\, x$. If $I \cap K \cap O = J \cap K \cap O = \emptyset$, then $K \cap O = \emptyset$, so $x + a \equiv -\infty$, and this means that $x \in \mathcal{H}_a^{IJ}$ and so $\mathrm{dist}_H(x, \mathcal{H}_a^{IJ}) = 0$.

If $I \cap K \cap O = \emptyset$ and $J \cap K \cap O \neq \emptyset$, then $\max_{i\in I}(x_i + a_i) = -\infty$ and $\max_{j\in J}(x_j + a_j) \neq -\infty$. Let $j \in J \cap K \cap O$, we have $\tilde{x}_j = \min\{x_j, -a_j + \max_{i\in I}(a_i + x_i)\} = -\infty$ and we have $x_j \neq -\infty$, then $d(x, \tilde{x}) = +\infty = |\max_{i\in I}(x_i + a_i) - \max_{j\in J}(x_j + a_j)|$. By symmetry we treat the case when $I \cap K \cap O \neq \emptyset$ and $J \cap K \cap O = \emptyset$.

Now, we suppose that $I \cap K \cap O \neq \emptyset$ and $J \cap K \cap O \neq \emptyset$. Let $i \in I \cap K \cap O$, we have $x_i - \tilde{x}_i = x_i + \max\{-x_i, a_i - \max_{j\in J}(a_j + x_j)\} = \max\{0, x_i + a_i - \max_{j\in J}(a_j + x_j)\}$. Then, we have $\max_{i\in I}(x_i - \tilde{x}_i) = \max_{i\in I\cap K\cap O}(x_i - \tilde{x}_i) = \max\{0, \max_{i\in I\cap K\cap O}(x_i + a_i) - \max_{j\in J}(a_j + x_j)\} = \max\{0, \max_{i\in I}(x_i + a_i) - \max_{j\in J}(a_j + x_j)\}$, and symmetrically, we have $\max_{j\in J}(x_j - \tilde{x}_j) = \max\{0, \max_{j\in J}(x_j + a_j) - \max_{i\in I}(a_i + x_i)\}$. Therefore, we deduce that $\max_{l\in[n]}(x_l - \tilde{x}_l) = |\max_{i\in I}(x_i + a_i) - \max_{j\in J}(x_j + a_j)|$.

To finish the proof we need now to show that $\min_{l\in[n]}(x_l - \tilde{x}_l) = 0$. This is a general property of the projection $\tilde{x} = P_M(x)$ of a vector on a closed tropical cone: since $\tilde{x} \leqslant x$, the minimum is nonnegative, and if the minimum is positive, adding a small constant ϵ to every entry of $\tilde{x}$, we get a vector $\tilde{x}^\epsilon$ which still belongs to M and satisfies $\tilde{x}^\epsilon \leqslant x$, contradicting $P_M(x) = \max\{z \in M \mid z \leqslant x\}$. □

In the sequel, we suppose that the following Assumption F holds.

Assumption F. We suppose that for each $l \in [n]$, there exists $v \in \mathcal{V}$, such that $v_l \neq -\infty$.

We now introduce the operator $T^{IJ} : (\mathbb{R}_{\max})^n \mapsto (\mathbb{R}_{\max})^n$, defined by:

$$T_l^{IJ}(x) := \begin{cases} \inf_{v\in\mathcal{V}, v_l\neq-\infty}\{-v_l + \max_{j\in J}(v_j + x_j)\}, \text{ if } l \in I \ , \\ \inf_{v\in\mathcal{V}, v_l\neq-\infty}\{-v_l + \max_{i\in I}(v_i + x_i)\}, \text{ if } l \in J \ . \end{cases} \tag{4.31}$$

The following result, analogous to Lemma 4.22, gives a metric interpretation of the sub-eigenspace of the operator T^{IJ}.

Lemma 4.36. *Let $\lambda \in [-\infty, 0]$ and $a \in \mathbb{P}(\mathbb{R}_{\max})^n$, we have*

$$T^{IJ}(a) \geqslant \lambda + a \Leftrightarrow \mathrm{dist}_H(\mathcal{V}, \mathcal{H}_a^{IJ}) \leqslant -\lambda \ .$$

Proof. The equivalence is trivial if $\lambda = -\infty$, so, we suppose that $\lambda \in (-\infty, 0]$. We have

$$T^{IJ}(a) \geqslant \lambda + a \Leftrightarrow \begin{cases} \forall l \in I, \forall v \in \mathcal{V}, v_l \neq -\infty; -v_l + \max_{j\in J}(v_j + a_j) \geqslant \lambda + a_l \\ \forall l \in J, \forall v \in \mathcal{V}, v_l \neq -\infty; -v_l + \max_{i\in I}(v_i + a_i) \geqslant \lambda + a_l \end{cases} \tag{4.32}$$

$$\Leftrightarrow \begin{cases} \forall v \in \mathcal{V}, \max_{l\in I}(v_l + a_l) \leqslant \max_{j\in J}(v_j + a_j) - \lambda \\ \forall v \in \mathcal{V}, \max_{l\in J}(v_l + a_l) \leqslant \max_{i\in I}(v_i + a_i) - \lambda \end{cases} \tag{4.33}$$

Let $\mathcal{V}'$ denote the set of vectors $v \in \mathcal{V}$ for which at least one of the latter maxima are finite, and observe that the vectors of $\mathcal{V} \setminus \mathcal{V}'$ trivially belong to $\mathcal{H}_a^{I,J}$. Then, using Proposition 4.35, we see that the last condition in (4.33) is equivalent to

$$\forall v \in \mathcal{V}', d(v, \mathcal{H}_a^{IJ}) = |\max_{i\in I}(v_i + a_i) - \max_{j\in J}(v_j + a_j)| \leqslant -\lambda$$

i.e., $\mathrm{dist}_H(\mathcal{V}, \mathcal{H}_a^{IJ}) \leqslant -\lambda$. □

Let $w \in \mathbb{R}^n$ and $r \geqslant 0$, we define the *vertical interval* of type I, J centered at point w and of radius r,

$$B_{IJ}(w, r) = \{\lambda + w + \mu e^J \mid \mu \in [-r, r], \lambda \in \mathbb{R}\} \ ,$$

where e^J is the vector of $\mathbb{R}^n$ such that $e_l^J = 0$ for $l \in I$ and $e_l^J = 1$ for $l \in J$. Using the identity $-\mu + \mu e^J = -\mu e^I$, we see

$$B_{IJ}(w, r) = \{\lambda + w + \mu e^I \mid \mu \in [-r, r], \lambda \in \mathbb{R}\} \ .$$

Lemma 4.37. *Let $\lambda \in [-\infty, 0]$ and $a \in \mathbb{R}^n$, we have*

$$B_{IJ}(-a, -\lambda) \subset \mathrm{Sp}(\mathcal{V}) \Rightarrow T^{IJ}(a) \leqslant \lambda + a \ .$$

Proof. Suppose first that λ is finite. If $B_{IJ}(-a, -\lambda) \subset \mathrm{Sp}(\mathcal{V})$, then

$$\forall \mu \in [-\lambda, \lambda], \exists (\alpha_v)_{v\in\mathcal{V}} \in \mathbb{R}^p, -a + \mu e^J = \max_{v\in\mathcal{V}}(\alpha_v + v).$$

Let $\mu \in [-\lambda, \lambda]$, we have

$$\forall i \in I, (\forall v \in \mathcal{V}, -a_i \geqslant \alpha_v + v_i \text{ and } \exists v^{(i)} \in \mathcal{V}, -a_i = \alpha_{v^{(i)}} + v_i^{(i)}) \ , \tag{4.34}$$

and also

$$\forall j \in J, (\forall v \in \mathcal{V}, -a_j + \mu \geqslant \alpha_v + v_j \text{ and } \exists v^{(j)} \in \mathcal{V}, -a_j + \mu = \alpha_{v^{(j)}} + v_j^{(j)}) \ . \tag{4.35}$$

From (4.35), we have $\forall v \in \mathcal{V}, \sup_{j\in J}(v_j + a_j) \leqslant -\alpha_v + \mu$, and from (4.34), we have $\forall i \in I, v_i^{(i)} \neq -\infty$ because $\alpha_{v^{(i)}} + v_i^{(i)} = -a_i \in \mathbb{R}$. Then, for all $i \in I$, we have $T_i^{IJ}(a) = \inf_{v\in\mathcal{V}, v_i\neq-\infty}\{-v_i + \sup_{j\in J}(v_j + a_j)\} \leqslant \inf_{v\in\mathcal{V}, v_i\neq-\infty}\{-v_i - \alpha_v + \mu\} \leqslant -v_i^{(i)} - \alpha_{v^{(i)}} + \mu = \mu + a_i$. This being true for all $\mu \in [-\lambda, \lambda]$, we take here $\mu = \lambda$ and we get that $\forall i \in I, T_i^{IJ}(a) \leqslant \lambda + a_i$.

Similarly, we have $\forall j \in J, T_j^{IJ}(a) = \inf_{v\in\mathcal{V}, v_j\neq-\infty}\{-v_j + \sup_{i\in I}(v_i + a_i)\} \leqslant \inf_{v\in\mathcal{V}, v_j\neq-\infty}\{-v_j - \alpha_v\} \leqslant -v_j^{(j)} - \alpha_{v^{(j)}} = -\mu + a_j$. By taking here $\mu = -\lambda$, we get that $\forall j \in J, T_j^{IJ}(a) \leqslant \lambda + a_j$. Therefore, we get that $T^{IJ}(a) \leqslant \lambda + a$.

The conclusion of the lemma is still true when $\lambda = -\infty$. This follows from $B_{IJ}(-a, +\infty) = \cup_{\mu\in(-\infty,0)} B_{IJ}(-a, -\mu)$ and $-\infty + a = \inf_{\mu\in(-\infty,0)} \mu + a$. □

Lemma 4.38. *Let $\lambda \in [-\infty, 0]$, we have*

$$\exists u \in \mathbb{R}^n; T^{IJ}(u) \leqslant \lambda + u \Rightarrow \exists w \in \mathbb{R}^n; B_{IJ}(w, -\lambda) \subset \mathrm{Sp}(\mathcal{V}) \ .$$

Proof. Suppose first that λ is finite. For simplicity of notation, we shall assume that $u = 0$. The general case reduces to this one by replacing every vector $v \in \mathcal{V}$ by the vector $v + u$. We denote by V the matrix whose columns are the elements of $\mathcal{V}$. Since $T^{IJ}(u) \leqslant \lambda + u$, denoting by σ a map $[n] \to [p]$ such that for all $l \in [n]$, $v = V_{\cdot\sigma(l)}$ achieves the minimum in (4.31), we get:

$$\forall l \in I,\ \forall j \in J, \qquad V_{j\sigma(l)} \leqslant V_{l\sigma(l)} + \lambda \ , \tag{4.36}$$

$$\forall l \in J,\ \forall i \in I, \qquad V_{i\sigma(l)} \leqslant V_{l\sigma(l)} + \lambda \ . \tag{4.37}$$

Consider the vectors

$$w^I := \vee_{i\in I} - V_{i\sigma(i)} + V_{\cdot\sigma(i)}, \qquad w^J := \vee_{j\in J} - V_{j\sigma(j)} + V_{\cdot\sigma(j)},$$

so that $w^I, w^J \in \mathrm{Sp}(\mathcal{V})$. By considering the values $i = l$ or $j = l$ in the suprema above, we get

$$w^I_l \geqslant 0, \forall l \in I, \qquad w^J_l \geqslant 0, \forall l \in J \ . \tag{4.38}$$

Moreover, using (4.36), we get

$$w^I_j = \vee_{i\in I} - V_{i\sigma(i)} + V_{j\sigma(i)} \leqslant \lambda, \qquad \text{for all } j \in J \ , \tag{4.39}$$

and similarly, using (4.37),

$$w^J_i \leqslant \lambda, \text{ for all } i \in I \ . \tag{4.40}$$

Define the vector w by

$$w_l = \begin{cases} w^I_l, \text{ if } l \in I \ , \\ w^J_l, \text{ if } l \in J \ . \end{cases}$$

Using (4.38)–(4.40), we deduce that for all $\mu \in [\lambda, -\lambda]$,

$$w + \mu e^I = (w^I + \mu) \vee w^J \in \mathrm{Sp}(\mathcal{V}) \ ,$$

and so $B_{IJ}(w, -\lambda) \subset \mathrm{Sp}(\mathcal{V})$.

We finally show that the conclusion of the lemma is still true when $\lambda = -\infty$. This follows from the fact that the above center w depends only on the vectors of $\mathcal{V}$ and does not depend on λ, and also from the facts that $B(w, +\infty) = \cup_{\mu\in(-\infty,0)} B(w, -\mu)$ and $-\infty + w = \inf_{\mu\in(-\infty,0)} \mu + w$. □

The next result is immediate from Lemmas 4.37 and 4.38. It is analogous to Lemma 4.8. It shows that the existence of a super-eigenvector of $T^{I,J}$ is equivalent to the existence of a *vertical interval* included in the module $\mathrm{Sp}(\mathcal{V})$.

Proposition 4.39. *Let $\lambda \in [-\infty, 0]$, and $a \in \mathbb{R}^n$, we have*

$$\exists u \in \mathbb{R}^n; T^{IJ}(u) \leqslant \lambda + u \Leftrightarrow \exists w \in \mathbb{R}^n; B_{IJ}(w, -\lambda) \subset \mathrm{Sp}(\mathcal{V}) \ . \qquad \square$$

We now derive a strong duality theorem for signed tropical regression.

Theorem 4.40. *We have*

$$\min_{a\in\mathbb{P}(\mathbb{R}_{\max})^n} \mathrm{dist}_H(\mathcal{V}, \mathcal{H}^{IJ}_a) = -\rho(T^{IJ}) = \sup\{r \geqslant 0 \mid \exists w \in \mathbb{R}^n, B_{IJ}(w, r) \subset \mathrm{Sp}(\mathcal{V})\} \ . \tag{4.41}$$

The minimum is achieved by any vector $b \in \mathbb{P}(\mathbb{R}_{\max})^n$ such that $T^{IJ}(b) \geqslant \rho(T^{IJ}) + b$. Moreover, if $\rho(T^{IJ})$ is finite, the supremum is achieved by a ball $B(c, \rho(T^{IJ}))$ where $c \in \mathbb{R}^n$ can be deduced from any vector u such that $T^{IJ}(u) \leqslant \rho(T^{IJ}) + u$.

Proof. From Proposition 4.39, Lemma 4.36 and the Collatz-Wielandt property (Theorem 4.1), we deduce the strong duality property (4.41). Moreover, the existence of a vector $b \in \mathbb{P}(\mathbb{R}_{\max})^n$ such $T^{IJ}(b) \geqslant \rho(T^{IJ}) + b$ follows from Theorem 4.1. Then, by Lemma 4.36, we have $\mathrm{dist}_H(\mathcal{V}, \mathcal{H}_b) \leqslant -\rho(T^{IJ})$, which implies that b such that $T^{IJ}(b) \geqslant \rho(T^{IJ}) + b$ achieves the minimum in (4.41).

Finally, if $\rho(T^{IJ})$ is finite, since the infimum is attained in the expression of the Collatz-Wielandt number of T^{IJ} (see Theorem 4.1), there exists a finite vector $u \in \mathbb{R}^n$ such that $T^{IJ}(u) \leqslant \rho(T^{IJ}) + u$. By the proof of Lemma 4.38, we can then construct a vector c such that $B_{IJ}(c, -\rho(T^{IJ})) \subset \mathrm{Sp}(\mathcal{V})$. □

Remark 4.41. When the set $I = \{i\}$ is of cardinality one, the regression problem for the signed hyperplane (4.26) has the following special form:

$$\operatorname*{Min}_{a\in\mathbb{R}^n} \max_{v\in\mathcal{V}} |v_i - (\max_{j\neq i} a_j - a_i + v_j)| \ . \tag{4.42}$$

This can be solved in a direct way [MCT21], avoiding the recourse to mean payoff games.Indeed, (4.42) reduces to the following "one-sided" tropical linear regression problem. Given sample points $(x^{(k)}, y^{(k)})$ in $\mathbb{R}^n \times \mathbb{R}^m$, for $k \in [p]$, compute

$$\operatorname*{Min}_{A} \max_{k\in[p]} \|y^{(k)} - Ax^{(k)}\|_\infty \ , \tag{4.43}$$

where the minimum is taken over tropical matrices A of size $m \times n$, and the product $Ax^{(k)}$ is understood tropically. Up to a straightforward duality, this problem was solved in [But10, Theorem 3.5.2], the result being attributed there to Cuninghame-Green [CG79]. Alternatively, this solution may be recovered by combining [CF00, Coro. 1] with the explicit formula of the tropical projection [CGQ04, Th. 5]. More precisely, define the matrix $\bar{A} \in \mathbb{R}^{m\times n}$ by $\bar{A}_{ij} := \min_{k\in[p]} y_i^{(k)} - x_j^{(k)}$, so that $\bar{A}$ is the maximal matrix such that $Ax^{(k)} \leqslant y^{(k)}$ for all $k \in [p]$. Let $\delta := \max_{k\in p} \|y^{(k)} - \bar{A}x^{(k)}\|_\infty$, and $A_{ij}^{\mathrm{opt}} = \bar{A}_{ij} + \delta/2$. Then, A^{opt} is the greatest optimal solution. It can be computed in $O(mnp)$ arithmetic operations. By specializing this formula, one can solve (4.42) in $O(np)$ arithmetic operations. We refer the reader to [MCT21] for more information, and for the solution of further problems of this category.

Remark 4.42. In contrast, when I, J are part of the input, the signed linear tropical regression problem is polytime Turing equivalent to mean payoff games. This can be seen as follows. The reduction to mean payoff games is a consequence of Theorem 4.40. Conversely, observe that finding a signed tropical hyperplane $\mathcal{H}_a^{I,J}$ containing a set $\mathcal{V} = \{v^{(1)}, \dots, v^{(p)}\}$ in $\mathbb{R}^n$ is equivalent to solving a tropical linear system of the form $Bx = Cy$, where $x \in (\mathbb{R}_{\max})^I$, $y \in (\mathbb{R}_{\max})^J$, $B \in (\mathbb{R}_{\max})^{p\times I}$, $C \in (\mathbb{R}_{\max})^{p\times J}$, $B_{ki} = v_i^{(k)}$ for $i \in I$ and $C_{kj} = v_j^{(k)}$ for $j \in I$. Indeed, the vector a defining $\mathcal{H}_a^{I,J}$ is given by $a_i = x_i$ for $i \in I$ and $a_j = y_j$ for $j \in J$. We know from [AGG12] that deciding whether a mean payoff game has an initial winning position is equivalent to the existence of a non-identically $-\infty$ solution $z \in (\mathbb{R}_{\max})^s$ of a system of tropical linear inequalities $Fz \leqslant Gz$, where $F, G \in (\mathbb{Z}_{\max})^{r\times s}$ are given. Such a system $Fz \leqslant Gz$ can be rewritten as $Bx = Cy$ by introducing lift variables $u, v \in (\mathbb{R}_{\max})^r$, so that $v = Fz$ and u is a slack variable. Setting $y := (u, v)$, identified to a column vector, $B := \left(\begin{smallmatrix} -\infty & \mathrm{I} \\ \mathrm{I} & \mathrm{I} \end{smallmatrix}\right)$ and $C := \left(\begin{smallmatrix} F \\ G \end{smallmatrix}\right)$, where $-\infty$ is a zero tropical matrix, and I the identity matrix, we see that $Fz \leqslant Gz$ has a non-identically $-\infty$ solution iff $By = Cz$ has a non-identically $-\infty$ solution. It follows that mean payoff games reduce to checking whether there is a solution of a signed tropical linear regression problem with zero error.

4.5.2 Tropical linear regression with type information

The following variant will be relevant to the application to economy considered below, to measure the "distance to equilibria" of a market. We suppose the set of points $\mathcal{V}$ is the disjoint union $\mathcal{V} = \cup_{i\in[n]} \mathcal{V}_i$,

where each $\mathcal{V}_i$ is non-empty. We shall say that the points of $\mathcal{V}_i$ are of *type* $i \in [n]$. Note that the set of types is the same as the set of indices of vectors. For each type $i \in [n]$, we consider the signed hyperplane:

$$\mathcal{H}_a^i := \mathcal{H}_a^{\{i\}\{i\}^c} = \{x \in (\mathbb{R}_{\max})^n \mid a_i + x_i = \max_{j \neq i}(a_j + x_j)\} \ .$$

The *typed tropical linear regression* problem associated to the partition $\mathcal{V}_1, \ldots, \mathcal{V}_n$ of $\mathcal{V}$, is defined as:

$$\operatorname*{Min}_{a \in \mathbb{P}(\mathbb{R}_{\max})^n} \max_{i \in [n]} \operatorname{dist}_H(\mathcal{V}_i, \mathcal{H}_a^i) \ . \tag{4.44}$$

The value of this problem is small if and only if for each $i \in [n]$, the points of $\mathcal{V}_i$ are close to the signed tropical hyperplane $\mathcal{H}_a^i$.

From Proposition 4.35, we know that $\operatorname{dist}_H(v, \mathcal{H}_a^i) = |v_i + a_i - \max_{j \neq i}(v_j + a_j)|$.

We suppose in the sequel that Assumption F holds. For each type $i \in [n]$, we consider the Shapley operator $T^{\mathrm{ty},i} : (\mathbb{R}_{\max})^n \mapsto (\mathbb{R}_{\max})^n$, given by (4.31) where the type considered is $(I, J) = (\{i\}, \{i\}^c)$ and to the set of points is $\mathcal{V}_i$:

$$T_l^{\mathrm{ty},i}(x) := \begin{cases} \inf_{v \in \mathcal{V}_i, v_i \neq -\infty}\{-v_i + \max_{j \neq i}(v_j + x_j)\}, \text{ if } l = i \ , \\ \inf_{v \in \mathcal{V}_i, v_l \neq -\infty}\{-v_l + v_i\} + x_i, \text{ if } l \neq i \ . \end{cases} \tag{4.45}$$

We consider now the Shapley operator $T^{\mathrm{ty}} : (\mathbb{R}_{\max})^n \mapsto (\mathbb{R}_{\max})^n$ given by the infimum of the operators $T^{\mathrm{ty},i}, i \in [n]$. It is given by:

$$T_l^{\mathrm{ty}}(x) := \min_{i \in [n]} T_l^{\mathrm{ty},i}(x) \ . \tag{4.46}$$

The following lemma, analogous to Lemma 4.22, gives a metric interpretation of the sub-eigenspace of the operator T.

Lemma 4.43. *Let $\lambda \in [-\infty, 0]$ and $a \in \mathbb{P}(\mathbb{R}_{\max})^n$, we have*

$$T^{\mathrm{ty}}(a) \geqslant \lambda + a \Leftrightarrow \max_{i \in [n]} \operatorname{dist}_H(\mathcal{V}_i, \mathcal{H}_a^i) \leqslant -\lambda \ .$$

Proof. Let $\lambda \in [-\infty, 0]$ and $a \in \mathbb{P}(\mathbb{R}_{\max})^n$. From (4.46) and Lemma 4.36, we deduce the equivalence:

$$\begin{aligned} T^{\mathrm{ty}}(a) \geqslant \lambda + a &\Leftrightarrow \forall i \in [n], T^{\mathrm{ty},i}(x) \geqslant \lambda + a \\ &\Leftrightarrow \forall i \in [n], \operatorname{dist}_H(\mathcal{V}_i, \mathcal{H}_a^i) \leqslant -\lambda \ , \\ &\Leftrightarrow \max_{i \in [n]} \operatorname{dist}_H(\mathcal{V}_i, \mathcal{H}_a^i) \leqslant -\lambda \ . \end{aligned}$$

□

From Lemma 4.43 and Theorem 4.1, we deduce the following result, showing that the tropical linear regression problem with types, associated to the sets $\mathcal{V}_1, \ldots, \mathcal{V}_n$, also reduces to a mean payoff game.

Theorem 4.44. *We have,*

$$\min_{a \in \mathbb{P}(\mathbb{R}_{\max})^n} \max_{i \in [n]} \operatorname{dist}_H(\mathcal{V}_i, \mathcal{H}_a^i) = -\rho(T^{\mathrm{ty}}) \ .$$

Moreover, the minimum is achieved by any vector $a \in \mathbb{P}(\mathbb{R}_{\max})^n$ such that $T^{\mathrm{ty}}(a) \geqslant \rho(T^{\mathrm{ty}}) + a$.

Remark 4.45. Typed tropical linear regression should be compared with the tropical SVM problem introduced in [GJ08]. In the tropical SVM setting, we have a partition of the set of points in n *color classes*, $\mathcal{V}_{c_1}, \dots, \mathcal{V}_{c_n}$, and we are looking for a tropical hyperplane $\mathcal{H}_a$, and for a permutation σ of $\{1, \dots, n\}$ such that for all $i \in [n]$, all the points of color c_i are in the same *sector* $S_{\sigma(i)}(a)$. In other words, we want the tropical hyperplane to separate the n color classes. This is not possible in general, so one needs to consider metric versions, modeling the minimization of classification errors [TWY20]. A possible metric formulation, in the spirit of the present approach, would be to consider

$$\min_{\sigma \in \mathfrak{S}_n} \min_{a \in \mathbb{R}^n} \max_{i \in [n]} \operatorname{dist}_H(\mathcal{V}_i, S_{\sigma(i)}(a)) \qquad \text{(Metric Tropical SVM)} \tag{4.47}$$

where $\mathfrak{S}_n$ denotes the symmetric group on n letters. By comparison with (4.44), we see that we have in addition a minimization over the symmetric group, but the subproblem with a fixed permutation σ arising in the SVM problem is simpler than the analogous problem of typed tropical linear regression, since the sector $S_{\sigma(i)}$ is convex, whereas the set $\mathcal{H}_a^i$ arising in (4.44) is not a convex one. In the application described below, it is the set $\mathcal{H}_a^i$ that is relevant to measure the "distance to equilibrium".

In Figure 4.5(a), we consider the following matrix $V \in \mathbb{R}^{3\times 11}$:

$$V = \begin{pmatrix} 1 & 1 & 2 & 0 & 0 & 0 & -3 & -1 & 0 & 0 & -2 \\ 0 & -2 & 0 & 1 & 1 & 2 & 1 & 0 & 0 & -3 & 0 \\ 0 & 0 & -2 & -2 & -1 & -2 & 0 & 2 & 3 & 1 & 1 \end{pmatrix}, \tag{4.48}$$

and the types are given by the subsets of $\mathcal{V} = [11]$ as follows $\mathcal{V}_1 = \{1, 2, 3, 4\}$, $\mathcal{V}_2 = \{5, 6, 7, 8\}$ and $\mathcal{V}_3 = \{9, 10, 11\}$.

The operators $T^{\mathrm{ty},i} : (\mathbb{R}_{\max})^n \mapsto (\mathbb{R}_{\max})^n$ given by (4.45) and associated to the above matrix V and partition $(\mathcal{V}_i)_{i\in[3]}$ are given by:

$$T^{\mathrm{ty},1}\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} \min[-1 + \max(-2 + x_2, x_3), -2 + \max(x_2, -2 + x_3), \max(1 + x_2, -2 + x_3)] \\ -1 + x_1 \\ 1 + x_1 \end{pmatrix},$$

$$T^{\mathrm{ty},2}\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 1 + x_2 \\ \min[-2 + \max(x_1, -2 + x_3), -1 + \max(-3 + x_1, x_3), \max(-1 + x_1, 2 + x_3)] \\ -2 + x_2 \end{pmatrix},$$

$$T^{\mathrm{ty},3}\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 1 + x_3 \\ 1 + x_3 \\ \min[-3 + \max(x_1, x_2), -1 + \max(x_1, -3 + x_2), -1 + \max(-2 + x_1, x_2)] \end{pmatrix}.$$

Then the operator $T^{\mathrm{ty}} : (\mathbb{R}_{\max})^n \mapsto (\mathbb{R}_{\max})^n$ given by (4.46) is in this example:

$$T^{\mathrm{ty}}\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} \min[-1 + \max(-2 + x_2, x_3), -2 + \max(x_2, -2 + x_3), 1 + x_2, 1 + x_3,] \\ \min[-2 + \max(x_1, -2 + x_3), -1 + \max(-3 + x_1, x_3), -1 + x_1, 1 + x_3] \\ \min[-3 + \max(x_1, x_2), -1 + \max(x_1, -3 + x_2), 1 + x_1, -2 + x_2] \end{pmatrix}.$$

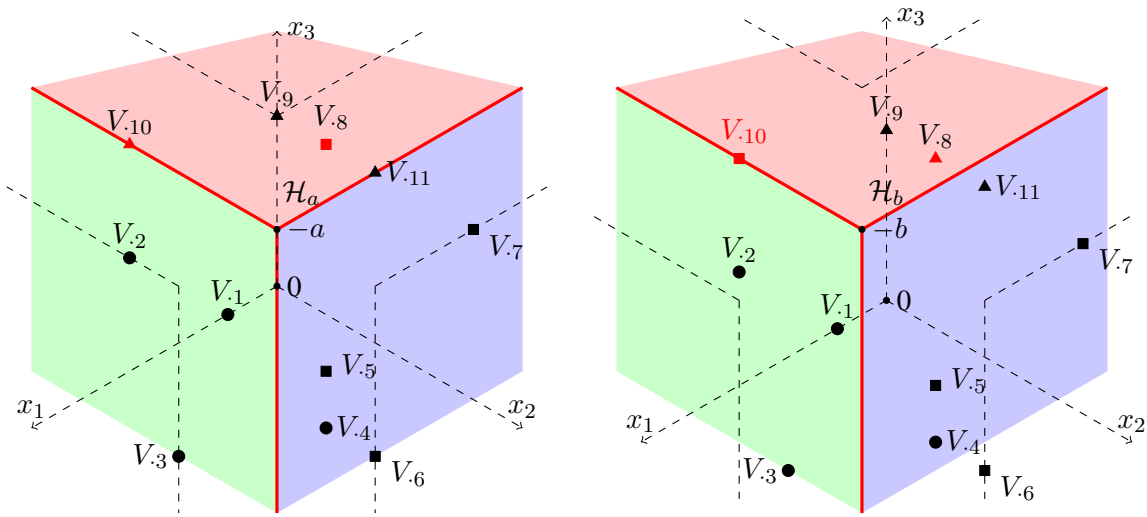
We verify easily that $\lambda = -2$ and $a = (0, 0, -1)^\top$ satisfy $T^{\mathrm{ty}}(a) = \lambda + a$, so that by Theorem 4.44 the apex a is optimal for the typed tropical linear regression problem (4.44).

We notice that in this case, the tropical hyperplane $\mathcal{H}_a$ has at least one witness point in each sector, which means, by Theorem 4.25, that $\mathcal{H}_a$ is also an optimal hyperplane in the sense of the usual tropical linear regression studied in Section 4.4.

Now, if we consider the same matrix V in (4.48), but we exchange the types of the points $V_{\cdot 8}$ and $V_{\cdot 10}$, i.e. we consider the partition $\widetilde{\mathcal{V}}_1 = \{1,2,3,4\}$, $\widetilde{\mathcal{V}}_2 = \{5,6,7,10\}$ and $\widetilde{\mathcal{V}}_3 = \{8,9,11\}$, then the new typed Shapley operator $\widetilde{T^{\mathrm{ty}}}$ is given by:

$$\widetilde{T^{\mathrm{ty}}}\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} \min[-3+x_2, 3+x_3] \\ \min[-2+\max(x_1,-2+x_3), -1+\max(-3+x_1,x_3), -1+x_1, 1+x_3] \\ \min[1+x_1, -4+x_2] \end{pmatrix} .$$

We verify easily that $\mu = -5/2$ and $b = (0, 1/2, -1)^\top$ satisfy $\widetilde{T^{\mathrm{ty}}}(b) = \mu + b$. This example is presented in Figure 4.5(b). Here, we notice that the hyperplane $\mathcal{H}_b$ that is optimal in the typed tropical linear regression sense (Section 4.5.2) does not have witness points in each sector, which means that it is not optimal in the usual tropical linear regression framework (Section 4.4).



(b)

Figure (4.5) Figure 4.5(a): A set of typed points $\mathcal{V}$ with three types in $\mathbb{P}(\mathbb{R}_{\max})^3$ with an optimal tropical hyperplane $\mathcal{H}_a$ in the sense of the typed tropical regression, where $a = (0,0,-1)^\top$ satisfies $T^{\mathrm{ty}}(a) = -2 + a$. Figure 4.5(b): The same set of typed points $\mathcal{V}$ as Figure 4.5(a) but with the types of the two points $V_{\cdot 8}$ and $V_{\cdot 10}$ being exchanged, and an optimal tropical hyperplane $\mathcal{H}_b$ in the sense of the typed tropical regression, where $b = (0, 1/2, -1)^\top$ satisfies $\widetilde{T^{\mathrm{ty}}}(b) = -5/2 + b$.

4.6 Algorithmic aspects

In this section, we explain how the tropical linear regression problem can be effectively solved by using mean-payoff games algorithms. Throughout the section, we assume that the set of points $\mathcal{V}$ is given by as the set of columns the matrix V. By Corollary 4.16, in theory, any algorithm solving mean payoff games in the weakest sense (deciding the inequality $\chi_i(T) \geqslant 0$) can be used. However, some game algorithms lead to more direct approaches, we next discuss some of these.

Considering the strong duality result, Theorem 4.23, and the result on the existence of witness points Theorem 4.25, the key algorithmic issues are:

(i) to compute the upper mean payoff, $\rho(T)$ (which is the opposite of the value of the tropical linear regression problem);

(ii) to decide whether there is a finite eigenvector $u \in \mathbb{R}^n$ such that $T(u) = \rho(T) + u$, and to compute such an eigenvector (when this is so, $-u$ is the center an an optimal ball included in $\mathrm{Sp}(\mathcal{V})$ and the apex of an optimal regression hyperplane, see Remark 4.26);

(iii) to find a sub-eigenvector $b \in (\mathbb{R}_{\max})^n \setminus \{\bot\}$, satisfying $T(b) \geqslant \rho(T) + b$ (then, $\mathcal{H}_b$ is an optimal regression hyperplane);

(iv) to find a super-eigenvector $c \in \mathbb{R}^n$ satisfying $T(c) \leqslant \rho(T) + c$ (then, $-c$ is the center of an optimal ball included in $\mathrm{Sp}(\mathcal{V})$.

For simplicity of the discussion, we assume that T sends $\mathbb{R}^n \to \mathbb{R}^n$. The case in which T sends $\mathbb{R}^n$ to $(\mathbb{R}_{\max})^n$ reduces to this one by considering the action of T on the parts of $(\mathbb{R}_{\max})^n$ and looking for invariant parts.

Then, problems (i)–(iv) are solved, simultaneously, as soon as we know an invariant half-line of T. Indeed, we observed after stating Theorem 4.1 that if (u, η) is an invariant half-line, then $\chi(T) = \eta$. In this way, $\rho(T) = \max_{i\in[n]} \chi_i(T)$ is determined, and this solves issue (i). Moreover, by Proposition 4.3 T admits a finite eigenvector if and only if η is a constant vector, i.e., $\eta = (\lambda, \dots, \lambda)$ for some $\lambda \in \mathbb{R}$, and u is an eigenvector. This solves issue (ii). We observed in the proof of Theorem 4.1 that u satisfies $T(u) \leqslant \rho(T) + u$, and so, this solves issue (iii). Finally, setting $I := \{i \in [n] \mid \chi_i(T) = \rho(T)\}$, and defining the vector $\bar{u}$ such that $\bar{u}_i = u_i$ for $i \in I$ and $\bar{u}_i = -\infty$ otherwise, it can be checked that $T(\bar{u}) \geqslant \bar{u} + \rho(T)$, which solves issue (iv).

More generally, the reduction in the second part of the proof of Corollary 4.16 shows that algorithm which returns an optimal policy σ of Player Min, i.e., a policy such that $\chi(T) = \chi(T^\sigma)$, can be used to produce a finite vector $c \in \mathbb{R}^n$ such that $T(c) \leqslant \overline{\chi}(T) + c$, by reduction to a tropical eigenvalue problem. Moreover, any algorithm which returns an optimal policy τ of Player Max, i.e., a policy such that $\chi(T) = \chi({}^\tau T)$, can be used to produce a vector $b \in (\mathbb{R}_{\max})^n \setminus \{\bot\}$, satisfying $T(b) \geqslant \rho(T) + b$, see the second part of the reduction in Corollary 4.31.

We refer the reader to [Cha09] for a comparative discussion of mean payoff game algorithms. The main known algorithms include the pumping algorithm of [GKK88], value iteration [ZP96], and different algorithms based on the idea of policy iteration [BV07, Sch08, DG06]. In particular, the algorithm of [DG06] returns an invariant half-line. The policy iterations algorithms [BV07, DG06] were reported in [Cha09] to have the best experimental behavior, although policy iteration is are generally exponential [Fri09a].

For the present application to tropical linear regression, we often know in advance that the operator T has a finite eigenvector; this occurs in particular if the entries of the matrix V are finite, and more generally, under the dominion condition of Theorem 4.52. Then, one can use another algorithm, projective Krasnoselkii-Mann value iteration [GS20], which is straightforward to implement and still effective. Starting from a vector $v^0 = (0, \cdots, 0)^\top$, this algorithm computes the following sequence:

$$\tilde{v}^{k+1} = T(v^k) - (\max_{i\in[n]} T(v^k)_i)e, \tag{4.49}$$

$$v^{k+1} = (1-\gamma)v^k + \gamma\tilde{v}^{k+1}. \tag{4.50}$$

where $e = (1, \cdots, 1)^\top \in \mathbb{R}^n$, and $\gamma \in (0, 1)$ is fixed, $1 - \gamma$ being interpreted as a *damping parameter*. In the original Krasnoselskii-Mann algorithm, one writes simply $v^{k+1} = (1-\gamma)v^k + \gamma T(v^k)$. It follows

from [GS20, Coro. 13], based on a general result of Baillon and Bruck [BB92] on the convergence of the original Krasnoselskii-Mann algorithm in normed spaces, see also [CSV14], that v^k does converge to an eigenvector of T as soon as such a (finite) eigenvector u exists. Moreover, $\|T(v^k) - v^k\|_H \leqslant 2\|u\|_H/\sqrt{\pi\gamma(1-\gamma)k}$. In practice, we fix a desired precision $\epsilon > 0$, and stop the computation of the sequence v^k when $\|T(v^k) - v^k\|_H \leqslant \epsilon$.

We now analyze the complexity of the projective Krasnoselskii-Mann algorithm in our special setting. The following observation, shows that, notwithstanding the quadratic size of $S^{\max}$ in the game associated with T (see the discussion after (4.9)), the operator T can be evaluated in linear time.

Proposition 4.46. *The operator T can be evaluated in $O(|E|)$ arithmetic operations.*

Proof. We write $T_i(x) = \min_{k\in[p],(i,k)\in E}(-V_{ik} + y_{ik})$ where $y_{ik} = \max_{j\in[n],j\neq i,(j,k)\in E}(V_{jk} + x_k)$. First, for each column k of the matrix V, we compute the column maximum $M_k := \max_{j\in[n],(j,k)\in E}(V_{jk} + x_k)$ together with an arbitrary index j_k that achieves this maximum, and also the second column maximum, $m_k := \max_{j\in[n],j\neq j_k,(j,k)\in E}(V_{jk} + x_k)$. This preprocessing requires $O(|E|)$ arithmetic operations. We observe that $y_{ik} = m_k$ if $i = j_k$ and $y_{i_k} = M_k$ otherwise. Hence, all the y_{ik} with $(i,k) \in E$ can be computed in $O(|E|)$ arithmetic operations. Finally, the $T_i(x)$ are obtained from the y_{ik} in $O(|E|)$ arithmetic operations. □

We set:

$$W := \max_{v\in\mathcal{V}} \|v\|_H \ .$$

Lemma 4.47. *Suppose that $\mathcal{V}$ is finite, then any finite eigenvector u of T satisfies $\|u\|_H \leqslant W$.*

Proof. By definition of W, we have $v \in B_H(0, W)$ for all $v \in \mathcal{V}$, and since $B_H(0, W)$ is stable by tropical linear combinations, we get $\mathrm{Sp}(\mathcal{V}) \subset B_H(0, W)$. Moreover, by Lemma 4.8, $B_H(u, -\rho(T)) \subset \mathrm{Sp}(\mathcal{V})$. Hence $u \in B_H(0, W)$, meaning that $\|u\|_H \leqslant W$. □

Remark 4.48. There are situations (Section 4.8) in which although some vectors of $\mathcal{V}$ have infinite entries, it is still the case that T has a finite eigenvector. Then, we may still show that there exists a finite eigenvector with not too large entries. To see this, we need to replace W by $W' := \max_{k\in[p]} \delta(V_{\cdot,k})$, where $\delta(V_{\cdot,k}) = \max_{i\in[n],(i,k)\in E} V_{ik} - \min_{j\in[n](j,k)\in E} V_{jk}$. We can always choose such an eigenvector u in such a way that $\|u\|_H = O(nW')$, by appealing to a Blackwell optimality argument, using the proof method of [Sko18, Lemma 8.51] (details are left to the reader). Note that in the special case in which V has finite entries, the bound on $\|u\|_H$ is improved by a factor n.

Corollary 4.49 (Approximate optimality certificate)**.** *Suppose that $\mathcal{V} \subset \mathbb{R}^n$ is of cardinality p. Then, the projective Krasnoselskii-Mann iteration returns in a number of arithmetic operations $O(npW/\epsilon^2)$ a vector $u \in \mathbb{R}^n$ such that $-u$ is both the center of a ball of radius $-\rho(T) - \epsilon$ included in $\mathrm{Sp}(\mathcal{V})$ and the apex of a regression hyperplane, $\mathcal{H}_u$, such that $\mathrm{dist}_H(\mathrm{Sp}(\mathcal{V}), \mathcal{H}_u) \leqslant -\rho(T) + \epsilon$.*

Proof. By [GS20, Coro. 13] and Lemma 4.47, after $k = O(\lceil W/\epsilon^2 \rceil)$ iterations, we end up with a vector $u := v^k$ which satisfies $\|T(u) - u\|_H \leqslant \epsilon$. Moreover, by Proposition 4.46, each iteration requires $O(np)$ arithmetic operations. Setting $\underline{\lambda} := \mathrm{bot}(T(u) - u)$, where $\mathrm{bot}(x) := \min_i x_i$, we deduce that $\underline{\lambda} + u \leqslant T(u) \leqslant \underline{\lambda}(T) + \epsilon + u$, which, by Theorem 4.1, entails that $\rho(T) \leqslant \underline{\lambda}(T) + \epsilon$. Then, by Theorem 4.23, $B(-u, -\rho(T) - \epsilon) \subset \mathrm{Col}(V)$. The proof that $\mathrm{dist}_H(\mathrm{Col}(V), \mathcal{H}_u) \leqslant -\rho(T) + \epsilon$ is dual. □

The following result shows that the factor in $1/\epsilon^2$ can be replaced by $1/\epsilon$ if we look separately for the center of a Hilbert's ball included in $\mathrm{Sp}(\mathcal{V})$ and for the apex of an approximate tropical linear regression hyperplane (in Corollary 4.49, the apex and the center coincide).

Corollary 4.50. *Suppose that $\mathcal{V} \subset \mathbb{R}^n$ is of cardinality p. Then, an ϵ-approximation of the inner radius of $\mathrm{Col}(V)$, as well as vectors $v, z \in \mathbb{R}^n$ satisfying $B_H(v, \text{in-rad}(\mathcal{V}) - \epsilon) \subset \mathcal{V}$ and $\mathrm{dist}_H(\mathcal{V}, \mathcal{H}_z) \leqslant \text{in-rad}(\mathcal{V}) + \epsilon$ can be obtained in $O(npW/\epsilon)$ arithmetic operations.*

Proof of Corollary 4.50. We now rely on the value iteration approach of [AGKS18, Sko18]. The latter computes the sequence given by $v^0 = 0$, $v^k := T(v^{k-1})$, together with the numbers $\bar{\lambda}^k := \max_{i\in[n]} v_i^k$, $\underline{\lambda}^k := \min_{i\in[n]} v_i^k$. The sequence v^k generally does not converge, even up to an additive constant. So, we rely on the following "regularized" sequence [GG04],

$$w^k := \inf(v^0, v^1 - \bar{\lambda}^k/k, \ldots, v^{k-1} - \bar{\lambda}^k(k-1)/k) \ . \tag{4.51}$$

Lemma 8.18 of [Sko18] entails that $\rho(T)$ satisfies $\underline{\lambda}^k/k \leqslant \rho(T) \leqslant \bar{\lambda}^k/k$ with $\bar{\lambda}^k/k - \rho(T) \leqslant \|u\|_H/k$ and $\rho(T) - \underline{\lambda}^k/k \leqslant \|u\|_H/k$, where $u \in \mathbb{R}^n$ is an arbitrary finite eigenvector of T. Hence, it suffices to execute the algorithm up to the iteration $k := \lceil W/\epsilon \rceil$ to make sure that $\bar{\lambda}^k \leqslant \rho(T)+\epsilon$ and $\underline{\lambda}^k \geqslant \rho(T)-\epsilon$. Moreover, Lemma 2 of [GG04] entails that $T(w^k) \leqslant \bar{\lambda}^k + w^k$. Hence, by Lemma 4.8, $-w^k$ is the center of a Hilbert's ball of radius $-\bar{\lambda}^k$ included in $\mathrm{Sp}(\mathcal{V})$. The construction of the apex of an approximate optimal regression hyperplane uses a dual argument, replacing inf by sup in (4.51). □

Remark 4.51. The conclusions of Corollary 4.49 and Corollary 4.50 can be extended to the situation in which some vectors of $\mathcal{V}$ have infinite entries, provided T has a finite eigenvector. Using Remark 4.48, we need to replace W by $W'n$ in the bounds of Corollary 4.49 and Corollary 4.50.

4.7 Illustration: inferring hidden information from equilibria in repeated invitations to tenders

We now illustrate our results on an example from auction theory, in which tropical linear regression allows one to identify secret information from the observation of prices offered in repeated invitations to tenders (ITT).

4.7.1 Auction model with hidden preference factors

We suppose a public decision maker chooses the best offer made by the firms responding to ITT. In accordance with market regulations, see e.g. [cod21, Art. R.2152-7], the best offer is not nessarily the one with the lowest price: other factors, like technical quality, respect of environment, of social impact, can also be taken into account. In the presence of corruption, decisions may be also influenced by bribes.

We assume that this ITT is done repeatedly for a similar service or product each time and in front of the same local firms. We label the firms by $1, 2, \cdots, n$, and we suppose that we have a history of q ITTs with the prices offered by each firm, that are revealed by the decision maker, after having made her choice.

More precisely, we denote the price offered by firm $i \in [n]$ for the ITT number $j \in [q]$ by p_{ij}. We assume that the decision maker has a non public preference factor $f_i > 0$ for each firm i, and that she selects the firm of index i minimizing the expression:

$$\min_{i\in[n]} p_{ij} f_i^{-1} \ . \tag{4.52}$$

In this way, the decision maker considers that for a requested price of p_{ij}, the final cost to be taken into account is $p_{ij} f_i^{-1}$, where $f_i^{-1} \geqslant 1$ is a proportional penalty depending on her estimate f_i of the technical, environmental, or social quality of the firm (the larger f_i, the better its quality).

The same model applies to the situation in which $f_i^{-1} = 1 - \alpha_i\beta$ for some $0 \leqslant \alpha_i \leqslant 1$ and $0 \leqslant \beta < 1$. Now, α_i may be interpreted as a proportional bribe: the firm promises to secretly give back $\alpha_i p_{ij}$ to the decision maker if its offer is accepted, and the parameter β measures how sensitive is the decision maker to bribery ($\beta = 0$ corresponds to a totally honest decision maker, and $\beta = 1^-$ to a totally dishonest one). This is a variant of the classical *first-price sealed-bid auction* [Kri02], incorporating the secret preference.

We suppose that the same firms answer in a recurrent manner to invitations from the same decision maker, and that the factors f_i secretly attached to each firm are kept constant. Then we expect that the prices to be offered to constitute an *equilibrium*, meaning that for each invitation $j \in [q]$, the minimum $\min_{i\in[q]} p_{ij} f_i^{-1}$ is achieved twice at least. Indeed, if the firm i that wins the invitation offers a price p_{ij} such that $p_{ij} f_i^{-1}$ is strictly smaller than $p_{kj} f_k^{-1}$ for all $k \in [n] \setminus \{i\}$, it may offer a higher price and still win the offer, so, in the long run, if an invitation of the same type is made recurrently, the firm will adapt its offer.

This can be modeled in terms of membership to a tropical hyperplane. We put $V_{ij} = -\log(p_{ij})$ and $a_i = \log f_i$, so that the decision maker selects the firm of index achieving the maximum in

$$\max_{i\in[n]}(V_{ij} + a_i) \ . \tag{4.53}$$

Assuming the prices p_{ij} are observed, our goal is to infer the secret information f_i, i.e. the preference factor for firm i, or the bribe offered by this firm.

We first suppose that for each invitation, the identity of the firm that wins the contract is not known to us. We want to infer the hidden information $f = (f_i)_{i\in[n]}$. So, we look for a tropical hyperplane $\mathcal{H}_b$ that is the best regression of the set $\mathcal{V}$ formed by the points $(V_{\cdot j})_{j\in[q]}$ following the analysis of Section 4.4, i..e, we solve a problem of the form (4.17). Following Theorem 4.23, we solve this problem by computing a super-eigenvector $b \in \mathbb{R}^n$ of T, i.e. such that $T(b) \geqslant \rho(T) + b$, where the operator T is given by (4.9).

We note that the decision maker cares only about the *relative* preference factors between the firms, in the sense that if all the preference factors f_i, $i \in [n]$ are multiplied by the same positive constant, the choices of the decision maker will not change. Therefore, we can suppose without loss of generality that $\max_{j\in[n]} f_j = 1$, or equivalently, $\max_{j\in[n]} a_j = 0$.

4.7.2 Numerical instance and experiments

In the following toy example, we take $n = 3$ firms, and a history of $q = 6$ ITTs. We suppose that the decision maker attributes to the firms the preference factors $f = (1, 0.8, 0.6)$, and we take $\forall i \in [3], a_i = \log(f_i)$.

We generated the matrix V_{ij} and the prices $p_{ij} = \exp(-V_{ij})$ by the following structured probabilistic model. We consider six types of products with prices of different order of magnitude. In Table 4.1 the reference prices of these products are $P = (1, 3, 9, 25, 70, 130)$. For each $j \in [6]$, we draw entries $A_{ij}, i \in [3]$ randomly in the interval $K_{ij} = [0.9\times P_j f_i, 1.1\times P_j f_i]$ following a log-uniform law, i.e. equal to the exponential of a variable generated uniformly on the logarithm of the interval K_{ij}. We choose the log-uniform law because it's in adequacy with Benford's law that is observed in real-life price instances. Then, we take $B_{ij} = -\log(A_{ij})$, and we project each column $B_{\cdot j}$ into the tropical hyperplane $\mathcal{H}_a$, to get a vector $C_{\cdot j}$, such that for a given $i \in \arg\max_{k\in[n]}(B_{kj} + a_k)$, we take $C_{ij} = \max_{k\neq i}(B_{kj} + a_k) - a_i$ and we take for all $k \neq i$, $C_{kj} = B_{kj}$. Now the columns $C_{\cdot j}$ belong to the tropical hyperplane $\mathcal{H}_a$. To model the inefficiency of the market, we perturb these columns by taking $V_{ij} = C_{ij} + \delta_{ij}$, with δ_{ij} generated randomly uniformly in $[-\delta, \delta]$, with $\delta = 0.05$. Then the prices are given by $p_{ij} = \exp(-V_{ij})$.

To solve our example, we used the projective Krasnoselskii-Mann iteration described in Section 4.6, with a damping parameter $\gamma = 1/2$. We take $b = v^N$ that gives the approximation of the preference factors by tropical linear regression: $f_i^{\text{reg}} = \exp(b_i)$, $i \in [n]$.

We define the error of the approximation e as the ratio between the Hilbert's distance of the set $\mathcal{V}$ to the hyperplane $\mathcal{H}_b$, which measures the "distance to equilibrium" in this market, and the maximal absolute value of the logarithm of the Hilbert's seminorms of the price vectors $(p_{\cdot j})_{j\in[q]}$:

$$e := \frac{\text{dist}_H(\mathcal{V}, \mathcal{H}_b)}{\max_{j\in[q]} |\log(\|p_{\cdot j}\|_H)|} \ .$$

The following Table 4.1 shows the preference factors f_i, the prices p_{ij} generated with this model and for each invitation we underlined the price of the firm wining that invitation in the sense of achieving the minimum in (4.52). Table 4.1 shows also the prediction f^{reg} of the preference factors that we find by tropical linear regression.

In this example, we set a target accuracy of $\epsilon = 10^{-8}$, and we get that the number of iterations N needed to get $\|T(v^N) - v^N\|_H \leqslant \epsilon$ is $N = 25$. By setting $b = v^N$, we have $\text{dist}_H(\mathcal{V}, \mathcal{H}_b) = 4.21 \times 10^{-2}$ and $\max_{j\in[q]} |\log(\|p_{\cdot j}\|_H)| = 3.84$, and this leads to an error equal to $e = 1.09 \times 10^{-2}$. Figure 4.6 shows the points $(V_{\cdot j})_{j\in[6]}$ in the projective space $\mathbb{P}(\mathbb{R}_{\max})^3$, with the tropical hyperplane $\mathcal{H}_b$ (in blue solid lines) and the points of the space that are at distance equal to $\text{dist}_H(\mathcal{V}, \mathcal{H}_b)$ from $\mathcal{H}_b$ (in blue dashed lines). Figure 4.6 shows in particular the existence of a witness point in each on the three sectors associated to the tropical hyperplane $\mathcal{H}_b$.

	individual houses	social housing	school	road	stadium	bridge	f	f^{reg}
Firm 1	1.02	3.21	<u>8.72</u>	26.2	69.8	<u>123</u>	1	1
Firm 2	0.81	2.65	7.49	20.3	<u>53.8</u>	106	0.8	0.81
Firm 3	<u>0.6</u>	<u>1.86</u>	5.5	<u>14.7</u>	41.8	76	0.6	0.605

Table (4.1) Prices proposed by firms in million euros, the vector of preference factors f and its estimation by tropical linear regression f^{reg} based on the observation of the prices.

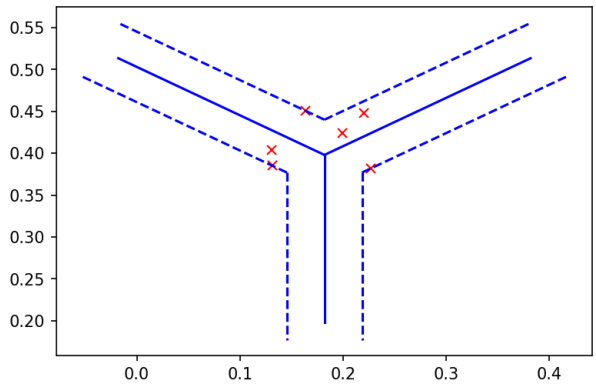


Figure (4.6) The points $(V_{\cdot j})_{j\in[6]}$ in the projective space $\mathbb{P}(\mathbb{R}_{\max})^3$, with the tropical hyperplane $\mathcal{H}_b$ (in blue solid lines) and the points of the space that are at distance equal to $\text{dist}_H(\mathcal{V}, \mathcal{H}_b)$ from $\mathcal{H}_b$ (in blue dashed lines).

Now we consider a similar example still with $n = 3$ firms, but with $q = 100$ invitations to tenders. We use the same generation model, the reference prices $P_j, j \in [100]$, being generated randomly

following a log-uniform law on the interval $[1, 100]$. We set a target accuracy of $\epsilon = 10^{-8}$, and we get that the number of iterations N needed to get $\|T(v^N) - v^N\|_H \leqslant \epsilon$ is $N = 24$. By setting $b = v^N$, we have $\mathrm{dist}_H(\mathcal{V}, \mathcal{H}_b) = 7.69 \times 10^{-2}$ and $\max_{j\in[q]} |\log(\|p_{\cdot j}\|_H)| = 3.72$, and this leads to an error equal to $e = 2.06 \times 10^{-2}$, and the approximation of the preference factors that we obtain is $f^{\mathrm{reg}} = (1, 0.7994, 0.6018)$. Figure 4.7 shows the points $(V_{\cdot j})_{j\in[100]}$ and the approximation hyperplane $\mathcal{H}_b$ obtained in this case with a history of $q = 100$ invitations. We observe also that we have at least a witness point in each sector defined by the tropical hyperplane $\mathcal{H}_b$.

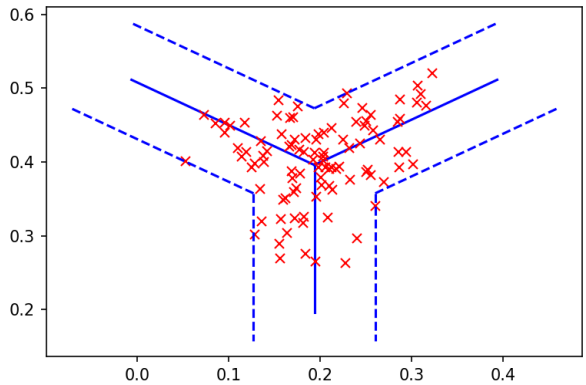


Figure (4.7) The points $(V_{\cdot j})_{j\in[100]}$ in the projective space $\mathbb{P}(\mathbb{R}_{\max})^3$, with the tropical hyperplane $\mathcal{H}_b$ (in blue solid lines) and the points of the space that are at distance equal to $\mathrm{dist}_H(\mathcal{V}, \mathcal{H}_b)$ from $\mathcal{H}_b$ (in blue dashed lines).

4.7.3 Example of regression with types – in which the identities of the winners of the invitations are known

We now suppose the decision maker makes public not only the bid prices that were offered to her, but also the identities of the firms that won the different invitations $j \in [q]$. Then, we can write the set of points $\mathcal{V}$ as a disjoint union $\mathcal{V} = \cup_{\ell\in[n]}\mathcal{V}_\ell$, where $\mathcal{V}_\ell$ is the set of invitations won by firm ℓ. This information can be exploited through the *typed* tropical linear regression of Section 4.5.2. Indeed, if $v = V_{\cdot j} \in \mathcal{V}_\ell$, and if the market is "at equilibria", we know not only that the maximum $\max_{i\in[n]}(V_{ij} + a_i)$ is achieved twice, but that it must be achieved by the firm that won the invitation, i.e., $i = \ell$. Thus, the vector $v \in \mathcal{V}_\ell$ should be close to the *signed* tropical hyperplane $\mathcal{H}_a^\ell$, a finer condition than being close to $\mathcal{H}_a \supset \mathcal{H}_a^\ell$. So, to infer the vector a, we now solve the typed regression problem (4.44), instead of the untyped problem (4.17). Following Theorem 4.44, we are looking for a super-eigenvector b such that $T^{\mathrm{ty}}(b) \geqslant \rho(T^{\mathrm{ty}}) + b$, where the operator T^{ty} is given by (4.46).

We use the same two examples above, and we generate the information of the firm winning each contract $j \in [q]$ by using the information f known by the decision maker. We construct the sets $\mathcal{V}_i$ and the operator T^{ty}, and we find a super-eigenvector of T^{ty} by using the projective Krasnoselkii-Mann value iteration algorithm described in Section 4.6.

After doing the numerical experiments, we find that, the apex b found by typed tropical linear regression, taking advantage of the knowledge of which firm won each invitation, is the same as the one found above by tropical linear regression, for both examples with $q = 6$ and $q = 100$. Hence, here, the additional information provided by the identity of the winners did not help to improve the inference of hidden preferences, by comparison with the basic model in which only the history of the bid prices is used.

4.8 Dominions of the two players and existence of a finite eigenvector

The strongest form of strong duality (Theorem 4.25), with the existence of *witness points*, is valid whenever the Shapley operator T in Equation (4.9) has a finite eigenvector. In this appendix, we provide a sufficient condition for the existence of this eigenvector, which is less demanding than the condition of Proposition 4.28 (requiring V to have only finite entries).

We recall that the operator T represents a game Γ with two players Min and Max, such that when we are at state i, player Min plays first by choosing a column $k \in [p]$ such that $(i, k) \in E$, then player Max chooses a state $j \in [n]$ such that $j \neq i$ and $(j, k) \in E$. Moreover, policies can be defined using (4.10). The game Γ is played repeatedly starting from a given initial position.

We call *dominion* of one player a nonempty subset of states $I \subset [n]$ such that from any initial position in I, that player can force the state to remain in I at each stage of the repeated game, whatever actions the other player chooses. This means that there exist a policy of that player such that for any strategy of the other player, a trajectory of the game starting in I is such that the states visited by Min are all contained in I. The next result, which follows from a more general result (which applies to arbitrary Shapley operators) relates the lack of disjoint dominions of the two players with the existence of a finite eigenvector of a polyhedral Shapley operator.

Theorem 4.52 (Corollary of Thm. 1.2 of [AGH20])**.** *The following assertions are equivalent:*

1. *The two players do not have disjoint dominions in the game Γ;*
2. *For all $r \in \mathbb{R}^n$, the operator $r + T$ has a finite eigenvector.*

Deciding the existence of disjoint dominions for (deterministic) mean payoff games is equivalent to deciding the existence of a non-trivial fixed point of a monotone Boolean function, which is a NP-complete problem, see the discussion in [AGH15a]. However, we next show that for the restricted class of games associated to the Shapley operator T_V, this problem can be solved in polynomial time.

We make the following assumption, which is required for the operator T to send $\mathbb{R}^n$ to $\mathbb{R}^n$, and a fortiori, to have a finite eigenvector.

Assumption G. Each column of the matrix V contains at least two finite entries.

Proposition 4.53. *Suppose that Assumption E and Assumption G hold. Then, the following assertions are equivalent:*

1. *There are disjoint dominions for the two players in the game Γ;*
2. *There exist nonempty subsets I, J of $[n]$, such that $I \cup J = [n]$, $I \cap J = \emptyset$, some columns of V have support included in I, and the other columns of V have at least two finite entries in J.*
3. *There exists a subset K of $[p]$, such that $K \neq \emptyset$ and $K \neq [p]$, and such that if we denote by I_K the union of supports of the columns of V in K, then all the columns not in K have at least two finite entries that are outside I_K. In this case, I_K together with its complement $[n] \setminus I_K$ constitute disjoint dominions of players Min and Max, respectively.*

Proof. We verify first that the assertion (2) and the first part of assertion (3) are equivalent. Indeed, it is straightforward that (3) implies (2) by taking $I = I_K$ and $J = [n] \setminus I_K$. Now, if (2) is true, we take $K = \{k \in [p] \mid \operatorname{supp} V_{\cdot k} \subset I\}$, so $I_K = \cup_{k \in K} \operatorname{supp} V_{\cdot k} \subset I$, and each column $k \notin K$ has at least two finite entries in J, i.e. outside I_K.

Now we suppose that assertion (1) is satisfied, i.e. there are disjoint dominions I and J respectively for Player Min and Player Max. Let us show that this implies (2). The set J is a dominion for Player Max, then there exists a policy τ for Max, such that, for all $i \in J$, for any possible action $(i,k) \in E$ of Player Min, the policy τ sends the state in J, that is $\tau((i,k)) \in J$. Since a policy for Max is a map from E to $[n]$ such that $j = \tau((i,k))$ satisfies $j \neq i$ and $(j,k) \in E$, this implies that, for all $(i,k) \in E$ with $i \in J$, there exists $(j,k) \in E$ with $j \in J$ such that $j \neq i$. Therefore, for all $k \in [p]$, $\operatorname{supp} V_{\cdot k} \cap J$ is either empty or it contains at least two elements. We take $I' = [n] \setminus J \supset I \neq \emptyset$, then the sets I', J satisfy the assertion (2).

Now, we suppose that the first part of assertion (3) is true, and show that I_K and $J = [n] \setminus I_K$ are disjoint dominions of players Min and Max respectively, which will imply (1). Indeed, if $i \in I_K$, then there exists $k \in K$, such that $i \in \operatorname{supp} V_{\cdot k} \subset I_K$. Let us consider a policy σ of Min such that if $i \in I_K$ then $\sigma(i) = (i,k)$ with $k \in K$. Then, if $i \in I_K$, and if Min plays the action $(i,k) = \sigma(i)$, for any possible action of player Max (which exists by Assumption G), that is a choice of $j \in \operatorname{supp} V_{\cdot k}$ such that $j \neq i$, we have $j \in I_K$. This shows that I_K is a dominion of Player Min. Now, let $i \in J$, for any action $(i,k) \in E$ of Min (which exists by Assumption E), we have $k \in K$, since $i \in \operatorname{supp} V_{\cdot k}$ and $i \notin I_K$, so by (3), there exits $j \in \operatorname{supp} V_{\cdot k} \setminus I_K$, with $j \neq i$. So $j \in J$ and j is a possible action of Max when the game is in state (i,k). Considering the policy τ for Max, such that $\tau(i,k) = j$ for i,k,j as before, we get that the set J is a dominion of Player Max. □

From the proof of Proposition 4.53, we deduce in a straightforward manner the following observation, which will be used in Algorithm 9. Note that in the present setting (deterministic mean payoff games), if I, J are disjoint dominions of the two players, then $[n] \setminus J$ and J are also dominions of the two players, hence we shall restrict our search to disjoint dominions that constitute partitions of $[n]$.

Lemma 4.54. *If $D^{\mathrm{Min}}, D^{\mathrm{Max}} \subset [n]$ are disjoint dominions of players Min and Max respectively, that constitute a partition of $[n]$, and K is a subset of columns of V such that the set $S = \cup_{k \in K} \operatorname{supp} V_{\cdot k}$ satisfies $S \subset D^{\mathrm{Min}}$, then for each column $k \notin K$ that has only one finite entry i outside of S, we have $S \cup \{i\} \subset D^{\mathrm{Min}}$.* □

Theorem 4.55. *Algorithm 9, which decides the existence of disjoint dominions in the game Γ associated to a matrix $V \in (\mathbb{R}_{\max})^{n \times p}$, is correct, and it makes $O(n^2 p^2)$ arithmetic operations.*

Proof. The algorithm looks for a set of columns $\bar{K}$ satisfying the last statement of Proposition 4.53. Since the set $\bar{K}$ is required to be nonempty, it suffices for each $k \in [p]$, to verify whether there is such a set $\bar{K} \ni k$ (for loop of the algorithm).

We next show that the algorithm admits the following invariants.

1. At Line 18, S is the union of supports of the columns of K.
2. If there is a subset $\bar{K} \ni k$ satisfying the last statement of Proposition 4.53, with associated then at line 18 of the algorithm, the set K satisfies $K \subset \bar{K}$ and the set S satisfies $S \subset D^{\mathrm{Min}}$.

The first invariant is enforced by lines 10, 12 and 16. We prove that the loop invariant at line 18 holds by induction on the cardinality of S. Let us assume that the condition of the first "if", i.e., $|S_\ell| = 1$ is satisfied. Then, by Lemma 4.54, and by the induction assumption, we must have $S \cup \{i\} \subset D^{\mathrm{Min}}$. Moreover, the last statement of Proposition 4.53 entails that $K \cup \{\ell\} \subset \bar{K}$, and so, the loop invariant is valid in this case. Moreover, if the condition of the second "if", i.e., $|S_\ell| = 0$ is satisfied, then, the second invariant is still valid. This shows that the loop invariant is always valid.

At the exit of the outer while loop, at line 21, we have by construction that every column of V with index outside K has at least two finite entries outside S. Then, by the last statement of Proposition 4.53,

Algorithm 9 Detecting dominions in the game arising from the tropical linear regression problem, for an input matrix $V \in (\mathbb{R}_{\max})^{n \times p}$.

```
for k ∈ [p] do
    K ← {k}
    S ← the support of column k of V
    Declare S to be AUGMENTED (Boolean flag)
    while S is declared as AUGMENTED do
        Declare S not to be AUGMENTED
        Declare all the elements of [p] \ K to be UNSCANNED (Boolean flags)
        while ([p] \ K) contains an UNSCANNED element do
            ℓ ← smallest UNSCANNED element of [p] \ K, declare ℓ to be SCANNED
            S_ℓ ← {i ∈ [n] \ S | V_iℓ is finite}
            if |S_ℓ| = 1 then ▷ column ℓ of V has precisely one finite entry outside S
                K ← K ∪ {ℓ}, S ← S ∪ S_ℓ
                Declare S to be AUGMENTED
            end
            if |S_ℓ| = 0 then ▷ column ℓ of V has no finite entries outside S
                K ← K ∪ {ℓ}
            end
            ▷ S is the union of supports of the columns of K
        done
    done
    if S ≠ [n] then return S and [n] \ S which are disjoint dominions of Players Min and Max respectively
    end
done
There are no disjoint dominions
```

if $S \neq [n]$, S and $[n] \setminus S$ provide disjoint dominions of Players Min and Max, whereas if $S = [n]$, there are no dominions arising from a set $\bar{K} \ni k$. This shows the correctness of the algorithm.

Each iteration of the inner "while" loop makes $O(n)$ arithmetic operations, and every outer "while loop" executes the inner while loop $O(p)$ times. Moreover, the number of outer "while loop" iterations is at most $n-1$. Finally, we have at most p iterations in the "for" loop, which leads to a complexity bound of $O(n^2p^2)$ arithmetic operations for the algorithm. □

We call *Boolean pattern* of the matrix $V \in (\mathbb{R}_{\max})^{n\times p}$ the matrix with entries in $\{0, -\infty\}$, obtained by replacing each finite entry of V by 0. Theorem 4.52 provides a sufficient condition involving the Boolean pattern on V, which guarantees that for all matrices V with this pattern, the operator T admits a finite eigenvector. This condition is not necessary. Consider the following Boolean pattern:

$$\begin{pmatrix} 0_{3,2} & 0_{3,2} \\ (-\infty)_{3,2} & 0_{3,2} \end{pmatrix} , \tag{4.54}$$

where for $\alpha \in \{0, -\infty\}$, $\alpha_{p,q}$ denotes the $p \times q$ matrix with entries identically equal to α.

Proposition 4.56. *If V is a matrix with Boolean pattern* (4.54)*, then, the operator T has a finite eigenvector, but the associated game admits disjoint dominions.*

Proof. First we have that the set $K = \{1, 2\}$ satisfies the condition (3) of Proposition 4.53, and from the proof of Proposition 4.53, we have that the sets $I = \{1, 2, 3\}$ and $J = \{4, 5, 6\}$ are disjoint dominions of Player Min and Player Max respectively.

To show that T has a finite eigenvector, by Proposition 4.3, it suffices to check that $\chi(T) = 0$. The inequality $\chi(T) \leqslant 0$ follows from Remark 4.4. We next show that $\chi(T) \geqslant 0$.

If the game starts from a state $i \in \{4, 5, 6\}$, Player Min must choose the next state to be a pair (i, k) with $k \in \{3, 4\}$, and Player Max can respond by choosing the next state j to belong to $\{4, 5, 6\}$. So, Player Max can force Min to play the same game as the one defined by the submatrix $X := (V_{ij})_{i\in\{4,5,6\},j\in\{3,4\}}$. Since the matrix X consists of only 2 columns of $(\mathbb{R}_{\max})^3$, it follows from Corollary 4.14 that the inner radius of $\mathrm{Col}(X)$ is equal to 0. Then by Theorem 4.1, $\rho(T_X) = 0$, and this entails that Player Max can ensure a payment equal to 0 in the original game, so that $\forall i \in \{4, 5, 6\}, \chi_i(T) = 0$.

Suppose now that the initial state $i \in \{1, 2, 3\}$. Since $\chi(T) = \min_\sigma \chi(T^\sigma)$ where the minimum is taken over the stationary policies of Player Min, it suffices to show that for any such policy, and for $i \in \{1, 2, 3\}$, $\chi_i(T^\sigma) \geqslant 0$. If this policy of Player Min chooses the column 3 or 4, Player Max can again enforce Player Min to play the game associated to the submatrix X, and then Player Max can ensure a payment 0 as before. Now, if the policy of Player Min does not choose the columns 3 and 4, Player Max is forced to play a subgame correspinding to the the submatrix $Y := (V_{ij})_{i\in\{1,2,3\},j\in\{1,2\}}$, and by the same reasoning as before, we know that the value of this game is equal to 0. Then $\forall i \in \{1, 2, 3\}, \chi_i(T) = 0$. □

We leave it as an open question to characterize the Boolean patterns of V which guarantee that the operator T has a finite eigenvector.

CHAPTER 5

Best tropical low-rank approximation of matrices

Abstract

In this chapter, we study the tropical analogue of low-rank approximation of matrices. We establish general properties of this problem, and we identify some tractable subclasses. In particular, we introduce a notion of outer radius of a column space of a matrix, and we show that it is equal to twice the tropical best rank-one approximation error of that matrix. This allows us to provide a strongly polynomial algorithm that gives a rank-one approximation of 3-way tensors. We provide an algebraic interpretation of the outer radius as a skew singular value. Finally, we propose a linear algorithm to compute the best rank-two approximation in dimension three. This chapter is based on a work in progress in collaboration with Marianne Akian, Stéphane Gaubert and Yang Qi.

5.1 Introduction

5.1.1 Motivation and Context

Classical low-rank approximation, allows one to reduce dimensionality and extract a concise linear structure from a given possibly high dimensional data set. It is commonly used in algorithms in data science. A basic tool for dimension reduction and low-rank approximation is Principal Component Analysis (PCA) [Pea01, Hot33], based on the properties of singular value decomposition (also known

as Eckart-Young decomposition) [EY36]. In particular, it provides a best approximation of a given rank, with respect to the Frobenius norm, or statistically speaking in the least squares sense. Another basic low-rank approximation method is based on a selection of subsets of rows and columns (CUR decomposition, also known as skeleton decomposition), see e.g. [GTZ97]. These approaches are useful when the data in question has a predominantly linear structure. However, in the case of intrinsically non-linear systems, approximation becomes more challenging. In this chapter, we aim to give a tropical (max-plus) analogue of low-rank approximation, which provides, with its max-plus structure, non-linear approximation of matrices and more generally of tensors. Like in the classical case, tropical approximation is most effective when the data have a compatible structure — i.e., a predominantly tropically linear structure.

An important case in which such a tropical linear structure arises is the numerical solution of optimal control problems. These problems can be solved using Bellman's Dynamic Programming Principle [Bel52]. The latter shows that the value function is the solution of the so called dynamic programming equation, and it provides an optimal control in feedback form. In the continuous space and time case, the dynamic programming equation takes the form of a partial differential equation called the Hamilton-Jacobi-Bellman (HJB) equation [FR12].

To avoid the *curse of dimensionality* from which suffer the grid based methods in solving HJB equation, *max-plus basis methods* have been developed (see [FM00, McE07, GMQ11, MKH11, Qu14]). The idea of these methods is to exploit the max-plus linearity of the Lax-Oleinik semigroup, which is the evolution semigroup of the HJB equation. In these works, the value function is approximated as a supremum of finitely many basis functions and the supremum is propagated forward in time. Max-plus decomposition methods have allowed to attenuate the curse of dimensionality, for classes of HJB equation [McE07, SMGJ14]. A key issue, in the efficient implementation of max-plus methods, is to select a "basis" of functions with a prescribed cardinality that best approximates a given collection of functions. A version of this is known as the pruning problem [GMQ11, Qu14]. It is equivalent to a problem of low-rank approximation in an infinite dimensional space (of functions). A discrete analogue of the pruning problem has been discussed in [TM19, TTM20].

Other more recent methods allowing to solve the HJB equation are based on tensor decomposition. Dolgov, Kalise and Kunish [DKK21] propose a method based on a classical tensor train approximation for the value function together with a Newton-like iterative method to solve the resulting nonlinear system. Oster, Sallandt and Schneider [OSS19] propose to use low-rank hierarchical tensor product approximation/tree-based tensor formats, together with high-dimensional quadrature, e.g. Monte-Carlo, to solve HJB equation, overcoming computational infeasibility. Recently, deep learning methods were also used to solve HJB equation by trying to find a feedback control law in the form of a neural network [KGNZ19] in the case of deterministic problems. Although, generating data for training the neural network and validating its accuracy remains challenging.

These recent developments motivate the study of the tropical analogue of low-rank approximation for matrices and tensors, and in particular, the discrete analogue of the max-plus basis synthesis problem or pruning problems studied in [MDG08, GMQ11] in the continuous space setting. In particular, we are interested in the fundamental problem of finding the best tropical approximation of matrices and tensors. For example, approximating a tropical matrix A of size $n \times p$ by a tropical product of a matrix B of size $n \times r$ by a matrix C of size $r \times p$, is approximating each entry A_{ij} by the maximum of sums $\max_{1\leqslant k\leqslant r}(B_{ik} + C_{kj})$. In the continuous case, where A is a function of two variables, we look for approximating the entry $A(x, y)$ by a maximum of type $\max_{1\leqslant k\leqslant r}(B_k(x) + C_k(y))$, where B_k and C_k are functions that we want to identify.

Approximating a matrix by a tropical product of two matrices is equivalent to approximately embedding a set of points (the columns of the matrix) by a column space of a matrix, i.e. a tropical

cone with a given number of vertices (see Proposition 5.10). In tropical geometry, many "best approximation" problems have been studied. The simplest one is the tropical projection [CGQ04, AGNS11], which allows to find the nearest point in a (closed) tropical module to a given point of the space, in the sense of Hilbert's metric. The best approximation in the space of ultrametrics, which is a fundamental example of tropical module in view of its application to phylogenetics, has been thoroughly studied [CF00, LSTY17, Ber20]. Another important special case is the best approximation of a point by a tropical linear space [Ard04, JSY07]. In contrast with the approximation of a set of points that we study here, these problems concern the approximation of a *single* point.

Hook proposes in [Hoo17] heuristic algorithms using a local descent method to find low-rank approximation of matrices. In [YZZ19], Yoshida et al. study tropical principal component analysis (PCA) by looking for a tropical polytope with a fixed number of vertices that minimizes the sum of tropical distances (associated to Hilbert's semi-norm) between each data point and its tropical projection into that tropical polytope. They develop a randomized heuristic method to solve this problem with a focus on the special case of tropical polytopes with three vertices. In [PYZ20], Page et al. study also tropical PCA, but with a focus on applying it to dimension reduction over the space of phylogenetic trees. They developed a stochastic optimization method using a Markov Chain Monte Carlo approach to estimate tropical principal components over the space of phylogenetic trees.

In [DSS05], Develin et al. proved that the factor rank of a matrix is at most 2 if and only if all its 3×3 submatrices have a factor rank at most 2. This entails that deciding whether the factor rank does not exceed 2 is polynomial time solvable, and even linear time solvable [Shi12]. In [Shi14], Shitov proves that it is NP-hard to decide whether the factor rank of a matrix is at most 7.

Tropical low-rank factorization is closely related to classical nonnegative factorization (since tropical numbers behave as nonnegative numbers). Moreover, it includes as a special case the problem of factorization for Boolean matrices. Deciding whether a Boolean matrix admits a factorization of a given rank is an NP-hard problem, see e.g. [MN20, Sect. 4.1]. Since the Boolean semiring can be embedded in the tropical semiring, tropical low-rank factorization is also NP-hard. Given the difficulty, it is of interest to identity tractable subclasses, and to develop efficient heuristic algorithms for tropical low-rank approximation.

5.1.2 Results

In this chapter, we establish general properties of tropical low-rank approximation, and identify classes of low-rank approximation problems that are polynomial-time solvable. In particular, we study the tropical low-rank approximation in the case of rank one and rank two. We introduce a notion of outer radius of a column space of a matrix, and we characterize, in Theorem 5.13, the outer radius of a given column space as the eigenvalue of some specific matrix. We show also, in Theorem 5.16, that the tropical best rank-one approximation of a given matrix is equal to half the outer radius of its column space. We provide also a strongly polynomial algorithm that gives a rank-one approximation of 3-way tensors. We provide an algebraic interpretation of the outer radius as a skew singular value. This yields a tropical analogue of a classical result in matrix theory, showing the error in spectral norm of the best rank-one approximation is given by the second singular value. We extend the best tropical rank-one approximation to the case of kernels. In dimension three, we give a linear algorithm that allows to compute the best rank-two approximation, based on signed tropical linear regression studied in Chapter 4.

This chapter is organized as follows. In the preliminaries Section 5.2, we recall the needed results concerning tropical algebra. In Section 5.3, we establish general properties of tropical low-rank approximation and its geometric interpretation. In Section 5.4, we study the outer radius of a column space of a matrix, its equivalence to finding the best rank-one approximation of a matrix, and its algebraic

interpretation. In Section 5.5, we study the best rank-two approximation in dimension three.

5.2 Preliminaries

The tropical (max-plus) semifield $\mathbb{R}_{\max}$ is the set $\mathbb{R} \cup \{-\infty\}$ equipped with the addition $(a,b) \mapsto \max(a,b)$ and the multiplication $(a,b) \mapsto a \odot b := a+b$. We shall occasionally use *the min-plus semifield* $\mathbb{R}_{\min}$, which is the set $\mathbb{R} \cup \{+\infty\}$, equipped with the addition $(a,b) \mapsto \min(a,b)$ and the multiplication $(a,b) \mapsto a+b$.

For any integer n, we denote $[n] := \{1, \cdots, n\}$. For all $x, y \in (\mathbb{R}_{\max})^n$, $A \in (\mathbb{R}_{\max})^{n\times p}$, $B \in (\mathbb{R}_{\max})^{p\times m}$ and $\lambda \in \mathbb{R}_{\max}$, $\lambda + x$ denotes the vector with entries $\lambda + x_i$, for $i \in [n]$, $\lambda + A$ denotes the matrix with entries $\lambda + A_{ij}$, for $i \in [n], j \in [p]$, $x \vee y = \max(x,y)$ denotes the vector with entries $\max(x_i, y_i)$, for $i \in [n]$, and $A \odot_{\max} B \in (\mathbb{R}_{\max})^{nm}$ denotes the matrix with entries $\max_{k\in[p]}(A_{ik} + B_{kj})$, for $i \in [n], j \in [m]$. By abuse of notation, we may denote $\odot_{\max}$ by $\odot$, or denote $A \odot_{\max} B$ by simply AB.

The set $(\mathbb{R}_{\max})^n$ equipped with the addition $(x,y) \mapsto x \vee y$ and the action $(\lambda, x) \mapsto \lambda + x$ of $\mathbb{R}_{\max}$ is a tropical module, i.e. a module over the semifield $\mathbb{R}_{\max}$. We write that $x \leqslant y$ if and only if $\forall i \in [n], x_i \leqslant y_i$. We denote by $A_{i\cdot} \in (\mathbb{R}_{\max})^p$ the i-th row of A and by $A_{\cdot j} \in (\mathbb{R}_{\max})^p$ the j-th column of A. Similarly for $x, y \in (\mathbb{R}_{\min})^n$, $A \in (\mathbb{R}_{\min})^{n\times p}$ and $B \in (\mathbb{R}_{\min})^{p\times m}$, $x \wedge y = \inf(x,y)$ denotes the vector with entries $\min(x_i, y_i)$, for $i \in [n]$, and $A \odot_{\min} B \in (\mathbb{R}_{\min})^{n\times m}$ denotes the matrix with entries $\min_{k\in[p]}(A_{ik} + B_{kj})$, for $i \in [n], j \in [m]$. The set $(\mathbb{R}_{\min})^n$ equipped with the addition $(x,y) \mapsto x \wedge y$ and the action $(\lambda, x) \mapsto \lambda + x$ of $\mathbb{R}_{\min}$ is a tropical module, i.e. a module over the semifield $\mathbb{R}_{\min}$.

A subset $\mathcal{C}$ of $(\mathbb{R}_{\max})^n$ is a *tropical (convex) cone* or equivalently a *tropical submodule* of $(\mathbb{R}_{\max})^n$ if it satisfies $x, y \in \mathcal{C}$ and $\lambda \in \mathbb{R}_{\max}$ implies $\lambda + x \in \mathcal{C}$ and $x \vee y \in \mathcal{C}$. We use a similar definition for tropical cones of $(\mathbb{R}_{\min})^n$ using $\wedge$ instead of $\vee$. For any given matrix $V \in (\mathbb{R}_{\max})^{n\times p}$, we denote by $\mathrm{Col}^{\max}(V) := \{V \odot_{\max} x \mid x \in (\mathbb{R}_{\max})^p\}$ the column space of V, that is the tropical cone of $(\mathbb{R}_{\max})^n$ generated by the columns of V, and we denote by $\mathrm{Row}^{\max}(V)$ the row space of V, that is the tropical cone of $(\mathbb{R}_{\max})^p$ generated by the rows of V. By abuse of notation, we may denote $\mathrm{Col}^{\max}$ by Col and denote $\mathrm{Row}^{\max}$ by Row. Similarly for a matrix $V \in (\mathbb{R}_{\min})^{n\times p}$, we denote by $\mathrm{Col}^{\min}(V) := \{V \odot_{\min} x \mid x \in (\mathbb{R}_{\min})^p\}$ the column space of V, that is the tropical cone of $(\mathbb{R}_{\min})^n$ generated by the columns of V.

We denote by $\perp$ the vector of $(\mathbb{R}_{\max})^n$ identically equal to $-\infty$. We consider the *Hilbert's projective metric*, defined for vectors $x, y \in (\mathbb{R}_{\max})^n$ where at least one of them is not equal to $\perp$, by

$$d_H(x,y) = \inf\{\lambda - \mu \mid \lambda, \mu \in \mathbb{R},\ \mu + y_i \leqslant x_i \leqslant \lambda + y_i\ \forall i \in [n]\} \in \mathbb{R}_{\geqslant 0} \cup \{+\infty\}\ .$$

In addition, we set $d_H(\perp, \perp) := 0$.

The *support* of a vector $x \in (\mathbb{R}_{\max})^n$ is defined as $\operatorname{supp} x := \{i \in [n] \mid x_i \neq -\infty\}$. Observe that $d_H(x,y)$ is finite if and only if x and y have the same support. In this case

$$d_H(x,y) = \max_{i\in I}(x_i - y_i) - \min_{i\in I}(x_i - y_i)\ .$$

We denote the Hilbert's semi-norm of a vector $x \in \mathbb{R}^n$ by

$$\|x\|_H = \max_{i\in[n]} x_i - \min_{i\in[n]} x_i\ . \tag{5.1}$$

Given any semi-norm $\|\cdot\|$ on $\mathbb{R}^n$ (or even any positively homogenous map $\mathbb{R}^n \to \mathbb{R}_+, u \mapsto \|u\|$), we may construct the distance $d_{\|\cdot\|}(u,v) = \|u - v\|$ for all $u, v \in \mathbb{R}^n$, associated to $\|\cdot\|$, and its extension

to $(\mathbb{R}_{\max})^n \times (\mathbb{R}_{\max})^n$ to get a lower semi-continuous map with values in $\mathbb{R}_+ \cup \{+\infty\}$. More precisely, the semi-continuous extension is obtained as follows

$$\forall u, v \in (\mathbb{R}_{\max})^n, \quad d_{\|\cdot\|}(u, v) = \inf_{\mathbb{R}^n \ni \bar{u}_k \to u, \mathbb{R}^n \ni \bar{v}_k \to v} \liminf_{k \to \infty} \|\bar{u}_k - \bar{v}_k\|_\phi \ ,$$

where the infimum is taken over all sequences of vectors $\bar{u}_k \to u$ and $\bar{v}_k \to v$. In this way, the Hilbert's projective metric d_H is obtained from the Hilbert's semi-norm $\|\cdot\|_H$. Similarly, starting from the sup-norm $\|x\|_\infty = \max_{i \in [n]} |x_i|$ for vectors $x \in \mathbb{R}^n$, we obtain the sup-norm distance between two vectors $x, y \in (\mathbb{R}_{\max})^n$, which is $d_\infty(x, y) = \max_{i \in \operatorname{supp} x} |x_i - y_i|$ if x and y have the same support and $d_\infty(x, y) = +\infty$ otherwise.

Given any norm again, we define the distance of a vector $x \in (\mathbb{R}_{\max})^n$ to a set $S \subset (\mathbb{R}_{\max})^n$ as $\operatorname{dist}_{\|\cdot\|}(x, S) := \inf_{a \in S} d_{\|\cdot\|}(x, a)$ and replace $\|\cdot\|$ by ∞ and H if the norm is $\|\cdot\|_\infty$ and $\|\cdot\|_H$ respectively.

We denote by $B(a, r) := \{x \in (\mathbb{R}_{\max})^n \mid d_H(a, x) \leqslant r\}$ the closed ball centered at $a \in \mathbb{R}^n$ with radius r under Hilbert's projective metric.

We denote by $\mathbb{P}(\mathbb{R}_{\max})^n$ the *tropical projective space*, i.e., the quotient of the set of non-identically $-\infty$ vectors of $(\mathbb{R}_{\max})^n$ by the equivalence relation $\sim$ which identifies tropically proportional vectors. We shall abuse notation and denote by the same symbol a vector and its equivalence class.

We shall need some notions and results from tropical spectral theory. The digraph $\Gamma(A)$ of a tropical matrix $A \in (\mathbb{R}_{\max})^{n \times n}$ is defined as the graph with nodes $1, \ldots, n$ and arcs (i, j) whenever $A_{ij} > -\infty$. The weight of the arc (i, j) is A_{ij}, and the weight of a circuit is the sum of the the weights of its arcs. For any integer k, $(A^{\odot k})_{ij}$ is the weight of the maximally weighted path of length k in the graph $\Gamma(A)$ from state i to state j.

When the digraph of A does not have a circuit of positive weight, we define its *Kleene star* by

$$A^* = \mathrm{I} \vee A \vee A^2 \vee \cdots ,$$

where $\mathrm{I} \in (\mathbb{R}_{\max})^{n \times n}$ is the tropical identity matrix with $\forall i \in [n], I_{ii} = 0$ and $\forall i, j \in [n]$, such that $i \neq j$, $I_{ij} = -\infty$, so that A^*_{ij} is the weight of the maximally weighted path from i to j.

We have also the following known result.

Lemma 5.1. *If the digraph of $A \in (\mathbb{R}_{\max})^{n \times n}$ does not have a circuit of positive weight, then*

$$\{x \in (\mathbb{R}_{\max})^n; \quad x \geqslant Ax\} = \operatorname{Col}(A^*) \ .$$

Proof. For any $x \in (\mathbb{R}_{\max})^n$ satisfying $x \geqslant Ax$, we have $x \geqslant A^k x$, which implies $x \geqslant A^* x$. On the other hand, since by definition $A^* \geqslant I$, then $A^* x \geqslant x$, forcing $x = A^* x$. Thus $x \in \operatorname{Col}(A^*)$.

Conversely, if $x \in \operatorname{Col}(A^*)$, $x = A^* y$ for some $y \in (\mathbb{R}_{\max})^n$. Since $AA^* \leqslant A^*$ by the definition of A^*, $Ax = AA^* y \leqslant A^* y = x$, namely $x \geqslant Ax$. □

A scalar μ is a *tropical eigenvalue* of a matrix $A \in (\mathbb{R}_{\max})^{n \times n}$ if there exists a vector $u \in (\mathbb{R}_{\max})^n$, not identically $-\infty$, such that $Au = \mu + u$ in the tropical sense. The eigenvalue is known to be unique when the digraph of A is strongly connected, then it coincides with the maximum weight-to-length ratio of the circuits of the digraph of A. We denote it by $\lambda(A)$. See [BCOQ92, But10] for more information.

5.3 Best tropical rank-r approximation and its geometric interpretation

Let $\phi : \mathbb{R}^n_{\geqslant 0} \mapsto \mathbb{R}_{\geqslant 0}$ be a function that is continuous, order-preserving (i.e. $\forall x, y \in \mathbb{R}^n_{\geqslant 0}$, $x \leqslant y$ implies that $\phi(x) \leqslant \phi(y)$), positively homogeneous of degree 1 (i.e. $\forall \lambda \in \mathbb{R}_{\geqslant 0}, \forall x \in \mathbb{R}^n_{\geqslant 0}$, $\phi(\lambda x) = \lambda \phi(x)$),

and such that $\phi(x) > 0$ for $x \neq 0$. Given a norm $\|\cdot\|$ on $\mathbb{R}^p$, we define the norm map $\|\cdot\|_\phi$ on $\mathbb{R}^{n\times p}$ as follows:

$$\forall A \in \mathbb{R}^{n\times p}, \|A\|_\phi = \phi(\|A_{\cdot 1}\|, \cdots, \|A_{\cdot p}\|) \ .$$

This map is a norm if we assume in addition that ϕ is subadditive, meaning that $\phi(x+y) \leqslant \phi(x)+\phi(y)$.

In this chapter, we are particularly interested in the sup-norm $\|\cdot\|_\infty$ and the Hilbert's semi-norm $\|\cdot\|_H$. In these two cases we denote $\|\cdot\|_\phi$ respectively by $\|\cdot\|_{\phi,\infty}$ and $\|\cdot\|_{\phi,H}$.

The choice of ϕ that will be the most interesting for our analysis is $\phi(x) = \max_{i\in[n]} x_i$. In this case, we replace ϕ by ∞ in the notations $\|\cdot\|_{\phi,\infty}$ and $\|\cdot\|_{\phi,H}$. Thus, $\forall A \in \mathbb{R}^{n\times p}, \|A\|_{\infty,\infty} = \max_{i\in[n],j\in[p]} |A_{ij}|$, and $\|A\|_{\infty,H} = \max_{j\in[p]} \|A_{\cdot j}\|_H$.

We define the distance $d_{\phi,\|\cdot\|}$ on $\mathbb{R}^{n\times p}$ by

$$\forall A, B \in \mathbb{R}^{n\times p}, \quad d_{\phi,\|\cdot\|}(A, B) = \|A - B\|_\phi \ ,$$

and we extend it to $(\mathbb{R}_{\max})^{n\times p}$ to get a lower semi-continuous map (as for distances over $\mathbb{R}^n_{\max}$). We shall use the notation $d_{\phi,\infty}$, $d_{\phi,H}$, $d_{\infty,\infty}$, and $d_{\infty,H}$ for the specializations of this construction.

For a tropical matrix, several notions of rank can be defined. In [DSS05], three natural notions of rank are introduced. In this chapter, we are interested in the *factor rank*, also known as *Schein rank* [Kim82] or *Barvinok rank* [DSS05]. It is defined as follows.

Definition 5.2. The *factor rank* of a matrix $V \in (\mathbb{R}_{\max})^{n\times p}$ is the smallest integer r such that the matrix V is the product of a matrix $A \in (\mathbb{R}_{\max})^{n\times r}$ by a matrix $B \in (\mathbb{R}_{\max})^{r\times p}$, i.e.

$$\forall i \in [n], j \in [p], V_{ij} = \max_{1\leqslant k\leqslant r} (A_{ik} + B_{kj}) \ . \tag{5.2}$$

In the following, we may say rank for the factor rank.

Remark 5.3. The factor rank of a matrix is the specialization of the notion of *tensor rank*, defined for a tensor t (over an arbitrary semiring) as the smallest integer r such that t can be written as the sum of at most r tensors of rank one. Here, the factor rank of V is the smallest integer r such that the matrix can be written as a tropical sum of r rank-one matrices. Indeed, in (5.2), the rank-one matrices are $(A_{ik} + B_{kj})_{i\in[n],j\in[p]} \in (\mathbb{R}_{\max})^{n\times p}$ for each $1 \leqslant k \leqslant r$.

Inspired by the factor rank, we define the best $\|\cdot\|_\phi$ approximation of rank r of a matrix $V \in (\mathbb{R}_{\max})^{n\times p}$ as the following minimization problem

$$\min_{A\in(\mathbb{R}_{\max})^{n\times r}, B\in(\mathbb{R}_{\max})^{r\times p}} d_{\phi,\|\cdot\|}(V, AB) \ . \tag{5.3}$$

We will show that this optimization problem does admit an optimal solution. To this end, we need to recall some properties.

Given a nonempty finite set N, we say that a subset of $\mathbb{R}^N$ is *basic semilinear* if it is of the form $\{x \in \mathbb{R}^N \mid \ell_i(x) \diamond_i b_i, i = 1, \ldots, K\}$ for some $K \geqslant 1$, where for all $i \in [K]$, $\diamond_i \in \{=, <, \leqslant\}$, $b_i \in \mathbb{R}$, and ℓ_i is a linear form with integer coefficients. A subset of $\mathbb{R}^N$ is *semilinear* if it is a finite union of basic semilinear sets. Given $\varnothing \neq I \subset N$, and $S \subset (\mathbb{R} \cup \{-\infty\})^N$, the stratum S_I is defined by $S_I = \{x \in S \mid x_i > -\infty \iff i \in I\}$, and we identify it to a subset of $\mathbb{R}^I$. We say that a subset of $(\mathbb{R}_{\max})^N$ is semilinear if all its strata are semilinear.

We have the following property.

Lemma 5.4. *Let $\varnothing \neq M \subset N$ be finite sets, and let S be a semilinear subset of $\mathbb{R}^N$. Then, the image of S by the projection mapping from $\mathbb{R}^N \to \mathbb{R}^M$ is semilinear. Moreover, if S is closed, then its projection is also closed.*

Proof. The semilinear sets of $\mathbb{R}^N$ where N is a finite nonempty set are precisely the sets that are definable in the first order theory of the divisible ordered group $\mathbb{R}$ and this theory admits quantifier elimination [Mar02, Corollary 3.1.17]. This entails in particular that the projection of a semilinear set S is semilinear. An inspection of the elimination argument shows, in addition, that if S is closed, then its projection is also closed. More precisely, it suffices to consider the case of a closed basic semilinear subset S of $\mathbb{R}^N$, which can be written as $S = \{x \in \mathbb{R}^N \mid Ax \leqslant b\}$ for some $A \in \mathbb{Z}^{p\times N}$ and $b \in \mathbb{R}^N$. Then, by applying the Fourier-Motzkin elimination algorithm (see e.g. [GM06, Sect. 6.7]) to the system of weak inequalities $Ax \leqslant b$, we obtain a finite system of weak inequalities characterizing the projection of S. This shows that the projection of S is a closed (and basic semilinear) set. □

Theorem 5.5. *For all $r \leqslant \min(n,p)$, the set of matrices $R \in (\mathbb{R}_{\max})^{n\times p}$ of rank at most r is semilinear and all its strata are closed.*

Proof. For all $\varnothing \subset I_R \subset [n]\times[p]$, $\varnothing \subset I_A \subset [n]\times[r]$, $\varnothing \subset I_B \subset [r]\times[p]$, let

$$S(I_R, I_A, I_B) := \{(R, A, B) \mid R = AB, R \in (\mathbb{R}_{\max})^{n\times p}, \operatorname{supp} R = I_R, \\ A \in (\mathbb{R}_{\max})^{n\times r}, \operatorname{supp} A = I_A,\ B \in (\mathbb{R}_{\max})^{r\times p} \operatorname{supp} B = I_B\}\ .$$

We note that $S(I_R, I_A, I_B)$ is nonempty iff for all $(i,j) \in I_R$, there exists $k \in [r]$ such that $(i,k) \in I_A$ and $(k,j) \in I_B$. Let Ψ be the set of maps ψ which associate to every $(i,j) \in I_R$ such an element k. Let

$$S^{\psi}(I_R, I_A, I_B) := \{(R, A, B) \in S(I_R, I_A, I_B) \mid \forall (i,j) \in I_R, \max_{1\leqslant k\leqslant r}(A_{ik}+B_{kj}) = A_{i\psi(i,j)}+B_{\psi(i,j)j}\}\ .$$

We have $S(I_R, I_A, I_B) = \cup_{\psi\in\Psi} S^{\psi}(I_R, I_A, I_B)$. Moreover, $S^{\psi}(I_R, I_A, I_B) = \{(R, A, B) \in \mathbb{R}^{I_R} \times \mathbb{R}^{I_A} \times \mathbb{R}^{I_B} \mid \forall (i,j) \in I_R, \forall k \in [r]$ s.t. $(i,k) \in I_A$ and $(k,j) \in I_B, R_{ij} \geqslant A_{ik} + B_{kj}$, and $R_{ij} = A_{i\psi(i,j)} + B_{\psi(i,j)j}\}$, which proves that $S^{\psi}(I_R, I_A, I_B)$ is basic semilinear and closed. So, $S(I_R, I_A, I_B)$ is semilinear and closed. Observe that the set of matrices of rank at most r and with support I_R is the image under the projection mapping $(R, A, B) \mapsto R$ of the union of the sets $S(I_R, I_A, I_B)$ over all admissible values of I_A, I_B. It follows from Lemma 5.4 that this set is semilinear and closed. □

Recall that by singular value decomposition, a best rank-r approximation of a real or complex matrix with respect to the Frobenius norm can be obtained by a succession of best rank-one approximations. On the other hand, owing to the lack of the subtraction operation in the tropical world, we cannot perform such a sequence of best factor rank-one approximations to obtain a best factor rank-r approximation for a tropical matrix, which forces us to consider separately the best approximation problem for all ranks $r \leqslant \min(n,p)$.

We obtain the following corollary showing that Problem (5.3) has an optimal solution.

Corollary 5.6. *Given $V \in (\mathbb{R}_{\max})^{n\times p}$, and $r \leqslant \min(n,p)$, the problem of computing a best $\|\cdot\|_{\phi}$ approximation of rank r of V has an optimal solution.*

Proof. We first observe that ϕ is coercive, meaning that $\phi(x) \to \infty$ when $\|x\| \to \infty$. Indeed, by taking α to be the minimum of $\phi(x)$ over the set $\{x \in \mathbb{R}^n_{\geqslant 0} \mid \|x\|_\infty = 1\}$, noting that the minimum is achieved and is positive, and using the homogeneity of ϕ, we deduce that $\phi(x) \geqslant \alpha\|x\|_\infty$.

We also note that $d_{\phi,\|\cdot\|}(V, W) = +\infty$ if V and W do not have the same support. This follows from the coercive character of ϕ and from the definition of $d_{\phi,\|\cdot\|}$ as a lower semi-continuous extension. Therefore, the best approximation problem (5.3) reduces to minimizing the map $d_{\phi,\|\cdot\|}(V, R)$ over the set S of matrices that lie in the same stratum as V and are of rank at most r. By Theorem 5.5, this set is closed. Hence, since ϕ is coercive, the map to be minimized has compact sublevel sets. It follows that its minimum is achieved. □

Remark 5.7. Theorem 5.5 and Corollary 5.6 carry over (with the same proof method) to the problem of best approximation of a tropical tensor by a tensor of tensor rank at most r. In contrast, over the field $\mathbb{C}$, the set of tensors of rank at most r is not closed, and the best approximation problem may not have an optimal solution.

For two vectors $u, v \in (\mathbb{R}_{\max})^n$, with same support I, we define the center of u and v as $\mathrm{ctr}(u, v) = (\max_{i\in I}(u_i - v_i) + \min_{i\in I}(u_i - v_i))/2$. For a vector $u \in \mathbb{R}^n$, we simply write $\mathrm{ctr}(u)$ for $\mathrm{ctr}(u, 0)$. We denote by $e = (1, \cdots, 1) \in \mathbb{R}^n$ the constant vector equal to 1.

The following lemma relates the Hilbert's semi-norm of a vector to the sup-norm of this vector up to an additive constant, and relates also the Hilbert's distance between two vectors to the sup-norm distance.

Lemma 5.8. *For $u \in \mathbb{R}^n$, we have*

$$\|u\|_H = 2 \inf_{\lambda\in\mathbb{R}} \|u - \lambda e\|_\infty = 2\|u - \mathrm{ctr}(u)e\|_\infty \ . \tag{5.4}$$

For $u, v \in (\mathbb{R}_{\max})^n$, we have

$$d_H(u, v) = 2 \inf_{\lambda\in\mathbb{R}} d_\infty(u, v + \lambda e) = \begin{cases} 2d_\infty(u, v + \mathrm{ctr}(u, v)e) & \textit{if } u \textit{ and } v \textit{ have same support,} \\ +\infty & \textit{otherwise.} \end{cases} \tag{5.5}$$

Proof. For $u \in \mathbb{R}^n$, we have $\inf_{\lambda\in\mathbb{R}} \|u - \lambda e\|_\infty = \inf_{\lambda\in\mathbb{R}} \max_{i\in[n]} |u_i - \lambda|$. We can see easily that the minimum is achieved for $\lambda^* = (\max_{i\in[n]} u_i + \min_{i\in[n]} u_i)/2 = \mathrm{ctr}(u)$ and that the minimum is $\inf_{\lambda\in\mathbb{R}} \|u - \lambda e\|_\infty = \max_{i\in[n]} u_i - \lambda^* = \lambda^* - \min_{i\in[n]} u_i = (\max_{i\in[n]} u_i - \min_{i\in[n]} u_i)/2 = \|u\|_H/2$, which proves (5.4). Similarly, we prove (5.5) for $u, v \in (\mathbb{R}_{\max})^n$; when u and v have the same support, by taking the index i in the support of u instead of $[n]$, and when u and v have different supports we have by definition of the Hilbert's and sup-norm distances that $d_H(u, v) = +\infty$ and $\forall\lambda \in \mathbb{R}, d_\infty(u, v+\lambda e) = +\infty$. □

Using the previous lemma we obtain the following proposition showing the link between the sup-norm distance and the Hilbert's distance of a vector to a tropical cone.

Proposition 5.9. *Let $\mathcal{C} \subset (\mathbb{R}_{\max})^n$ be a tropical cone and $u \in (\mathbb{R}_{\max})^n$. We have*

$$\mathrm{dist}_\infty(u, \mathcal{C}) = \frac{1}{2} \mathrm{dist}_H(u, \mathcal{C}) \ .$$

Moreover, if $\mathrm{dist}_H(u, \mathcal{C}) = d_H(u, v^H) < +\infty$ for some $v^H \in \mathcal{C}$, then $\mathrm{dist}_\infty(u, \mathcal{C}) = d_\infty(u, v^\infty)$ for $v^\infty = v^H + \mathrm{ctr}(u, v^H)e$.

Proof. First we note that $\mathrm{dist}_H(u, \mathcal{C}) = +\infty$ iff $\mathrm{dist}_\infty(u, \mathcal{C}) = +\infty$. In this case no vector of $\mathcal{C}$ has the same support as u. Now, we suppose that $\mathrm{dist}_H(u, \mathcal{C})$ is finite. Using Lemma 5.8, we have $\mathrm{dist}_H(u, \mathcal{C}) = \inf_{v\in\mathcal{C}} d_H(u, v) = 2 \inf_{v\in\mathcal{C}} \inf_{\lambda\in\mathbb{R}} d_\infty(u, v + \lambda e) = 2 \inf_{\lambda\in\mathbb{R}} \inf_{v\in\mathcal{C}} d_\infty(u, v + \lambda e)$. Knowing that $\mathcal{C}$ is a tropical cone, the map $v \mapsto v + \lambda e$ is an isomorphism from $\mathcal{C}$ to itself and then $\mathrm{dist}_H(u, \mathcal{C}) = 2 \inf_{\lambda\in\mathbb{R}} \inf_{w\in\mathcal{C}} d_\infty(u, w) = 2 \inf_{w\in\mathcal{C}} d_\infty(u, w) = 2 \mathrm{dist}_\infty(u, \mathcal{C})$.

If $\mathrm{dist}_H(u, \mathcal{C}) = d_H(u, v^H) < +\infty$, then we have, by Lemma 5.8, $\mathrm{dist}_\infty(u, \mathcal{C}) = \frac{1}{2} \mathrm{dist}_H(u, \mathcal{C}) = \frac{1}{2} d_H(u, v^H) = d_\infty(u, v^H + \mathrm{ctr}(u, v^H)e) = d_\infty(u, v^\infty)$ for $v^\infty = v^H + \mathrm{ctr}(u, v^H)e$. □

The following proposition gives a geometric interpretation to the best rank-r approximation problem for a matrix $V \in (\mathbb{R}_{\max})^{n\times p}$, in terms of the approximation of the columns of V by the column space $\mathrm{Col}(A)$ of some matrix $A \in (\mathbb{R}_{\max})^{n\times r}$.

Proposition 5.10. *Let $V \in (\mathbb{R}_{\max})^{n\times p}$ and $r \leqslant \min(n,p)$. We have,*

$$\min_{A\in(\mathbb{R}_{\max})^{n\times r}, B\in(\mathbb{R}_{\max})^{r\times p}} d_{\phi,\|\cdot\|}(V, AB) = \min_{A\in(\mathbb{R}_{\max})^{n\times r}} \phi(\mathrm{dist}_{\|\cdot\|}(V_{\cdot 1}, \mathrm{Col}(A)), \cdots, \mathrm{dist}_{\|\cdot\|}(V_{\cdot p}, \mathrm{Col}(A))) \ .$$

Moreover, selecting $A \in (\mathbb{R}_{\max})^{n\times r}$ minimizing the right hand side, as well as a matrix $B \in (\mathbb{R}_{\max})^{r\times p}$ such that $\mathrm{dist}_{\|\cdot\|}(V_{\cdot k}, \mathrm{Col}(A)) = d_{\|\cdot\|}(V_{\cdot k}, AB_{\cdot k})$ for all $k \in [p]$, we obtain an optimal solution (A, B) for the above best rank-r approximation of V.

Proof. We have $\min_{A,B} d_{\phi,\|\cdot\|}(V, AB) = \min_A \min_{B_{\cdot 1},\cdots,B_{\cdot p}} \phi(d_{\|\cdot\|}(V_{\cdot 1}, AB_{\cdot 1}), \cdots, d_{\|\cdot\|}(V_{\cdot p}, AB_{\cdot p}))$.

Since the function ϕ is order preserving we have, for A fixed,

$$\min_{B_{\cdot 1},\cdots,B_{\cdot p}} \phi(d_{\|\cdot\|}(V_{\cdot 1}, AB_{\cdot 1}), \cdots, d_{\|\cdot\|}(V_{\cdot p}, AB_{\cdot p})) = \phi(\min_{B_{\cdot 1}} d_{\|\cdot\|}(V_{\cdot 1}, AB_{\cdot 1}), \cdots, \min_{B_{\cdot p}} d_{\|\cdot\|}(V_{\cdot p}, AB_{\cdot p})) \ .$$

Moreover, if $B_{\cdot 1}, \cdots, B_{\cdot p}$ achieve respectively the minimums in the right hand side, then $(B_{\cdot 1}, \cdots, B_{\cdot p})$ achieves the minimum in the left hand side. We have also, for all $k \in [p]$, $\min_{B_{\cdot k}} d_{\|\cdot\|}(V_{\cdot k}, AB_{\cdot k}) = \min_{w\in\mathrm{Col}(A)} d_{\|\cdot\|}(V_{\cdot k}, w) = \mathrm{dist}_{\|\cdot\|}(V_{\cdot k}, \mathrm{Col}(A))$, so that

$$\min_{B_{\cdot 1},\cdots,B_{\cdot p}} \phi(d_{\|\cdot\|}(V_{\cdot 1}, AB_{\cdot 1}), \cdots, d_{\|\cdot\|}(V_{\cdot p}, AB_{\cdot p})) = \phi(\mathrm{dist}_{\|\cdot\|}(V_{\cdot 1}, \mathrm{Col}(A)), \cdots, \mathrm{dist}_{\|\cdot\|}(V_{\cdot p}, \mathrm{Col}(A))).$$

Finally, by taking the minimum over $A \in (\mathbb{R}_{\max})^{n\times r}$, we deduce the desired result. □

By specifying Proposition 5.10 to the case of the Hilbert's semi-norm and the sup-norm, we obtain the following corollary.

Corollary 5.11. *Let $V \in (\mathbb{R}_{\max})^{n\times p}$ and $r \leqslant \min(n,p)$. We have,*

$$\begin{aligned}\min_{A\in(\mathbb{R}_{\max})^{n\times r}, B\in(\mathbb{R}_{\max})^{r\times p}} d_{\phi,H}(V, AB) &= \min_{A\in(\mathbb{R}_{\max})^{n\times r}} \phi(\mathrm{dist}_H(V_{\cdot 1}, \mathrm{Col}(A)), \cdots, \mathrm{dist}_H(V_{\cdot p}, \mathrm{Col}(A))) \\ &= 2 \min_{A\in(\mathbb{R}_{\max})^{n\times r}} \phi(\mathrm{dist}_\infty(V_{\cdot 1}, \mathrm{Col}(A)), \cdots, \mathrm{dist}_\infty(V_{\cdot p}, \mathrm{Col}(A))) \\ &= 2 \min_{A\in(\mathbb{R}_{\max})^{n\times r}, B\in(\mathbb{R}_{\max})^{r\times p}} d_{\phi,\infty}(V, AB) \ .\end{aligned}$$

Proof. We have by Proposition 5.9 that $\forall k \in [p], \mathrm{dist}_H(V_{\cdot k}, \mathrm{Col}(A)) = 2\,\mathrm{dist}_\infty(V_{\cdot k}, \mathrm{Col}(A))$. Combining that with the fact that ϕ is positively homogeneous of degree 1, we obtain the second equality of this proposition. The first and third equalities are special cases of Proposition 5.10. □

The following corollary allows to deduce an optimal solution of the best $\|\cdot\|_{\phi,\infty}$ approximation of rank r from a known optimal solution of the best $\|\cdot\|_{\phi,H}$ approximation of rank r.

Corollary 5.12. *Let $V \in (\mathbb{R}_{\max})^{n\times p}$ and $r \leqslant \min(n,p)$. Let (A^H, B^H) be an optimal solution for $\min_{A\in(\mathbb{R}_{\max})^{n\times r}, B\in(\mathbb{R}_{\max})^{r\times p}} d_{\phi,H}(V, AB)$ constructed as in Proposition 5.10 for the Hilbert seminorm $\|\cdot\|_H$. Then, an optimal solution (A^∞, B^∞) of the problem $\min_{A\in(\mathbb{R}_{\max})^{n\times r}, B\in(\mathbb{R}_{\max})^{r\times p}} d_{\phi,\infty}(V, AB)$ is given by the matrices $A^\infty = A^H$ and B^∞ such that $\forall k \in [p], B^\infty_{\cdot k} = B^H_{\cdot k} + \mathrm{ctr}(V_{\cdot k}, A^H B^H_{\cdot k})e$. This solution is also satisfying the conditions of Proposition 5.10 for the sup-norm.*

Proof. By Proposition 5.10, A^H is optimal in

$$\min_{A\in(\mathbb{R}_{\max})^{n\times r}} \phi(\mathrm{dist}_H(V_{\cdot 1}, \mathrm{Col}(A)), \cdots, \mathrm{dist}_H(V_{\cdot p}, \mathrm{Col}(A))).$$

By using Proposition 5.9 and the positive homogeneity of ϕ, we deduce that $A^H = A^\infty$ is also optimal in $\min_{A\in(\mathbb{R}_{\max})^{n\times r}} \phi(\mathrm{dist}_\infty(V_{\cdot 1}, \mathrm{Col}(A)), \cdots, \mathrm{dist}_\infty(V_{\cdot p}, \mathrm{Col}(A)))$. Since (A^H, B^H) is constructed as in Proposition 5.10, we have $\forall k \in [p]$, $B^H_{\cdot k}$ achieves the minimum in $\min_{B_{\cdot k}} d_H(V_{\cdot k}, A^H B_{\cdot k})$, i.e. that $\mathrm{dist}_H(V_{\cdot k}, \mathrm{Col}(A^H)) = d_H(V_{\cdot k}, A^H B^H_{\cdot k})$. Then, by Proposition 5.9, we have

$$A^H B^H_{\cdot k} + \mathrm{ctr}(V_{\cdot k}, A^H B^H_{\cdot k})e = A^H(B^H_{\cdot k} + \mathrm{ctr}(V_{\cdot k}, A^H B^H_{\cdot k})e) = A^\infty B^\infty_{\cdot k},$$

so that $\mathrm{dist}_\infty(V_{\cdot k}, \mathrm{Col}(A^\infty)) = d_\infty(V_{\cdot k}, A^\infty B^\infty_{\cdot k})$. Therefore, by Proposition 5.10, we deduce that A^∞ and B^∞ are optimal in the problem $\min_{A\in(\mathbb{R}_{\max})^{n\times r}, B\in(\mathbb{R}_{\max})^{r\times p}} d_{\phi,\infty}(V, AB)$. □

5.4 Outer radius of a tropical polytope and best tropical rank-one approximation

5.4.1 Outer radius of a tropical polyhedron

In this section, we study the outer radius of a tropical polyhedron for the Hilbert's semi-norm. We consider a matrix $V \in (\mathbb{R}_{\max})^{n\times p}$ and its column space $\mathrm{Col}(V)$ which is a tropical $\mathbb{R}_{\max}$-submodule of $(\mathbb{R}_{\max})^n$. We are now looking for a Hilbert's ball that contains the columns of V and that has the smallest possible radius. This problem is non trivial only when all the entries of V have the same support. So it suffices to consider the case in which V has finite entries, and in the rest of this section, we assume that $V \in \mathbb{R}^{n\times p}$. Since a Hilbert's ball is stable by the supremum operation, and by translation by constants, this problem is equivalent to finding a Hilbert's ball of minimal radius containing $\mathrm{Col}_*(V) := \mathrm{Col}(V) \cap \mathbb{R}^n$, the set of finite vectors of the tropical polyhedral cone $\mathrm{Col}(V)$ (observe that when V has finite entries, $\mathrm{Col}(V) = \mathrm{Col}_*(V) \cup \{\bot\}$ where $\bot$ is the vector with entries $-\infty$). We will show in the following that the outer radius of $\mathrm{Col}_*(V)$ coincides with the outer radius of $\mathrm{Row}_*(V) := \mathrm{Row}(V) \cap \mathbb{R}^n$, that is the set of finite vectors of the row space $\mathrm{Row}(V)$. We call *outer ball* of the tropical polyhedron $\mathrm{Col}_*(V)$ any Hilbert's ball containing $\mathrm{Col}_*(V)$ and having the smallest possible radius, we call this radius the *outer radius* of $\mathrm{Col}_*(V)$, and we denote it by $\mathrm{outerradius}(\mathrm{Col}_* V)$.

Finding the outer radius of $\mathrm{Col}_*(V)$ can be expressed as a minimization problem:

$$\inf\{r \in \mathbb{R}_{\geqslant 0} \mid \exists u \in \mathbb{R}^n,\ \forall j \in [p],\ d_H(V_{\cdot j}, u) \leqslant r\} \ . \tag{5.6}$$

We consider the square matrix $H = V \odot_{\max} (-V^\top) \in \mathbb{R}^{n\times n}$, where

$$H_{ik} = \max_{j\in[p]}(V_{ij} - V_{kj}), \quad i, k \in [n] \ . \tag{5.7}$$

The following theorem characterizes the outer radius of the polyhedron $\mathrm{Col}_*(V)$.

Theorem 5.13. *The outer radius of* $\mathrm{Col}_*(V)$ *is the eigenvalue* λ *of the matrix* $H = V \odot_{\max} (-V^\top)$*, and the set of centers of all Hilbert outer balls of* $\mathrm{Col}_*(V)$ *is the column space* $\mathrm{Col}_*((-\lambda + H)^*)$.

Proof of Theorem 5.13. The condition $\forall j \in [p],\ d_H(V_{\cdot j}, u) \leqslant r$ is equivalent to $\forall j \in [p], \forall i \in [n], \forall k \in [n], (V_{ij} - u_i) - (V_{kj} - u_k) \leqslant r$, and also to $\forall i \in [n], \forall k \in [n], \max_{j\in[p]}(V_{ij} - V_{kj}) + u_k \leqslant r + u_i$, i.e. $\forall i \in [n], \forall k \in [n], H_{ik} + u_k \leqslant r + u_i$. This is equivalent to $\forall i \in [n], \max_{k\in[n]}(H_{ik} + u_k) \leqslant r + u_i$, and also to $Hu \leqslant r + u$. Therefore the outer radius of $\mathrm{Col}_*(V)$ is $\min\{r \mid \exists u \in \mathbb{R}^n,\ Hu \leqslant r + u\}$, which, by the Collatz-Wielandt property (Theorem 4.1),is exactly the tropical eigenvalue λ of the matrix H.

Let $u \in \mathbb{R}^n$, $\lambda \in \mathbb{R}_{\geqslant 0}$, the Hilbert's ball $B(u, \lambda)$ is an outer ball of $\mathrm{Col}_*(V)$ if and only $Hu \leqslant \lambda + u$, i.e. $(-\lambda + H)u \leqslant u$. Therefore, by Lemma 5.1, the set of possible centers is $\mathrm{Col}_*((-\lambda + H)^*)$. □

Given a tropical submodule $\mathcal{V}$ of $(\mathbb{R}_{\max})^n$, let $\mathrm{Lat}(\mathcal{V})$ denote the sublattice of $(\mathbb{R}_{\max})^n$ generated by $\mathcal{V}$, which coincides with the set of infima of finite families of elements of $\mathcal{V}$. We observe that a Hilbert's ball is a sublattice of $(\mathbb{R}_{\max})^n$. It follows that if a Hilbert's ball contains $\mathrm{Col}_*^{\max}(V)$, it contains also $\mathrm{Lat}(\mathrm{Col}_*^{\max} V)$. Hence, the following result, which shows in particular that the columns of the matrix H are generators of $\mathrm{Lat}(\mathrm{Col}_*^{\max} V)$, explains the role of the matrix H in Theorem 5.13.

Proposition 5.14. *Let $V \in \mathbb{R}^{n\times p}$. We have*

$$\mathrm{Col}_*^{\max}(V \odot_{\min} (-V^\top)) = \mathrm{Lat}(\mathrm{Col}_*^{\max} V) = \mathrm{Lat}(\mathrm{Col}_*^{\min} V) = \mathrm{Col}_*^{\min}(V \odot_{\max} (-V^\top)) \ .$$

Proof. We first observe that the matrix $X := V \odot_{\min} (-V^\top)$ satisfies the following properties

$$X \odot_{\max} V = V, \qquad X \odot_{\max} X = X, \qquad X \geqslant I_{\max} \ ,$$

where $I_{\max}$ is the tropical identity matrix (with 0 on the diagonal and $-\infty$ elsewhere). These properties originate from residuation theory, see [CGQ04, Sect. 2]. They can also be checked in a straightforward manner from the definition of $\odot_{\max}$ and $\odot_{\min}$.

It follows from $X \odot_{\max} V = V$ that $\mathrm{Col}_*^{\max}(V) \subset \mathrm{Col}_*^{\max}(X)$. We have, for all $u, v \in \mathbb{R}^n$,

$$\begin{aligned}(X \odot_{\max} u) \wedge (X \odot_{\max} v) &= (X \odot_{\max} X \odot_{\max} u) \wedge (X \odot_{\max} X \odot_{\max} v)\\ &\geqslant X \odot_{\max} ((X \odot_{\max} u) \wedge (X \odot_{\max} v))\\ &\qquad\qquad \text{(because } w \mapsto X \odot_{\max} w \text{ is order preserving)},\\ &\geqslant (X \odot_{\max} u) \wedge (X \odot_{\max} v) \qquad \text{(because } X \geqslant I_{\max}).\end{aligned}$$

This implies that $(X \odot_{\max} u) \wedge (X \odot_{\max} v) = X \odot_{\max} ((X \odot_{\max} u) \wedge (X \odot_{\max} v)) \in \mathrm{Col}_*^{\max}(X)$. Hence, $\mathrm{Col}_*^{\max}(X)$ is a sublattice of $(\mathbb{R}_{\max})^n$ containing $\mathrm{Col}_*^{\max}(V)$, and so, $\mathrm{Col}_*^{\max}(X) \supset \mathrm{Lat}(\mathrm{Col}_*^{\max}(V))$.

We also observe that

$$X_{ij} = \inf_k (V_{ik} - V_{jk})$$

and so, the column $X_{\cdot j}$ is the infimum of the vectors $V_{\cdot k} - V_{jk}$, which belong to $\mathrm{Col}_*^{\max}(V)$, showing that $\mathrm{Col}_*^{\max}(X) \subset \mathrm{Lat}(\mathrm{Col}_*^{\max}(V))$.

The proof that $\mathrm{Lat}(\mathrm{Col}^{\min} V) = \mathrm{Col}^{\min}(V \odot_{\max} (-V^\top))$ follows by a dual argument, notting that the matrix $H := V \odot_{\max} (-V^\top)$ satisfies $H \odot_{\min} V = V$, $H \odot_{\min} H = H$, and that $H \leqslant I_{\min}$, where $I_{\min}$ denotes the dual tropical identity matrix (with 0 on the diagonal and $+\infty$ elsewhere). Finally, $\mathrm{Lat}(\mathrm{Col}_*^{\max} V) = \mathrm{Lat}(\mathrm{Col}_*^{\min} V)$ follows from the distributivity property for infima and suprema. □

5.4.2 Best tropical rank-one approximation of matrices and tensors

In this section, we show that computing the outer radius of the tropical cone $\mathrm{Col}_*(V)$ for the Hilbert's semi-norm is equivalent to finding the best tropical rank-one approximation of the matrix V for the semi-norm $\|\cdot\|_{\infty,H}$ or the norm $\|\cdot\|_{\infty,\infty}$. We will use this analysis to give a 2-approximation rank-one decomposition for 3-way tensors of the form $T \in \mathbb{R}^{n\times m\times p}$.

The following proposition links the rank-one approximation error of the matrix V to the radii of Hilbert's balls containing the column and row spaces of V.

Proposition 5.15. *For a matrix $V \in \mathbb{R}^{n\times p}$ and $r \geqslant 0$, the following properties are equivalent:*

1. *There exists $u \in \mathbb{R}^n$ such that $\mathrm{Col}_*(V) \subset B(u, r)$.*
2. *There exists $v \in \mathbb{R}^p$ such that $\mathrm{Row}_*(V) \subset B(v, r)$.*

3. *There exists* $u \in \mathbb{R}^n, v \in \mathbb{R}^p$ *such that* $\|V - u \odot v\|_{\infty,H} \leqslant r$.

4. *There exists* $u \in \mathbb{R}^n, v \in \mathbb{R}^p$ *such that* $\|V - u \odot v\|_{\infty,\infty} \leqslant r/2$.

Proof. If we have assertion 1, then $\forall j \in [p],\ d_H(V_{\cdot j}, u) \leqslant r$. Then, taking the vector v identically equal to 0, we get $\|V - u \odot v\|_{\infty,H} \leqslant r$, which shows Assertion 3.

Conversely, if we have Assertion 3, then for all $j \in [p]$, $d_H(V_{\cdot j}, u) \leqslant r$, that is $V_{\cdot j} \in B(u, r)$, and since the Hilbert's ball $B(u, r)$ is stable by max plus combinations, we get that $\mathrm{Col}_*(V) \subset B(u, r)$.

Applying Corollary 5.11 and Corollary 5.12 to $r = 1$, and remarking that when V has finite entries, any vectors $u \in \mathbb{R}^n_{\max}$ and $v \in \mathbb{R}^p_{\max}$ minimizing $d_{\infty,\infty}(V, u \odot v)$ or $d_{\infty,H}(V, u \odot v)$ have finite entries, we get the equivalence between Assertions 3 and 4.

Finally, using the symmetry of the $\|\cdot\|_{\infty,\infty}$ norm for rows and columns, we deduce the equivalence between the assertions 2 and 4. □

From Proposition 5.15, we deduce readily the following theorem showing that the best rank-one approximation error for a tropical matrix V is given by half the outer radii of its column and row spaces.

Theorem 5.16. *For any matrix* $V \in \mathbb{R}^{n\times p}$*, we have*

$$\begin{aligned}\min_{A\in\mathbb{R}^{n\times p},\mathrm{rank}\,A=1} \|V - A\|_{\infty,\infty} &= \frac{1}{2} \min_{A\in\mathbb{R}^{n\times p},\mathrm{rank}\,A=1} \|V - A\|_{\infty,H} \\ &= \frac{1}{2}\,\mathrm{outerradius}(\mathrm{Col}_*\,V) = \frac{1}{2}\,\mathrm{outerradius}(\mathrm{Row}_*\,V)\ .\end{aligned}$$

From Proposition 5.15 and Theorem 5.13, we deduce the following practical corollary.

Corollary 5.17. *Let* $V \in \mathbb{R}^{n\times p}$*. If* $u \in \mathbb{R}^n$ *is an eigenvector of* $V \odot (-V^\top)$ *and* $v \in \mathbb{R}^p$ *is such that* $\forall j \in [p], v_j = \mathrm{ctr}(V_{\cdot j}, u) = \frac{1}{2}(\max_{k\in[n]}(V_{kj} - u_k) + \min_{k\in[n]}(V_{kj} - u_k))$*, then*

$$\|V - u \odot v\|_{\infty,\infty} = \frac{1}{2}\lambda(V \odot_{\max} (-V^\top)) = \min_{A\in\mathbb{R}^{n\times p},\mathrm{rank}\,A=1} \|V - A\|_{\infty,\infty}. \tag{5.8}$$

Proof. By Theorem 5.13, if $u \in \mathbb{R}^n$ is an eigenvector of $V \odot (-V^\top)$, and λ is its eigenvalue, then $\mathrm{Col}_*(V) \subset B(u, r)$ with $r = \lambda$, and r is minimal. By Proposition 5.15, we get that there exists $w \in \mathbb{R}^p$ such that $\|V - u \odot w\|_{\infty,H} \leqslant r$, and since r is minimal, we get the equality. The value of $\|V - u \odot w\|_{\infty,H}$ is independent of the (finite) vector w. Applying Corollary 5.12 to $A^H = u$ and $B^H = w$ identically equal to 0, we obtain that the couple (u, v) with $v = B^\infty$ is minimizing $\|V - u \odot w\|_{\infty,\infty}$ and that the minimum is $r/2$. Using the expression of B^∞ for $A^H = u$ and $B^H = w$ identically 0, we get that the entries of v are equal to $v_j = \mathrm{ctr}(V_{\cdot j}, u)$. □

Corollary 5.18. *An exact solution of the best tropical rank-one approximation problem for the matrix* $V \in \mathbb{R}^{n\times p}$ *can be obtained in* $O(n^2(n + p))$ *arithmetic operations.*

Proof. Corollary 5.17 shows that computing a best tropical rank-one approximation reduces to computing an eigenvector of the matrix H given by (5.7). First, the computation of the matrix H takes $O(n^2p)$ arithmetic operations. Then, computing an eigenvector of this matrix can be done in $O(n^3)$ arithmetic operations; indeed, the eigenvalue can be obtained in $O(n^3)$ time using Karp's algorithm [Kar78], and then, an eigenvector can be derived, still in $O(n^3)$ time, by solving a shortest path problem, see e.g. [BCOQ92, Th. 3.101]. □

We consider a 3-way tropical tensor $T \in \mathbb{R}^{n\times m\times p}$. We can look to it as a collection of vectors $T_{\cdot jk} \in \mathbb{R}^n$ with $j \in [m], k \in [p]$, as a collection of vectors $T_{i\cdot k} \in \mathbb{R}^m$, $i \in [n], k \in [p]$ or also as a collection of vectors $T_{ij\cdot} \in \mathbb{R}^p$, $i \in [n], j \in [m]$. For a tensor T, we shall denote by $\|T\|_\infty$ the sup-norm $\max_{i\in[n],j\in[m],k\in[p]} |T_{ijk}|$.

As in (5.7), we define the following matrices associated to the tensor T:

$$H^{(1)}_{il} = \max_{j\in[m],k\in[p]} (T_{ijk} - T_{ljk}), \quad i, l \in [n],$$
$$H^{(2)}_{jq} = \max_{i\in[n],k\in[p]} (T_{ijk} - T_{iqk}), \quad j, q \in [m],$$
$$H^{(3)}_{ks} = \max_{i\in[n],j\in[m]} (T_{ijk} - T_{ijs}), \quad k, s \in [p] \ .$$

The following corollary shows an equivalence between Hilbert's balls containing the column space $\mathrm{Col}^{\max}_*((T_{\cdot jk})_{jk})$, generated by the vectors $(T_{\cdot jk})_{j\in[m],k\in[p]}$, and the approximation of the tropical tensor T by a tropical product of type $u \odot A$, with $u \in \mathbb{R}^n$ and $A \in \mathbb{R}^{m\times p}$. This corollary follows from Theorem 5.13 and Proposition 5.15. The proof of the equivalence is similar to the one of Proposition 5.15, by taking here

$$A_{jk} = \frac{1}{2}(\max_{i\in[n]}(T_{ijk} - u_i) + \min_{i\in[n]}(T_{ijk} - u_i)) \ . \tag{5.9}$$

Corollary 5.19. *For a tensor $T \in \mathbb{R}^{n\times m\times p}$ and $r \geqslant 0$, the following properties are equivalent:*

1. *There exists $u \in \mathbb{R}^n$ such that $\mathrm{Col}^{\max}_*((T_{\cdot jk})_{jk}) \subset B(u, r)$.*
2. *There exists $u \in \mathbb{R}^n, A \in \mathbb{R}^{m\times p}$ such that $\|T - u \odot A\|_\infty = \max_{i\in[n],j\in[m],k\in[p]} |T_{ijk} - (u_i + A_{jk})| \leqslant r/2$.*

Moreover, we have $\min_{u\in\mathbb{R}^n,A\in\mathbb{R}^{m\times p}} \|T - u \odot A\|_\infty = \frac{1}{2}\,\mathrm{outerradius}(\mathrm{Col}^{\max}_*((T_{\cdot jk})_{jk})) = \frac{1}{2}\lambda(H^{(1)})$.

Remark 5.20. We obtain a similar result to Corollary 5.19 when we consider the tensor T as the collection of vectors $(T_{i\cdot k})_{i\in[n],k\in[p]}$ or $(T_{ij\cdot})_{i\in[n],j\in[m]}$. Moreover we deduce that

$$\frac{1}{2}\lambda(H^{(1)}) = \min_{u\in\mathbb{R}^n,A\in\mathbb{R}^{m\times p}} \|T - u \odot A\|_\infty \leqslant \min_{u\in\mathbb{R}^n,v\in\mathbb{R}^m,w\in\mathbb{R}^p} \|T - u \odot v \odot w\|_\infty =: e^*,$$

where e^* is the best rank-one approximation of the tensor T. By applying the same reasoning to $H^{(2)}$ and $H^{(3)}$, we deduce that $\frac{1}{2}\max(\lambda(H^{(1)}), \lambda(H^{(2)}), \lambda(H^{(3)})) \leqslant e^*$.

The following proposition gives a method to find a rank-one approximation of the tensor T that has an error smaller than twice the best rank-one error e^*.

Proposition 5.21. *Let $u \in \mathbb{R}^n$ be an eigenvector of $H^{(1)}$, and take $A \in \mathbb{R}^{m\times p}$ as in* (5.9)*. Let $v \in \mathbb{R}^m$ be an eigenvector of the matrix $H^A = A \odot (-A^\top)$, and take $w \in \mathbb{R}^p$, defined by $\forall k \in [p], w_k := \frac{1}{2}(\max_{j\in[m]}(A_{jk} - v_j) + \min_{j\in[m]}(A_{jk} - v_j))$. We have*

$$\|T - u \odot v \odot w\|_\infty \leqslant \frac{1}{2}(\lambda(H^{(1)}) + \lambda(H^A)) \leqslant 2e^* \ . \tag{5.10}$$

Proof. From Corollary 5.19, we have $\|T - u \odot A\|_\infty \leqslant \frac{1}{2}\lambda(H^{(1)})$, and from Corollary 5.17, we have $\|A - v \odot w\|_\infty \leqslant \frac{1}{2}\lambda(H^A)$. Therefore, $\|T - u \odot v \odot w\|_\infty \leqslant \|T - u \odot A\|_\infty + \|A - v \odot w\|_\infty \leqslant \frac{1}{2}(\lambda(H^{(1)}) + \lambda(H^A))$.

Now, we will prove that $H^A \leqslant H^{(2)}$. Indeed, by definition, we have $\forall j, q \in [m], H^A_{jq} = \max_{k\in[p]}(A_{jk} - A_{qk})$, and $A_{jk} - A_{qk} = \frac{1}{2}[\max_{i\in[n]}(T_{ijk} - u_i) - \max_{i\in[n]}(T_{iqk} - u_i) + \min_{i\in[n]}(T_{ijk} - u_i) - \min_{i\in[n]}(T_{iqk} - u_i)]$. Using that for any $\forall x, y \in \mathbb{R}^n, \max_{i\in[n]} x_i - \max_{i\in[n]} y_i \leqslant \max_{i\in[n]}(x_i - y_i)$ and $\min_{i\in[n]} x_i - \min_{i\in[n]} y_i \leqslant \max_{i\in[n]}(x_i - y_i)$, we get easily that $A_{jk} - A_{qk} \leqslant \max_{i\in[n]}(T_{ijk} - T_{iqk})$. Therefore, $\forall j, q \in [m], H^A_{jq} = \max_{k\in[p]}(A_{jk} - A_{qk}) \leqslant \max_{i\in[n],k\in[p]}(T_{ijk} - T_{iqk}) = H^{(2)}_{jq}$.

Then, we deduce that $\lambda(H^A) \leqslant \lambda(H^{(2)})$, and so by Remark 5.20, we obtain that $\|T - u\odot v\odot w\|_\infty \leqslant \frac{1}{2}(\lambda(H^{(1)}) + \lambda(H^{(2)})) \leqslant 2e^*$. □

Remark 5.22. The problem of finding the best sup-norm approximation of a tensor $T \in \mathbb{R}^{n\times m\times p}$ by a rank-one tensor $u \odot v \odot w$ can be formulated as a linear program, namely, to minimizing a real parameter s under the constraints $-s \leqslant T_{ijk} - u_i - v_j - w_k \leqslant s$, for $i \in [n], j \in [m], k \in [p]$. This program involves $2nmp$ inequalities and $n+m+p+1$ variables. Standard interior point methods make $O(L(nmp)^{3/2}(n+m+p)^2)$ arithmetic operations, where L is the total bit size of T (see [Vai90]).

By comparison, the bound of (5.10) can be computed in a number of arithmetic operations $O(n^3 + m^3 + m^2p + n^2mp)$.

5.4.3 Algebraic interpretation of the outer radius as a skew singular value

Owing to Theorem 5.16, we are able to give an algebraic interpretation of the outer radius. Given a tropical matrix $V \in \mathbb{R}^{n\times p}$, if there are vectors $u \in (\mathbb{R}_{\max})^n \setminus \{\perp\}, v \in (\mathbb{R}_{\max})^p \setminus \{\perp\}$ and a real scalar λ such that $Vv = \lambda u$ and $u^\top(-V) = \lambda v^\top$ in the tropical sense, namely

$$\begin{cases} \max\limits_{j\in[p]}(V_{ij} + v_j) & = \lambda + u_i \\ \max\limits_{j\in[n]}(-V_{jk} + u_j) & = \lambda + v_k \end{cases} \tag{5.11}$$

for all $i \in [n]$ and $k \in [p]$, then λ is called a *tropical skew singular value* of V, u is called a *tropical left skew singular vectors* of V associated with λ, and v is called a *tropical right skew singular vectors* of V associated with λ. Besides, we will call (λ, u, v) a *tropical skew singular tuple* of V. In fact, such a skew singular tuple (λ, u, v) satisfies

$$\begin{pmatrix} \perp & V \\ -V^\top & \perp \end{pmatrix} \begin{pmatrix} u \\ v \end{pmatrix} = \lambda \begin{pmatrix} u \\ v \end{pmatrix}$$

in the tropical sense. Recall that the tropical spectral theorem shows that a matrix with a strongly connected digraph has a unique eigenvalue, and that all associated eigenvectors are finite. Since V has finite entries, the digraph of $\left(\begin{smallmatrix} \perp & V \\ -V^\top & \perp \end{smallmatrix}\right)$ is strongly connected, then it follows the tropical skew singular value is unique, and also that the skew singular vectors are finite. This should be compared with the classical characterization of the singular values of a complex matrix W as the nonnegative eigenvalues of the matrix $\left(\begin{smallmatrix} 0 & W \\ W^* & 0 \end{smallmatrix}\right)$, where W^* denotes the Hermitian conjugate of W. A different notion of tropical singular value has been studied in [DSDM02].

Note that $\|V - u \odot v\|_{\infty,\infty} \leqslant \lambda$ if and only if

$$-\lambda \leqslant V_{ij} - u_i - v_j \leqslant \lambda$$

for all i and j, i.e., the following inequalities

$$\begin{cases} V_{ij} - v_j - u_i & \leqslant \lambda \\ V_{kl} - v_l - u_k & \geqslant -\lambda \end{cases} \tag{5.12}$$

hold for all i, j, k and l. On the other hand, (5.12) is equivalent to

$$\begin{cases} \max_j (V_{ij} - v_j) & \leqslant \lambda + u_i \\ \max_j (-V_{jk} + u_j) & \leqslant \lambda - v_k \end{cases} \tag{5.13}$$

for all i and k. Thus by (5.13) we have the following proposition.

Proposition 5.23. *If (λ, u, v) is a skew singular tuple of V, then*

$$\|V - u \odot (-v)\|_{\infty,\infty} \leqslant \lambda. \tag{5.14}$$

Conversely, if (5.14) holds, we have

$$\begin{pmatrix} \bot & V \\ -V^\top & \bot \end{pmatrix} \begin{pmatrix} u \\ v \end{pmatrix} \leqslant \lambda \begin{pmatrix} u \\ v \end{pmatrix}, \tag{5.15}$$

by the tropical analogue of the Collatz-Wielandt theorem (see e.g. [ABG13]), the infimum λ^* of λ such that (5.15) holds is achieved by an eigenvector $\left(\begin{smallmatrix} u^* \\ v^* \end{smallmatrix}\right)$, i.e.,

$$\begin{pmatrix} \bot & V \\ -V^\top & \bot \end{pmatrix} \begin{pmatrix} u^* \\ v^* \end{pmatrix} = \lambda^* \begin{pmatrix} u^* \\ v^* \end{pmatrix},$$

which, together with Proposition 5.23, gives us the following theorem.

Theorem 5.24. *Let $V \in \mathbb{R}^{n\times p}$. Then, the infimum*

$$\lambda = \inf_{A \in \mathbb{R}^{n\times p}_{\max}, \operatorname{rank} A = 1} \|V - A\|_{\infty,\infty}$$

is achieved by some $A = u \odot (-v)$, where (λ, u, v) is a tropical skew singular tuples of V.

Thus by Theorem 5.13 and Theorem 5.16 we have

Theorem 5.25. *A matrix $V \in \mathbb{R}^{n\times p}$ has a unique skew singular value λ, which equals both one half of* $\operatorname{outerradius}(\operatorname{Col}(V))$ *and one half of the eigenvalue of $V \odot (-V^\top)$.*

Remark 5.26. Theorem 5.24 shows that the error in the sup-norm of the best tropical rank-one approximation of a matrix $V \in \mathbb{R}^{n\times p}$ is given by its skew singular value. This is a tropical analogue of the classical result in matrix theory, showing that the error in the spectral norm of the best rank-one approximation of a matrix $M \in \mathbb{R}^{n\times p}$ is given by the second singular value of M. Also, under the Frobenius norm, the best rank-one approximation error is $(\sigma_2^2 + \cdots + \sigma_m^2)^{1/2}$, where $\sigma_1 \geqslant \sigma_2 \cdots \geqslant \sigma_m$ are the singular values of M with $m = \min(n, p)$. Thus, the tropical skew singular value plays the role of the *second* singular value.

5.4.4 Rank one approximation of kernels

We now establish an infinite dimensional analogue of the previous results on rank-one approximation.

Let X be a compact metric space and let Y be a non-empty set, we denote by $\mathcal{C}(X)$ the set of continuous functions defined from X to $\mathbb{R}$, and by $\mathcal{B}(Y)$ the set of bounded functions defined from Y to $\mathbb{R}$.

Theorem 5.27 (Best tropical rank-one approximation)**.** *Suppose that $V : X \times Y \to \mathbb{R}$ is bounded, and that the family of functions $\{V(\cdot, y)\}_{y\in Y}$ is equicontinuous. Then, the best tropical rank-one approximation problem*

$$\inf_{f\in\mathcal{C}(X),\, g\in\mathcal{B}(Y)} \sup_{x\in X, y\in Y} |V(x, y) - f(x) - g(y)| \tag{5.16}$$

admits an optimal solution, obtained as follows. We set

$$H(x, z) = \sup_{y\in Y}(V(x, y) - V(z, y)) \ . \tag{5.17}$$

The value of the infimum is equal to one half of the tropical eigenvalue of H; it is achieved by any tropical eigenvector f of H and by the function g defined by:

$$g(y) = \frac{1}{2}\Big(\max_{z\in X}(V(z, y) - f(z)) + \min_{z\in X}(V(z, y) - f(z))\Big) \ . \tag{5.18}$$

Moreover, if Y is a compact metric space, and if the family of functions $\{V(x, \cdot)\}_{x\in X}$ is equicontinuous, the same conclusion holds with in addition $g \in \mathcal{C}(Y)$.

We shall need the following result from tropical spectral theory.

Theorem 5.28 (Corollary of Th.2.4 and Th. 2.7 of [KM97])**.** *Suppose X is a compact metric space, that $H : X \times X$ is bounded, and that the family of functions $\{H(\cdot, y)\}_{y\in Y}$ is equicontinuous in $\mathcal{C}(X)$. Then, the operator $f \mapsto Hf$ admits a tropical eigenvector in $\mathcal{C}(X)$. Moreover, the tropical eigenvalue is unique.* □

Note that Th. 2.4 and Th. 2.7 of [KM97] allow H to take the $-\infty$ value, and relax the compactness assumptions on X, modulo technical assumptions on H which are automatically satisfied when X is compact and H is bounded.

Proof of Theorem 5.27. Let H be defined as in (5.17). The assumption that the family of functions $\{V(\cdot, y)\}_{y\in Y} is$ equicontinuous entails that the family of functions $\{H(\cdot, z)\}_{z\in Z}$ is also equicontinuous. Moreover, if V is bounded, H is also bounded. Then, it follows from Theorem 5.28 that H admits a tropical eigenvector $f \in \mathcal{C}(X)$, associated to the unique tropical eigenvalue λ . Then, defining g by (5.18), we conclude, arguing as in the proof of Theorem 5.16, that (f, g) is an optimal solution of the optimization problem (5.16) and that $\lambda/2$ is the value of this optimization problem. Finally, if Y is a compact metric space, and if the family of functions $\{V(x, \cdot)\}_{x\in X}$ is equicontinuous, the function g defined by (5.18) is readily seen to belong to $\mathcal{C}(X)$. □

Proposition 5.21 carries over the infinite dimensional setting in a similar manner.

5.5 Rank-two approximation in dimension three

In this section we consider a matrix $V \in (\mathbb{R}_{\max})^{3\times p}$. We will show how we can do rank-two approximation of such matrices in linear time.

Given $i \in [n]$ and $a \in \mathbb{P}(\mathbb{R}_{\max})^n$, we define $\mathcal{H}_a^i$ the signed tropical hyperplane of type i as following:

$$\mathcal{H}_a^i := \{x \in (\mathbb{R}_{\max})^n \mid x_i + a_i = \max_{j\neq i}(x_j + a_j)\} \ .$$

In Proposition 4.35 of Chapter 4, it is shown that the distance of a vector $v \in (\mathbb{R}_{\max})^n$ to the signed tropical hyperplane $\mathcal{H}_a^i$ is given by:

$$\mathrm{dist}_H(v, \mathcal{H}_a^i) = |v_i + a_i - \max_{j \neq i}(v_j + a_j)| \ . \tag{5.19}$$

Now, we consider the case of matrices $V \in \mathbb{R}^{3\times p}$, so that the columns $V_{\cdot k}$ are in $\mathbb{R}^3$. We want to find a best rank-two approximation of V, i.e. solving the problem (5.3):

$$\min_{A \in (\mathbb{R}_{\max})^{3\times 2}, B \in (\mathbb{R}_{\max})^{2\times p}} d_{\infty,H}(V, AB) \ ,$$

from which we will deduce also the case of the $d_{\infty,\infty}$.

Owing to Corollary 5.11, we will look for a matrix $A \in (\mathbb{R}_{\max})^{3\times 2}$ that minimizes the distance of the columns of V to the column space $\mathrm{Col}(A)$:

$$\min_{A \in (\mathbb{R}_{\max})^{3\times 2}} \max_{k \in [p]} \mathrm{dist}_H(V_{\cdot k}, \mathrm{Col}(A)) \ . \tag{5.20}$$

The following lemma shows that we can restrict the minimization in (5.20) to matrices A such that the column space $\mathrm{Col}(A)$ is a signed tropical hyperplane of some type $i \in [3]$.

Lemma 5.29. *Let $A \in (\mathbb{R}_{\max})^{3\times 2}$. We take $i_1 \in \arg\min_{i\in[3]}(A_{i2} - A_{i1})$, $i_2 \in \arg\max_{i\in[3]\setminus\{i_1\}}(A_{i2} - A_{i1})$ and $i_3 \in [3] \setminus \{i_1, i_2\}$, we consider the matrix $\tilde{A} \in (\mathbb{R}_{\max})^{3\times 2}$ given by:*

$$\tilde{A}_{ij} = \begin{cases} A_{ij} & \text{if } \{i,j\} \notin \{(i_1,2),(i_2,1)\} \\ -\infty & \text{otherwise} \end{cases} \ ,$$

and the vector $a \in (\mathbb{R}_{\max})^3$ such that $a_{i_1} = A_{i_3 1} - A_{i_1 1}$, $a_{i_2} = A_{i_3 2} - A_{i_2 2}$ and $a_{i_3} = 0$. We have,

$$\mathrm{Col}(A) \subset \mathrm{Col}(\tilde{A}) = \mathcal{H}_a^{i_3} \ .$$

Proof. First we prove that $\mathrm{Col}(A) \subset \mathrm{Col}(\tilde{A})$ by showing that $A_{\cdot 1}, A_{\cdot 2} \in \mathrm{Col}(\tilde{A})$. We consider $\alpha_1 = \arg\min_{i\in[3]}(A_{i2} - A_{i1}) = A_{i_1 2} - A_{i_1 1}$ and $\alpha_2 = \arg\max_{i\in[3]}(A_{i2} - A_{i1}) = A_{i_2 2} - A_{i_2 1}$. We have $A_{\cdot 1} = \max(\tilde{A}_{\cdot 1}, -\alpha_2 + \tilde{A}_{\cdot 2}) \in \mathrm{Col}(\tilde{A})$. Indeed, we have $A_{i_1 1} = \tilde{A}_{i_1 1} = \max(\tilde{A}_{i_1 1}, -\alpha_2 + \tilde{A}_{i_1 2})$, $A_{i_2 1} = -\alpha_2 + A_{i_2 2} = -\alpha_2 + \tilde{A}_{i_2 2} = \max(\tilde{A}_{i_2 1}, -\alpha_2 + \tilde{A}_{i_2 2})$ and also $A_{i_3 1} = \max(A_{i_3 1}, -\alpha_2 + A_{i_3 2}) = \max(\tilde{A}_{i_3 1}, -\alpha_2 + \tilde{A}_{i_3 2})$. Similarly we can prove that $A_{\cdot 2} = \max(\alpha_1 + \tilde{A}_{\cdot 1}, \tilde{A}_{\cdot 2}) \in \mathrm{Col}(\tilde{A})$. Therefore, we deduce that $\mathrm{Col}(A) \subset \mathrm{Col}(\tilde{A})$.

Now, we will prove that $\mathrm{Col}(\tilde{A}) = \mathcal{H}_a^{i_3}$. We check first that $\tilde{A}_{\cdot 1}, \tilde{A}_{\cdot 2} \in \mathcal{H}_a^{i_3}$. Indeed, using the definition of a and that $\tilde{A}_{i_2 1} = -\infty$, we have $\tilde{A}_{i_3 1} + a_{i_3} = A_{i_3 1} = A_{i_1 1} + a_{i_1} = \max(\tilde{A}_{i_1 1} + a_{i_1}, \tilde{A}_{i_2 1} + a_{i_2})$, then $\tilde{A}_{\cdot 1} \in \mathcal{H}_a^{i_3}$. Similarly, we have also $\tilde{A}_{i_3 2} + a_{i_3} = A_{i_3 2} = A_{i_2 2} + a_{i_2} = \max(\tilde{A}_{i_2 2} + a_{i_2}, \tilde{A}_{i_1 2} + a_{i_1})$, then $\tilde{A}_{\cdot 2} \in \mathcal{H}_a^{i_3}$. Therefore, since $\mathcal{H}_a^{i_3}$ is a tropical cone, we get that $\mathrm{Col}(\tilde{A}) \subset \mathcal{H}_a^{i_3}$. Conversely, let $v \in \mathcal{H}_a^{i_3}$. We have $v_{i_3} = v_{i_3} + a_{i_3} = \max(v_{i_1} + a_{i_1}, v_{i_2} + a_{i_2}) = \max(v_{i_1} + A_{i_3 1} - A_{i_1 1}, v_{i_2} + A_{i_3 2} - A_{i_2 2}) = \max(\beta_1 + \tilde{A}_{i_3 1}, \beta_2 + \tilde{A}_{i_3 2})$, where $\beta_1 = v_{i_1} - A_{i_1 1}$ and $\beta_2 = v_{i_2} - A_{i_2 2}$. We have also $v_{i_1} = \max(\beta_1 + \tilde{A}_{i_1 1}, \beta_2 + \tilde{A}_{i_1 2})$ and $v_{i_2} = \max(\beta_1 + \tilde{A}_{i_2 1}, \beta_2 + \tilde{A}_{i_2 2})$. Therefore $v \in \mathrm{Col}(\tilde{A})$, and so we proved that $\mathrm{Col}(\tilde{A}) = \mathcal{H}_a^{i_3}$. □

Example 5.30. We illustrate Lemma 5.29 with the following matrix A, and the corresponding matrix $\tilde{A}$:

$$A = \begin{pmatrix} 0 & 0 \\ -1 & 0 \\ 0 & -1 \end{pmatrix}, \ \tilde{A} = \begin{pmatrix} 0 & 0 \\ -\infty & 0 \\ 0 & -\infty \end{pmatrix} \ .$$

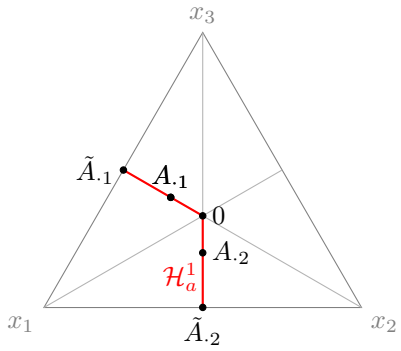


Figure (5.1) Illustration of Example 5.30.

Here, we have $i_1 = 3$, $i_2 = 2$, $i_3 = 1$ and $a = (0,0,0)^\top$. We have also $\mathrm{Col}(A) \subset \mathrm{Col}(\tilde{A}) = \mathcal{H}_a^1$ as shown in Figure 5.1.

Remark 5.31. From Lemma 5.29, we deduce that given $i \in [3]$ and $a \in \mathbb{P}(\mathbb{R}_{\max})^3$, we denote $[3] \setminus \{i\} = \{i_1, i_2\}$, then the matrix $A \in (\mathbb{R}_{\max})^{3\times 2}$ given by $A_{i_1 1} = a_i - a_{i_1}, A_{i_2 1} = -\infty, A_{i1} = 0$ and $A_{i_1 2} = -\infty, A_{i_2 2} = a_i - a_{i_2}, A_{i2} = 0$ satisfies $\mathcal{H}_a^i = \mathrm{Col}(A)$.

So now the problem that we want to solve is:

$$\min_{i\in[3]} \min_{a\in\mathbb{P}(\mathbb{R}_{\max})^3} \max_{k\in[p]} \mathrm{dist}_H(V_{\cdot k}, \mathcal{H}_a^i) \ . \tag{5.21}$$

Using the expression of the distance of a vector to a signed hyperplane given by (5.19), the problem to solve becomes:

$$\min_{i\in[3]} \min_{a\in\mathbb{P}(\mathbb{R}_{\max})^3} \max_{k\in[p]} |V_{ik} + a_i - \max_{j\neq i}(V_{jk} + a_j)| \ .$$

For $i \in [3]$ fixed, this is a special case of the "one-sided" tropical linear regression problem that we already discussed in Remark 4.41. We put this discussion again here (Proposition 5.32) for completeness. Given sample points $(x^{(k)}, y^{(k)})$ in $\mathbb{R}^n \times \mathbb{R}^m$, for $k \in [p]$, compute

$$\min_F \max_{k\in[p]} \|y^{(k)} - Fx^{(k)}\|_\infty \ , \tag{5.22}$$

where the minimum is taken over tropical matrices F of size $m \times n$, and the product $Fx^{(k)}$ is understood tropically. Up to a straightforward duality, this problem was solved in [But10, Theorem 3.5.2], the result being attributed there to Cuninghame-Green [CG79]. Alternatively, this solution may be recovered by combining [CF00, Coro. 1] with the explicit formula of the tropical projection [CGQ04, Th. 5]. The following proposition gives the explicit solution.

Proposition 5.32 (Theorem 3.5.2 of [But10])**.** *We consider the matrix $\bar{F} \in \mathbb{R}^{m\times n}$ defined by $\bar{F}_{ij} := \min_{k\in[p]}(y_i^{(k)} - x_j^{(k)})$. Let $\delta := \max_{k\in[p]} \|y^{(k)} - \bar{F}x^{(k)}\|_\infty$, and $F_{ij}^{opt} = \bar{F}_{ij} + \delta/2$. Then, F^{opt} is the greatest optimal solution of* (5.22)*, and it can be computed in $O(mnp)$ arithmetic operations.*

for a fixed $i \in [3]$ in our problem (5.21), we can take $a_i = 0$ because a is defined up to an additive constant, and we identify F as $(a_j)_{j\neq i}$, $x^{(k)}$ as $(V_{jk})_{j\neq i}$ and $y^{(k)}$ as V_{ik}. Then, by applying Proposition 5.32, we get the following result:

Corollary 5.33. *Let $V \in \mathbb{R}^{3\times p}$. For $i \in [3]$ fixed, we consider $\forall j \in [3] \setminus \{i\}, \bar{a}_j = \min_{k\in[p]}(V_{ik} - V_{jk})$ and $\delta = \max_{k\in[p]} |V_{ik} - \max_{j\neq i}(\bar{a}_j + V_{jk})|$. We take $\forall j \in [3] \setminus \{i\}, a_j^{opt} = \bar{a}_j + \delta/2$ and $a_i^{opt} = 0$. We have a^{opt} is the greatest optimal solution of the problem $\min_{a\in\mathbb{P}(\mathbb{R}_{\max})^3} \max_{k\in[p]} \mathrm{dist}_H(V_{\cdot k}, \mathcal{H}_a^i)$, and it can be computed in $O(p)$ arithmetic operations. Moreover, $\max_{k\in[p]} \mathrm{dist}_H(V_{\cdot k}, \mathcal{H}_{a^{opt}}^i) = \delta/2$.*

Algorithm 10 Rank two decomposition in dimension three

1: Input: matrix $V \in \mathbb{R}^{3\times p}$.
2: **for** $i \in [3]$ **do**:
3: $\quad \forall j \in [3] \setminus \{i\}, \bar{a}_j^i = \min_{k\in[p]}(V_{ik} - V_{jk})$ and $\delta^i = \max_{k\in[p]} |V_{ik} - \max_{j\neq i}(\bar{a}_j^i + V_{jk})|$,
4: $\quad \forall j \in [3] \setminus \{i\}, a_j^i = \bar{a}_j^i + \delta^i/2$ and $a_i^i = 0$,
5: $\quad e^i = \delta^i/2$,
6: **done**
7: take $i^* \in \arg\min_{i\in[3]} e^i$, and $a^* = a^{i^*}$,
8: take $\{i_1, i_2\} = [3] \setminus \{i^*\}$ and $A_{i_1 1}^* = -a_{i_1}^*, A_{i_2 1}^* = -\infty, A_{i^* 1}^* = 0$ and $A_{i_1 2}^* = -\infty, A_{i_2 2}^* = -a_{i_2}^*, A_{i^* 2}^* = 0$,
9: $\forall k \in [p], B_{1k}^H = \min(V_{i_1 k} - A_{i_1 1}^*, V_{i^* k})$ and $B_{2k}^H = \min(V_{i_2 k} - A_{i_2 2}^*, V_{i^* k})$,
10: $\forall k \in [p], B_{\cdot k}^\infty = B_{\cdot k}^H + \mathrm{ctr}(V_{\cdot k}, A^* B_{\cdot k}^H)e$,
11: **return** i^*, a^*, A^*, B^H and B^∞.

Theorem 5.34. *Let $V \in \mathbb{R}^{3\times p}$, Algorithm 10 returns the best rank-two approximation of the matrix V, such that*

$$\begin{aligned} d_{\infty,H}(V, A^* B^H) &= \min_{A\in(\mathbb{R}_{\max})^{3\times 2}, B\in(\mathbb{R}_{\max})^{2\times p}} d_{\infty,H}(V, AB) \\ &= \min_{A\in(\mathbb{R}_{\max})^{3\times 2}} \max_{k\in[p]} \mathrm{dist}_H(V_{\cdot k}, \mathrm{Col}(A)) = \max_{k\in[p]} \mathrm{dist}_H(V_{\cdot k}, \mathcal{H}_{a^*}^{i^*}) \\ &= 2d_{\infty,\infty}(V, A^* B^\infty) = 2 \min_{A\in(\mathbb{R}_{\max})^{3\times 2}, B\in(\mathbb{R}_{\max})^{2\times p}} d_{\infty,\infty}(V, AB) \ , \end{aligned}$$

and needs $O(p)$ arithmetic operations.

Proof. From Lemma 5.29, we know that we can restrict the minimization in (5.20) to the column spaces $\mathrm{Col}(A)$ that are signed hyperplanes of the form $\mathcal{H}_a^i$ with $a \in \mathbb{P}(\mathbb{R}_{\max})^3$ and $i \in [3]$, i.e. we need to solve (5.21). For $i \in [3]$ fixed, by Corollary 5.33, lines 3 and 4 give an a^i that is optimal in $\min_{a\in\mathbb{P}(\mathbb{R}_{\max})^3} \max_{k\in[p]} \mathrm{dist}_H(V_{\cdot k}, \mathcal{H}_a^i)$, and, line 7, ensures that (i^*, a^*) is the solution to (5.21).

Now, by Remark 5.31, line 8 gives a matrix A^* such that $\mathcal{H}_{a^*}^{i^*} = \mathrm{Col}(A^*)$, so that A^* is optimal in $\min_{A\in(\mathbb{R}_{\max})^{3\times 2}} \max_{k\in[p]} \mathrm{dist}_H(V_{\cdot k}, \mathrm{Col}(A))$.

Let $k \in [p]$, we have by definition of $\mathrm{Col}(A^*)$,

$$\mathrm{dist}_H(V_{\cdot k}, \mathcal{H}_{a^*}^{i^*}) = \mathrm{dist}_H(V_{\cdot k}, \mathrm{Col}(A^*)) = \min_{B_{\cdot k}\in(\mathbb{R}_{\max})^2} d_H(V_{\cdot k}, \max_{j\in[2]}(A_{\cdot j}^* + B_{jk})) \ . \tag{5.23}$$

We have $\mathrm{dist}_H(V_{\cdot k}, \mathcal{H}_{a^*}^{i^*}) = d_H(V_{\cdot k}, P_{\mathcal{H}_{a^*}^{i^*}}(V_{\cdot k}))$, where the vector $p = P_{\mathcal{H}_{a^*}^{i^*}}(V_{\cdot k})$ is the projection of $V_{\cdot k}$ into the hyperplane $\mathcal{H}_{a^*}^{i^*}$. Therefore, a vector $B_{\cdot k}$ that achieves the minimum in (5.23) is such that $\max_{j\in[2]}(A_{\cdot j}^* + B_{jk}) = p$. From Lemma 4.33 of Chapter 4, we get the expression of the projection into a hyperplane; we have $p_{i^*} = \min(V_{i^* k}, -a_{i^*}^* + \max(a_{i_1}^* + V_{i_1 k}, a_{i_2}^* + V_{i_2 k})) = \min(V_{i^* k}, \max(V_{i_1 k} - A_{i_1 1}^*, V_{i_2 k} - A_{i_2 2}^*))$, $p_{i_1} = \min(V_{i_1 k}, -a_{i_1}^* + a_{i^*}^* + V_{i^* k}) = \min(V_{i_1 k}, A_{i_1 1}^* + V_{i^* k})$ and $p_{i_2} = \min(V_{i_2 k}, -a_{i_2}^* + a_{i^*}^* + V_{i^* k}) = \min(V_{i_2 k}, A_{i_2 2}^* + V_{i^* k})$.

We have $p_{i_1} = \max_{j\in[2]}(A^*_{i_1 j} + B_{jk}) = A^*_{i_1 1} + B_{1k}$ because $A^*_{i_1 2} = -\infty$. Then, $B_{1k} = \min(V_{i_1 k} - A^*_{i_1 1}, V_{i^* k})$. Similarly using p_{i_2} we get that $B_{2k} = \min(V_{i_2 k} - A^*_{i_2 2}, V_{i^* k})$. Finally, by using the property $\max(\min(a,c), \min(a,c)) = \min(\max(a,b), c)$, we have $\max_{j\in[2]}(A^*_{i^* j} + B_{jk}) = \max(B_{1k}, B_{2k}) = \max(\min(V_{i_1 k} - A^*_{i_1 1}, V_{i^* k}), \min(V_{i_2 k} - A^*_{i_2 2}, V_{i^* k})) = \min(\max(V_{i_1 k} - A^*_{i_1 1}, V_{i_2 k} - A^*_{i_2 2}), V_{i^* k}) = p_{i^*}$. Therefore, this optimal matrix B, for (5.23), that we are looking for is exactly the one defined as B^H in line 9, and it satisfies $\max_{k\in[p]} \operatorname{dist}_H(V_{\cdot k}, \operatorname{Col}(A^*)) = \min_{B\in(\mathbb{R}_{\max})^{2\times p}} \max_{k\in[p]} d_H(V_{\cdot k}, \max_{j\in[2]}(A^*_{\cdot j} + B_{jk})) = \min_{B\in(\mathbb{R}_{\max})^{2\times p}} d_{\infty,H}(V, A^* B) = d_{\infty,H}(V, A^* B^H)$.

Finally, using Corollary 5.12, we deduce that i^*, a^*, A^*, B^H and B^∞ satisfy the desired property.

We check easily that each of the lines 3, 9 and 10 needs $O(p)$ arithmetic operations, and that the other lines need $O(1)$ arithmetic operations. Therefore Algorithm 10 has a complexity of $O(p)$. □

We observe that $\max(A^*_{\cdot 1}, A^*_{\cdot 2}) = -a^*$ is the apex of the optimal signed hyperplane $\mathcal{H}^{i^*}_{a^*}$.

Example 5.35. We illustrate Algorithm 10 with the following matrix :

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

Using Algorithm 10, we obtain the best $\|\cdot\|_{\infty,H}$ and $\|\cdot\|_{\infty,\infty}$ approximations of rank 2 of the matrix V, given by the following:

$$A^* = \begin{pmatrix} 0 & 0 \\ 0 & -\infty \\ -\infty & 0 \end{pmatrix}, B^H = \begin{pmatrix} 0 & 0 & -1 & -1 \\ -1 & -1 & 0 & 0 \end{pmatrix} \text{ and } B^\infty = \begin{pmatrix} 0.5 & 0.5 & -0.5 & -0.5 \\ -0.5 & -0.5 & 0.5 & 0.5 \end{pmatrix} ,$$

with the following approximation errors:

$$\min_{A\in(\mathbb{R}_{\max})^{3\times 2}, B\in(\mathbb{R}_{\max})^{2\times 4}} d_{\infty,H}(V, AB) = d_{\infty,H}(V, A^* B^H) = 1 \ ,$$
$$\min_{A\in(\mathbb{R}_{\max})^{3\times 2}, B\in(\mathbb{R}_{\max})^{2\times 4}} d_{\infty,\infty}(V, AB) = d_{\infty,\infty}(V, A^* B^\infty) = 0.5 \ .$$

We get also that $\operatorname{Col}(A^*) = \mathcal{H}^{i^*}_{a^*}$, where $\mathcal{H}^{i^*}_{a^*}$ is a signed hyperplane with a sign $i^* = 1$ and and apex $a^* = (0,0,0)^\top$. Figure 5.2 shows the signed hyperplane $\mathcal{H}^1_{a^*}$ that best approximates the set of point $(V_{\cdot k})_{k\in[4]}$ and also A^*.

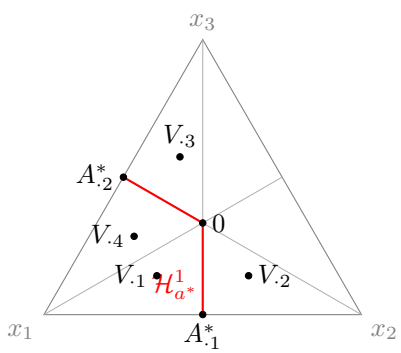


Figure (5.2) Illustration of Example 5.35.

CHAPTER 6

Conclusion and perspectives

In Chapter 2, we showed that a Nesterov accelerated scheme allows to accelerate the convergence of value iteration for non-symmetric affine fixed point problems (0-player), when the involved matrix has its spectrum in a specific region of the complex plane. We applied this result to obtain an accelerated policy iteration algorithm for Markov decision processes (1-player). An interesting open problem is to identify the conditions under which this Nesterov's type of scheme still yields an acceleration in the case of non-linear fixed point problems. The case of piecewise linear operators like the Bellman operator, is particularly interesting, since it will correspond to accelerating the value iteration algorithm for Markov decision processes. While the numerical experiments that we did in the case of accelerated value iteration for MDPs are promising, the theoretical analysis of the convergence of this accelerated scheme remains inherently difficult. Indeed, the characterization of the set of "accelerable" 0-player problems that we provided explains why this problem is difficult: in the 0-player problem, the convergence conditions are governed by fine spectral properties which have no known non-linear analogue in the one-player case.

In Chapter 3, we developed a deflation technique allowing to reduce a mean payoff problem to a discounted one under the hypothesis of existence of a distinguished state that is accessible from all other states and under all policies. This reduction allows to transfer results from the discounted case to the mean payoff case which is generally more difficult. In particular, we show that, as in the discounted case [SWWY18], we can develop a sublinear algorithm solving mean payoff stochastic games based on variance reduction techniques. An interesting question is related to the variance reduction value iteration algorithm proposed in [SWWY18]. This algorithm has a number of iterations that is fixed a priori which do not allow to profit from the natural speedup in convergence for generic instances.

So, it would be interesting to develop a variance reduction value iteration algorithm with an adaptive number of iterations first in the case of discounted games. Then, one could combine it with the deflation technique that we developed, to obtain a variance reduction deflated value iteration algorithm with an adaptive number of iterations in the case of mean payoff problems.

In Chapter 4, we solved the tropical linear regression problem, when the metric is of *sup-norm* type, and for tropical linear spaces of codimension 1 (tropical hyperplanes), but for a configuration of points of arbitrary cardinality. Several open problems related to the present work arise when changing either the class of metrics or of tropical spaces.

For instance, we may replace Hilbert's metric by the L_p-projective metric, i.e., the metric obtained by modding out the L_p normed space $\mathbb{R}^n$ by the action of additive constants, or by replacing the Hausdorff distance in (4.16) by a L_p type distance, for $p \in [1, \infty)$. Approaches based on mixed linear programming, or on local descent, have been proposed in [YZZ19, PYZ20, Hoo17] in some specific cases.

Another generalization consists in replacing hyperplanes by tropical linear spaces of a codimension not necessarily 1. We recall that a general *tropical linear space*, can be defined as

$$L(p) = \bigcap_I \{x \in (\mathbb{R}_{\max})^n \mid \max_{i \in I}(p_{I \setminus \{i\}} + x_i) \text{ is achieved at least twice}\} \ ,$$

where $p = (p_I) \in (\mathbb{R} \cup \{-\infty\})^{\binom{n}{k}}$ represents the *tropical Plücker coordinates* of an element of the *tropical Grassmannian* $\mathrm{Gr}^{\mathrm{trop}}_{k,n}$, see [SS04, FR15], and where the intersection is taken over all subsets I of $[n]$ of cardinality $k+1$. When $k = n-1$, $L(p)$ is a tropical hyperplane. Hence, given a set of points $\mathcal{V}$, a general version of tropical linear regression problem can be written as

$$\min_{p \in \mathrm{Gr}^{\mathrm{trop}}_{k,n}} \max_{v \in \mathcal{V}} \min_{x \in L(p)} \|v - x\|_H \ . \tag{6.1}$$

We solved this problem when $k = n-1$ in Chapter 4. When $k = 1$, $L(p)$ is reduced to a single point, and problem (6.1) is exactly the rank-one approximation of the matrix V whose columns are the elements of the set $\mathcal{V}$. We proved in Chapter 5 that this problem can be solved in strongly polynomial time. An interesting open question is to solve the problem (6.1) when $1 < k < n-1$. The same problem may be considered when p is a valuated matroid, or when it is inside the image of the Stiefel map [FR15], meaning that p is given by the maximal tropical minors of a matrix. A version of the latter problem (with a L_1-type error) is considered in [YZZ19].

In Chapter 5, we studied a tropical analogue of low-rank approximation for matrices. We established general properties of tropical low-rank approximation, and identified some very special classes, corresponding to rank-one and rank-two approximations, that are polynomial-time solvable. It will be interesting to identify wider classes of tractable low-rank approximation problems. Here, the linear space $L(p)$ in (6.1) is replaced by the column space $\mathrm{Col}(A)$ of a tropical matrix A with r columns

$$\min_{A \in (\mathbb{R}_{\max})^{n \times r}} \max_{v \in \mathcal{V}} \mathrm{dist}_H(v, \mathrm{Col}(A)) \ , \tag{6.2}$$

which is equivalent to the tropical rank-r approximation of the matrix $V \in (\mathbb{R}_{\max})^{n \times m}$ (see Proposition 5.10)

$$\min_{A \in (\mathbb{R}_{\max})^{n \times r}, B \in (\mathbb{R}_{\max})^{r \times m}} d_{\infty,H}(V, AB) \ . \tag{6.3}$$

Here also, another generalization consists is replacing Hilbert's metric or sup-norm metric by the L_p-projective metric, for $p \in [1, \infty)$.

APPENDIX $\mathcal{A}$

Résumé en français

Dans cette thèse, nous développons des algorithmes accélérés pour les processus de décision Markoviens (MDP) et plus généralement pour les jeux stochastiques à somme nulle (SG). Nous abordons également les problèmes de meilleure approximation qui se posent en géométrie tropicale. Ces deux domaines sont étroitement liés. Comme nous le montrerons dans la thèse, nous pouvons calculer des classes de problèmes de meilleure approximation tropicale en les réduisant à la résolution de jeux à somme nulle.

La programmation dynamique est l'une des principales approches utilisées pour résoudre les problèmes MDP et SG. Elle permet de transformer un jeu en un problème de point fixe faisant intervenir un opérateur appelé opérateur de Shapley (ou opérateur de Bellman dans le cas de MDP). L'itération sur les valeurs (VI) et l'itération sur les politiques (PI) sont les deux principaux algorithmes permettant de résoudre ces problèmes de point fixe. Cependant, dans le cas d'instances à grande échelle, ou lorsque l'on veut résoudre un problème à paiement moyen (où il n'y a pas de facteur d'escompte pour les paiements reçus dans le futur), les méthodes classiques deviennent lentes.

Dans la première partie de cette thèse, nous développons deux raffinements des algorithmes classiques d'itération sur les valeurs ou sur les politiques. Nous proposons d'abord une version accélérée de l'itération sur les valeurs (AVI) permettant de résoudre des problèmes de point fixe affines avec des matrices non auto-adjointes, ainsi qu'une version accélérée de l'itération sur les politiques (API) pour MDP, basée sur AVI. Cette accélération étend l'algorithme de gradient accéléré de Nesterov à une classe de problèmes de point fixe qui ne peuvent pas être interprétés en termes de programmation convexe. Nous caractérisons les spectres des matrices pour lesquelles cet algorithme converge avec un taux de convergence traduisant une accélération. Nous introduisons également un algorithme accéléré de degré d, et montrons qu'il donne un taux de convergence multi-accéléré sous des conditions plus exigeantes

sur le spectre des matrices. Nous montrons les performances de ces schémas d'accélération simples et multiples, sur des classes d'instances dans lesquelles les conditions spectrales d'accélération sont remplies. Ces classes comprennent un cadre de matrices aléatoires ainsi que l'équation de Hamilton Jacobi-Bellman dans le cas de petites dérives. Une autre contribution est une version déflatée de l'itération sur les valeurs (DVI) pour résoudre la version à paiement moyen des jeux stochastiques. Cette méthode permet de transformer un problème à paiement moyen en un problème escompté sous l'hypothèse d'existence d'un état distingué accessible depuis tous les autres états et sous toutes les politiques. Cette réduction combine un argument d'échelle (une version combinatoire de la h-transformation de Doob issue de la théorie des frontières des processus de Markov) et une technique de déflation: à un problème à paiement moyen, on associe un problème escompté, avec un taux d'escompte qui dépend de l'état. Nous obtenons un taux de contraction explicite pour le problème escompté qui dépend du temps de frappe maximal de l'état distingué, et qui apparaît dans nos bornes de complexité. En combinant cette méthode de déflation avec des techniques de réduction de la variance, nous obtenons également un algorithme sous-linéaire permettant de résoudre les jeux stochastiques à paiement moyen sous la même hypothèse d'accessibilité d'un état distingué.

Dans la deuxième partie de cette thèse, nous étudions différents problèmes de meilleure approximation tropicale. Nous résolvons d'abord un problème de régression linéaire tropicale consistant à trouver la meilleure approximation d'un ensemble de points par un hyperplan tropical. Notre résultat principal est un théorème de dualité forte montrant que la valeur de ce problème de régression coïncide avec le rayon maximal d'une boule de Hilbert incluse dans un polyèdre tropical. Cela nous permet de fournir un certificat d'optimalité qui peut être interprété géométriquement comme une collection de n points "témoins" parmi l'ensemble des points considéré où n est la dimension. Notre approche implique également que la régression linéaire tropicale est équivalente en temps polynomial à la résolution de jeux à paiement moyen. Nous appliquons ces résultats à un problème inverse de la théorie des enchères. Nous étudions également un analogue tropical de l'approximation de petit rang pour les matrices. Ceci est motivé par les méthodes approchées en programmation dynamique, dans lesquelles la fonction valeur est approximée par un supremum de fonctions élémentaires. Nous établissons des propriétés générales de l'approximation tropicale de petit rang et identifions des classes particulières de problèmes d'approximation de petit rang qui peuvent être résolus en temps polynomial. Nous donnons une interprétation géométrique au problème d'approximation matricielle de petit rang, en termes d'approximation d'une collection de points par un sous-module tropical avec peu de générateurs. Nous montrons en particulier que la meilleure approximation matricielle tropicale de rang un équivaut à trouver le rayon minimal d'une boule de Hilbert contenant un polyèdre tropical.

Bibliography

[ABG13] M. Akian, R. Bapat, and S. Gaubert. Max-plus algebras. In L. Hogben, editor, *Handbook of Linear Algebra (Discrete Mathematics and Its Applications)*, volume 39. Chapman & Hall/CRC, 2013. Chapter 25, Second Edition.

[AG13] Marianne Akian and Stéphane Gaubert. Policy iteration for perfect information stochastic mean payoff games with bounded first return times is strongly polynomial, 2013. arXiv:1310.4953.

[AGG09] M. Akian, S. Gaubert, and A. Guterman. Linear independence over tropical semirings and beyond. In G.L. Litvinov and S.N. Sergeev, editors, *Proceedings of the International Conference on Tropical and Idempotent Mathematics*, volume 495 of *Contemporary Mathematics*, pages 1–38. American Mathematical Society, 2009.

[AGG12] M. Akian, S. Gaubert, and A. Guterman. Tropical polyhedra are equivalent to mean payoff games. *International Journal of Algebra and Computation*, 22(1):125001 (43 pages), 2012.

[AGH15a] M. Akian, S. Gaubert, and A. Hochart. Ergodicity conditions for zero-sum games. *Discrete & Continuous Dynamical Systems*, 35(9):3901–3931, 2015.

[AGH15b] Marianne Akian, Stéphane Gaubert, and Antoine Hochart. Ergodicity conditions for zero-sum games. *Discrete and Continuous Dynamical Systems, series A*, 35(9):3901–3931, 2015.

[AGH18] Marianne Akian, Stéphane Gaubert, and Antoine Hochart. Generic uniqueness of the bias vector of finite stochastic games with perfect information. *Journal of Mathematical Analysis and Applications*, 457(2):1038–1064, 2018.

[AGH20] M. Akian, S. Gaubert, and A. Hochart. A game theory approach to the existence and uniqueness of nonlinear perron-frobenius eigenvectors. *Discrete & Continuous Dynamical Systems - A*, 40:207–231, 2020.

[AGK11] X. Allamigeon, S. Gaubert, and R. Katz. The number of extreme points of tropical polyhedra. *J. Comb. Theory Series A*, 118(1):162–189, 2011.

[AGKS18] X. Allamigeon, S. Gaubert, R. Katz, and M. Skomra. Condition numbers of stochastic mean payoff games and what they say about nonarchimedean semidefinite programming. In *Proceedings of the 23rd International Symposium on Mathematical Theory of Networks and Systems (MTNS2018)*, 2018.

[AGN11] Marianne Akian, Stéphane Gaubert, and Roger Nussbaum. A Collatz-Wielandt characterization of the spectral radius of order-preserving homogeneous maps on cones, 2011. arXiv:1112.5968.

[AGNS11] M. Akian, S. Gaubert, V. Niţică, and I. Singer. Best approximation in max-plus semimodules. *Linear Algebra and its Applications*, 435(12):3261–3296, 2011.

[AGQS19] Marianne Akian, Stéphane Gaubert, Zheng Qu, and Omar Saadi. Solving ergodic markov decision processes and perfect information zero-sum stochastic games by variance reduced deflated value iteration. In *2019 IEEE 58th Conference on Decision and Control (CDC)*, pages 5963–5970. IEEE, 2019.

[AGQS20] Marianne Akian, Stéphane Gaubert, Zheng Qu, and Omar Saadi. Multiply accelerated value iteration for non-symmetric affine fixed point problems and application to markov decision processes. *arXiv preprint arXiv:2009.10427*, 2020.

[AGQS21] Marianne Akian, Stéphane Gaubert, Yang Qi, and Omar Saadi. Tropical linear regression and mean payoff games: or, how to measure the distance to equilibria. *arXiv preprint arXiv:2106.01930*, 2021.

[Ami03] Rabah Amir. Stochastic games in economics and related fields: an overview. *Stochastic games and applications*, pages 455–470, 2003.

[And65] Donald Anderson. Iterative procedures for nonlinear integral equations. *Journal of the ACM*, 12(4):547–560, 1965.

[AP19] Hedy Attouch and Juan Peypouquet. Convergence of inertial dynamics and proximal algorithms governed by maximally monotone operators. *Mathematical Programming*, 174(1):391–432, 2019.

[Ard04] F. Ardila. Subdominant matroid ultrametrics. *Annals of Combinatorics*, 8:379–389, 2004.

[Att21] Hedy Attouch. Fast inertial proximal ADMM algorithms for convex structured optimization with linear constraint. *Minimax Theory and its Application*, 6(1):1–24, 2021. hal-02501604.

[BB92] J. B. Baillon and R. E. Bruck. Optimal rates of asymptotic regularity for averaged nonexpansive mappings. In K. K. Tan, editor, *Proceedings of the Second International Conference on Fixed Point Theory and Applications*, pages 27–66. World Scientific Press, 1992.

[BB96] J. B. Baillon and R. E. Bruck. The rate of asymptotic regularity is $\mathrm{O}(1/\sqrt{n})$. *Lecture Notes in Pure and Applied Mathematics*, pages 51–81, 1996.

[BCC08] Charles Bordenave, Pietro Caputo, and Djalil Chafai. Circular law theorem for random Markov matrices. *Probability Theory and Related Fields*, 152, 08 2008.

[BCOQ92] F. Baccelli, G. Cohen, G.J. Olsder, and J.P. Quadrat. *Synchronization and Linearity*. Wiley, 1992.

[Bel52] Richard Bellman. On the theory of dynamic programming. *Proceedings of the National Academy of Sciences of the United States of America*, 38(8):716, 1952.

[Bel57] R. Bellman. *Dynamic Programming*. Princeton University Press, Princeton, NJ, 1957.

[Ber11] Dimitri P Bertsekas. Approximate policy iteration: A survey and some new methods. *Journal of Control Theory and Applications*, 9(3):310–335, 2011.

[Ber17] Dimitri P Bertsekas. *Dynamic programming and optimal control, Vol. I and Vol. II, 4th ed*. Athena scientific Belmont, 2012 & 2017.

[Ber20] D. I. Bernstein. L-infinity optimization to Bergman fans of matroids with an application to phylogenetics. *Siam J. Discrete Math.*, 34(1):701–720, 2020.

[BH13] Casey C Bennett and Kris Hauser. Artificial intelligence framework for simulating clinical decision-making: A markov decision process approach. *Artificial intelligence in medicine*, 57(1):9–19, 2013.

[BK76] Truman Bewley and Elon Kohlberg. The asymptotic theory of stochastic games. *Mathematics of Operations Research*, 1(3):197–208, 1976.

[BK19] E. Baldwin and P. Klemperer. Understanding preferences: "demand types", and the existence of equilibrium with indivisibilities. *Econometrica*, 87(3):867–932, 2019.

[BP07] R. Burguet and Martin K. Perry. Bribery and favoritism by auctioneers in sealed-bid auctions. *The B.E. Journal of Theoretical Economics*, 7(1):1–27, June 2007.

[BR11] Nicole Bäuerle and Ulrich Rieder. *Markov decision processes with applications to finance*. Springer Science & Business Media, 2011.

[BSS07] P. Butkovič, H. Schneider, and S. Sergeev. Generators, extremals and bases of max cones. *Linear Algebra Appl.*, 421(2-3):394–406, 2007.

[BT91a] Dimitri P Bertsekas and John N Tsitsiklis. An analysis of stochastic shortest path problems. *Mathematics of Operations Research*, 16(3):580–595, 1991.

[BT91b] Dimitri P Bertsekas and John N Tsitsiklis. An analysis of stochastic shortest path problems. *Mathematics of Operations Research*, 16(3):580–595, 1991.

[BT96] D. P. Bertsekas and J. N. Tsitsiklis. *Neuro-Dynamic Programming*. Athena Scientific, 1996.

[But03a] P. Butkovič. Max-algebra: the linear algebra of combinatorics? *Linear Algebra Appl.*, 367:313–335, 2003.

[But03b] Peter Butkovič. Max-algebra: the linear algebra of combinatorics? *Linear Algebra and its applications*, 367:313–335, 2003.

[But10] P. Butkovič. *Max-linear Systems: Theory and Algorithms*. Springer Monogr. Math. Springer, London, 2010.

[BV07] H. Bjorklund and S. Vorobyov. A combinatorial strongly subexponential strategy improvement algorithm for mean payoff games. *Discrete Appl. Math.*, 155:210229, 2007.

[CF00] V. Chepoi and B. Fichet. ℓ_∞-approximation via subdominants. *Journal of Mathematical Psychology*, 44:600–616, 2000.

[CG79] R. A. Cuninghame-Green. *Minimax algebra*, volume 166 of *Lecture Notes in Economics and Mathematical Systems*. Springer-Verlag, Berlin, 1979.

[CGQ99] Guy Cohen, Stéphane Gaubert, and Jean-Pierre Quadrat. Max-plus algebra and system theory: where we are and where to go now. *Annual reviews in control*, 23:207–219, 1999.

[CGQ04] G. Cohen, S. Gaubert, and J. P. Quadrat. Duality and separation theorems in idempotent semimodules. *Linear Algebra and Appl.*, 379:395–422, 2004.

[Cha09] J. Chaloupka. Parallel algorithms for mean-payoff games: an experimental evaluation. In *Algorithms—ESA 2009*, volume 5757 of *Lecture Notes in Comput. Sci.*, pages 599–610. Springer, Berlin, 2009.

[CLMV05] O. Compte, A. Lambert-Mogiliansky, and T. Verdier. Corruption and competition in procurement auctions. *The RAND Journal of Economics*, 36(1):1–15, 2005.

[CLSW98] Francis H. Clarke, Yuri S. Ledyaev, Ronald J. Stern, and Peter R. Wolenski. *Nonsmooth Analysis and Control Theory*. Springer, 1998.

[CM18] Vasileios Charisopoulos and Petros Maragos. A tropical approach to neural networks with piecewise linear activations. *arXiv preprint arXiv:1805.08749*, 2018.

[CMY15] Caihua Chen, Shiqian Ma, and Junfeng Yang. A general inertial proximal point algorithm for mixed variational inequality problem. *SIAM Journal on Optimization*, 25(4):2120–2142, 2015.

[cod21] *Code de la commande publique*. Journal officiel, 2021. Available from legifrance.gouv.fr.

[Con92] Anne Condon. The complexity of stochastic games. *Information and Computation*, 96(2):203–224, 1992.

[CSV14] R. Cominetti, J. A. Soto, and J. Vaisman. On the rate of convergence of Krasnosel'skii-Mann iterations and their connection with sums of Bernoullis. *Israel Journal of Mathematics*, 199(2):757–772, 2014.

[DG06] V. Dhingra and S. Gaubert. How to solve large scale deterministic games with mean payoff by policy iteration. In *Proceedings of the 1st international conference on Performance evaluation methodolgies and tools (VALUETOOLS)*, volume 180, Pisa, Italy, 2006. article No. 12.

[DGJ17] J. Depersin, S. Gaubert, and M. Joswig. A tropical isoperimetric inequality. *Séminaire Lotharingien de Combinatoire*, 78B, 2017. Article #27, 12 pp., Proceedings of FPSAC 2017 (29th Conference on Formal Power Series and Algebraic Combinatorics, London).

[DKK21] Sergey Dolgov, Dante Kalise, and Karl K Kunisch. Tensor decomposition methods for high-dimensional hamilton–jacobi–bellman equations. *SIAM Journal on Scientific Computing*, 43(3):A1625–A1650, 2021.

[DKM01] V. Danilov, G. Koshevoy, and K. Murota. Discrete convexity and equilibria in economies with indivisible goods and money. *Mathematical Social Sciences*, 41(3):251–273, May 2001.

[DS04] M. Develin and B. Sturmfels. Tropical convexity. *Doc. Math.*, 9:1–27, 2004. (Erratum pp. 205–206).

[DSDM02] Bart De Schutter and Bart De Moor. The QR decomposition and the singular value decomposition in the symmetrized max-plus algebra revisited. *SIAM Rev.*, 44(3):417–454, 2002. Reprint of SIAM J. Matrix Anal. App. **1**9 (1998), no. 2, 378–406 (electronic) [MR1614050 (98m:15021)].

[DSS05] M. Develin, F. Santos, and B. Sturmfels. On the rank of a tropical matrix. In *Combinatorial and computational geometry*, volume 52 of *Math. Sci. Res. Inst. Publ.*, pages 213–242. Cambridge Univ. Press, Cambridge, 2005.

[DT14] Yoel Drori and Marc Teboulle. Performance of first-order methods for smooth convex minimization: a novel approach. *Mathematical Programming*, 145(1):451–482, 2014.

[Dyn69] E.B. Dynkin. Boundary theory of Markov processes (the discrete case). *Russian Math. Surveys*, 24(7):1–42, 1969.

[EB92] J. Eckstein and D. P. Bertsekas. On the Douglas-Rachford splitting method and the proximal point algorithm for maximal monotone operators. *Mathematical Programming*, 55:293–318, 1992.

[EKL06] M. Einsiedler, M. Kapranov, and D. Lind. Non-Archimedean amoebas and tropical varieties. *J. Reine Angew. Math.*, 601:139–157, 2006.

[EM79] A. Ehrenfeucht and J. Mycielski. Positional strategies for mean payoff games. *Internat. J. Game Theory*, 8(2):109–113, 1979.

[EY36] C. Eckart and G. Young. The approximation of one matrix by another of lower rank. *Psychometrika*, 1(3):211–218, 1936.

[FABG13] O. Fercoq, M. Akian, M. Bouhtou, and S. Gaubert. Ergodic control and polyhedral approaches to pagerank optimization. *IEEE-TAC*, 58(1):134–148, 2013.

[FB15] Nicolas Flammarion and Francis Bach. From averaging to acceleration, there is only a step-size. In Peter Grünwald, Elad Hazan, and Satyen Kale, editors, *Proceedings of The 28th Conference on Learning Theory*, volume 40 of *Proceedings of Machine Learning Research*, pages 658–695, 03–06 Jul 2015.

[Fea10] John Fearnley. Exponential lower bounds for policy iteration. In *International Colloquium on Automata, Languages, and Programming*, pages 551–562. Springer, 2010.

[FH13] Eugene A. Feinberg and Jefferson Huang. Strong polynomiality of policy iterations for average-cost MDPs modeling replacement and maintenance problems. *Operations Research Letters*, 41(3):249–251, May 2013.

[FH14] Eugene A Feinberg and Jefferson Huang. The value iteration algorithm is not strongly polynomial for discounted dynamic programming. *Operations Research Letters*, 42(2):130–131, 2014.

[FM00] Wendell H Fleming and William M McEneaney. A max-plus-based algorithm for a hamilton–jacobi–bellman equation of nonlinear filtering. *SIAM Journal on Control and Optimization*, 38(3):683–710, 2000.

[FR12] Wendell H Fleming and Raymond W Rishel. *Deterministic and stochastic optimal control*, volume 1. Springer Science & Business Media, 2012.

[FR15] A. Fink and F. Rincón. Stiefel tropical linear spaces. *Journal of Combinatorial Theory, Series A*, 135:291 – 331, 2015.

[Fri09a] O. Friedmann. An exponential lower bound for the parity game strategy improvement algorithm as we know it. In *Proceedings of the 24th Annual IEEE Symposium on Logic in Computer Science (LICS 2009)*, pages 145–156. IEEE Computer Society, 2009.

[Fri09b] Oliver Friedmann. An exponential lower bound for the parity game strategy improvement algorithm as we know it. In *Proceedings of the Twenty-Fourth Annual IEEE Symposium on Logic in Computer Science (LICS 2009)*, pages 145–156. IEEE Computer Society Press, August 2009.

[Fri11] Oliver Friedmann. An exponential lower bound for the latest deterministic strategy iteration algorithms. *Logical Methods in Computer Science*, 7(3), 2011.

[FS06] Wendell H Fleming and Halil Mete Soner. *Controlled Markov processes and viscosity solutions*, volume 25. Springer Science & Business Media, 2006.

[FV12] Jerzy Filar and Koos Vrieze. *Competitive Markov decision processes*. Springer Science & Business Media, 2012.

[GFJ15] E. Ghadimi, H. R. Feyzmahdavian, and M. Johansson. Global convergence of the heavy-ball method for convex optimization. In *2015 European Control Conference (ECC)*, pages 310–315, 2015.

[GG04] S. Gaubert and J. Gunawardena. The Perron-Frobenius theorem for homogeneous, monotone functions. *Trans. of AMS*, 356(12):4931–4950, 2004.

[GGC19] Vineet Goyal and Julien Grand-Clement. A first-order approach to accelerated value iteration, 2019. arXiv:1905.09963.

[GJ08] B. Gärtner and M. Jaggi. Tropical support vector machines. Technical report, ACSTR-362502-01, 2008.

[GJG15] A. Gupta, R. Jain, and P. W. Glynn. An empirical algorithm for relative value iteration for average-cost MDPs. In *54th IEEE Conference on Decision and Control (CDC)*, pages 5079–5084, Dec 2015.

[GK07] S. Gaubert and R. Katz. The Minkowski theorem for max-plus convex sets. *Linear Algebra and Appl.*, 421:356–369, 2007.

[GK11] S. Gaubert and R. Katz. Minimal half-spaces and external representation of tropical polyhedra. *Journal of Algebraic Combinatorics*, 33(3):325–348, 2011.

[GKK88] V. A. Gurvich, A. V. Karzanov, and L. G. Khachiyan. Cyclic games and finding minimax mean cycles in digraphs. *Zh. Vychisl. Mat. i Mat. Fiz.*, 28(9):1407–1417, 1439, 1988.

[GLPN93] Giorgio Gallo, Giustino Longo, Stefano Pallottino, and Sang Nguyen. Directed hypergraphs and applications. *Discrete applied mathematics*, 42(2-3):177–201, 1993.

[GM84] Michel Gondran and Michel Minoux. Linear algebra in dioids: a survey of recent results. In *North-Holland mathematics studies*, volume 95, pages 147–163. Elsevier, 1984.

[GM06] Bernd Gärtner and Jiří Matoušek. *Understanding and Using Linear Programming*. Springer, 2006.

[GMQ11] Stephane Gaubert, William McEneaney, and Zheng Qu. Curse of dimensionality reduction in max-plus based approximation methods: Theoretical estimates and improved pruning algorithms. In *2011 50th IEEE Conference on Decision and Control and European Control Conference*, pages 1054–1061. IEEE, 2011.

[GP13] D. Grigoriev and V. V. Podolskii. Complexity of tropical and min-plus linear prevarieties. *Computational Complexity*, pages 1–34, 2013.

[GS20] S. Gaubert and N. Stott. A convergent hierarchy of non-linear eigenproblems to compute the joint spectral radius of nonnegative matrices. *Mathematical Control and Related Fields*, 10(3):573–590, 2020.

[GTZ97] S. A. Goreinov, E. E. Tyrtyshnikov, and N. L. Zamarashkin. A theory of pseudoskeleton approximations. *Linear Algebra Appl.*, 261:1–21, 1997.

[HMZ13] Thomas Dueholm Hansen, Peter Bro Miltersen, and Uri Zwick. Strategy iteration is strongly polynomial for 2-player turn-based stochastic games with a constant discount factor. *Journal of the ACM*, 60(1):1–16, February 2013.

[Hoo17] J. Hook. Linear regression over the max-plus semiring: algorithms and applications, 2017. arXiv:1712.03499.

[Hot33] H. Hotelling. Analysis of a complex of statistical variables into principal components. *J. Educ. Psychol.*, 24:417–441, 1933.

[HOVDW14] Bernd Heidergott, Geert Jan Olsder, and Jacob Van Der Woude. *Max plus at work*. Princeton University Press, 2014.

[How60] Ronald A Howard. *Dynamic programming and Markov processes*. John Wiley, 1960.

[IH19] F. Iutzeler and J. M. Hendrickx. A generic online acceleration scheme for optimization algorithms via relaxation and inertia. *Optimization Methods and Software*, 34(2):383–405, 2019.

[IMS09] Ilia Itenberg, Grigory Mikhalkin, and Eugenii I Shustin. *Tropical algebraic geometry*, volume 35. Springer Science & Business Media, 2009.

[IR09] Z. Izhakian and L. Rowen. The tropical rank of a tropical matrix. *Communications in Algebra*, 37(11):3912–3927, 2009.

[IS14] A. Izmailov and M. Solodov. *Newton-Type Methods for Optimization and Variational Problems*. Springer, 03 2014.

[JSY07] M. Joswig, B. Sturmfels, and J. Yu. Affine buildings and tropical convexity. *Alban. J. Math.*, 1:187–211, 2007.

[Kar78] R.M. Karp. A characterization of the minimum mean-cycle in a digraph. *Discrete Maths.*, 23:309–311, 1978.

[KGNZ19] Wei Kang, Qi Gong, and Tenavi Nakamura-Zimmerer. Algorithms of data development for deep learning and feedback design. *arXiv preprint arXiv:1912.00492*, 2019.

[Kim82] K.H. Kim. *Boolean Matrix Theory and Applications*. Marcel Dekker, New York, 1982.

[Kim19] Donghwan Kim. Accelerated proximal point method for maximally monotone operators, 2019.

[KKDD01] Harold Joseph Kushner Kushner, Harold J Kushner, Paul G Dupuis, and Paul Dupuis. *Numerical methods for stochastic control problems in continuous time*, volume 24. Springer Science & Business Media, 2001.

[KM97] V. N. Kolokoltsov and V. P. Maslov. *Idempotent analysis and its applications*, volume 401 of *Mathematics and its Applications*. Kluwer Academic Publishers Group, Dordrecht, 1997. Translation of *Idempotent analysis and its application in optimal control* (Russian), "Nauka" Moscow, 1994 [MR1375021 (97d:49031)], Translated by V. E. Nazaikinskii, With an appendix by Pierre Del Moral.

[Koh80] E. Kohlberg. Invariant half-lines of nonexpansive piecewise-linear transformations. *Math. Oper. Res.*, 5(3):366–372, 1980.

[Kra55] M. A. Krasnosel'skiň. Two remarks on the method of successive approximations. *Uspekhi Matematicheskikh Nauk*, 10:123–127, 1955.

[Kri02] V. Krishna. *Auction theory*. Academic Press, 2002.

[Lie21] Felix Lieder. On the convergence rate of the Halpern-iteration. *Optimization Letters*, 15:405–418, 2021.

[Lit94] Michael L Littman. Markov games as a framework for multi-agent reinforcement learning. In *Machine learning proceedings 1994*, pages 157–163. Elsevier, 1994.

[LL69a] T. M. Liggett and S. A. Lippman. Stochastic games with perfect information and time average payoff. *SIAM Rev.*, 11:604–607, 1969.

[LL69b] T. M. Liggett and S. A. Lippman. Stochastic games with perfect information and time average payoff. *SIAM Rev.*, 11(4):604–607, 1969.

[LSTY17] B. Lin, B. Sturmfels, X. Tang, and R. Yoshida. Convexity in tree spaces. *SIAM Journal on Discrete Mathematics*, 31(3):2015–2038, January 2017.

[Man53] W. R. Mann. Mean value methods in iteration. *Proceedings of the American Mathematical Society*, 4:506–510, 1953.

[Mar02] D. Marker. *Model Theory: An Introduction*, volume 217 of *Grad. Texts in Math.* Springer, New York, 2002.

[McE07] William M McEneaney. A curse-of-dimensionality-free numerical method for solution of certain HJB PDEs. *SIAM journal on Control and Optimization*, 46(4):1239–1276, 2007.

[MCT21] P. Maragos, V. Charisopoulos, and E. Theodosis. Tropical geometry and machine learning. *Proceedings of the IEEE*, 109(5):728–755, May 2021.

[MDG08] William M McEneaney, Ameet Deshpande, and Stephane Gaubert. Curse-of-complexity attenuation in the curse-of-dimensionality-free method for HJB PDEs. In *2008 American Control Conference*, pages 4684–4690. IEEE, 2008.

[MKH11] William M McEneaney, Hidehiro Kaise, and Seung Hak Han. Idempotent method for continuous-time stochastic control and complexity attenuation. *IFAC Proceedings Volumes*, 44(1):3216–3221, 2011.

[MN81a] J.-F. Mertens and A. Neyman. Stochastic games. *Internat. J. Game Theory*, 10(2):53–66, 1981.

[MN81b] J.-F. Mertens and A. Neyman. Stochastic games. *Internat. J. Game Theory*, 10(2):53–66, 1981.

[MN20] Pauli Miettinen and Stefan Neumann. Recent developments in boolean matrix factorization. In Christian Bessiere, editor, *Proceedings of the Twenty-Ninth International Joint Conference on Artificial Intelligence, IJCAI-20*, pages 4922–4928. International Joint Conferences on Artificial Intelligence Organization, 7 2020. Survey track.

[MPN02] John Mallet-Paret and Roger D Nussbaum. Eigenvalues for a class of homogeneous cone maps arising from max-plus operators. *Discrete and Continuous Dynamical Systems*, 8(3):519–562, 2002.

[MS15a] D. Maclagan and B. Sturmfels. *Introduction to Tropical Geometry*, volume 161 of *Grad. Stud. Math.* AMS, Providence, RI, 2015.

[MS15b] Diane Maclagan and Bernd Sturmfels. *Introduction to tropical geometry*, volume 161. American Mathematical Soc., 2015.

[MSZ15a] J.-F. Mertens, S. Sorin, and S. Zamir. *Repeated games*. Cambridge University Press, 2015.

[MSZ15b] Jean-François Mertens, Sylvain Sorin, and Shmuel Zamir. *Repeated games*, volume 55. Cambridge University Press, 2015.

[Nes83] Y. Nesterov. A method of solving a convex programming problem with convergence rate $o(1/k^2)$. *Soviet Mathematics Doklady*, 27(2):372–376, 1983.

[Nes04] Yurii Nesterov. *Introductory Lectures on Convex Optimization: A Basic Course (Applied Optimization)*. Kluwer Academic Publishers, 2004.

[NS03] A. Neyman and S. Sorin, editors. *Stochastic games and applications*, volume 570 of *NATO Science Series C: Mathematical and Physical Sciences*. Kluwer Academic Publishers, Dordrecht, 2003.

[Nus86a] R. D. Nussbaum. Convexity and log convexity for the spectral radius. *Linear Algebra Appl.*, 73:59–122, 1986.

[Nus86b] Roger D Nussbaum. Convexity and log convexity for the spectral radius. *Linear Algebra and its Applications*, 73:59–122, 1986.

[OSS19] Mathias Oster, Leon Sallandt, and Reinhold Schneider. Approximating the stationary hamilton-jacobi-bellman equation by hierarchical tensor products. *arXiv preprint arXiv:1911.00279*, 2019.

[Pea01] K. Pearson. On lines and planes of closest fit to systems of points in space. *Philos. Mag.*, 2(6):559–572, 1901.

[Pol64] B.T. Polyak. Some methods of speeding up the convergence of iteration methods. *USSR Computational Mathematics and Mathematical Physics*, 4(5):1 – 17, 1964.

[Put14] Martin L Puterman. *Markov decision processes: discrete stochastic dynamic programming*. John Wiley & Sons, 2014.

[PYZ20] R. Page, R. Yoshida, and L. Zhang. Tropical principal component analysis on the space of phylogenetic trees. *Bioinformatics*, 36(17):4590–4598, 06 2020.

[Qu14] Zheng Qu. A max-plus based randomized algorithm for solving a class of HJB PDEs. In *53rd IEEE Conference on Decision and Control*, pages 1575–1580. IEEE, 2014.

[Rac13] S. Rachmilevitch. Bribing in first-price auctions. *Games and Economic Behavior*, 77(1):214–228, January 2013.

[RF91] TES Raghavan and Jerzy A Filar. Algorithms for stochastic games—a survey. *Zeitschrift für Operations Research*, 35(6):437–472, 1991.

[rFS09] Haw ren Fang and Yousef Saad. Two classes of multisecant methods for nonlinear acceleration. *Numer. Linear Algebra Appl.*, 16:197–221, 2009.

[RGST05] Jurgen Richter-Gebert, Bernd Sturmfels, and Thorsten Theobald. First steps in tropical geometry. *Contemporary Mathematics*, 377:289–318, 2005.

[Roc76a] R. T. Rockafellar. Augmented Lagrangians and applications of the proximal point algorithm in convex programming. *Mathematics of operations research*, 1(2):97–116, 1976.

[Roc76b] R. Tyrrell Rockafellar. Monotone operators and the proximal point algorithm. *SIAM Journal on Control and Optimization*, 14(5):877–898, 1976.

[SBD17] Damien Scieur, Francis Bach, and Alexandre D'Aspremont. Nonlinear acceleration of stochastic algorithms. In I. Guyon, U. V. Luxburg, S. Bengio, H. Wallach, R. Fergus, S. Vishwanathan, and R. Garnett, editors, *Advances in Neural Information Processing Systems*, volume 30. Curran Associates, Inc., 2017.

[Sch08] S. Schewe. An optimal strategy improvement algorithm for solving parity and payoff games. In M. Kaminski and S. Martini, editors, *Computer Science Logic*, pages 369–384, Berlin, Heidelberg, 2008. Springer Berlin Heidelberg.

[Sch13] Bruno Scherrer. Improved and generalized upper bounds on the complexity of policy iteration. In *Advances in Neural Information Processing Systems*, pages 386–394, 2013.

[Sch16] Bruno Scherrer. Improved and generalized upper bounds on the complexity of policy iteration. *Math. Oper. Res.*, 41(3):758–774, 2016.

[Ser07] S. Sergeev. Max-plus definite matrix closures and their eigenspaces. *Linear Algebra Appl.*, 421(2):182–201, 2007.

[SGG+15] Bruno Scherrer, Mohammad Ghavamzadeh, Victor Gabillon, Boris Lesner, and Matthieu Geist. Approximate modified policy iteration and its application to the game of tetris. *J. Mach. Learn. Res.*, 16:1629–1676, 2015.

[Sha53] Lloyd S. Shapley. Stochastic games. *Proceedings of the national academy of sciences*, 39(10):1095–1100, 1953.

[Shi12] Yaroslav Shitov. On tropical matrices of small factor rank. *Linear algebra and its applications*, 437(11):2727–2732, 2012.

[Shi14] Yaroslav Shitov. The complexity of tropical matrix factorization. *Advances in Mathematics*, 254:138–156, 2014.

[Sko18] M. Skomra. *Spectraèdres tropicaux : application à la programmation semi-définie et aux jeux à paiement moyen*. Theses, Université Paris-Saclay, December 2018.

[SMGJ14] Srinivas Sridharan, William M McEneaney, Mile Gu, and Matthew R James. A reduced complexity min-plus solution method to the optimal control of closed quantum systems. *Applied Mathematics & Optimization*, 70(3):469–510, 2014.

[SS04] D. Speyer and B. Sturmfels. The tropical Grassmannian. *Adv. Geom.*, 4(3):389–411, 2004.

[SWW+18] Aaron Sidford, Mengdi Wang, Xian Wu, Lin F. Yang, and Yinyu Ye. Near-optimal time and sample complexities for solving discounted Markov decision process with a generative model, 2018. arXiv:1806.01492.

[SWWY18] Aaron Sidford, Mengdi Wang, Xian Wu, and Yinyu Ye. Variance reduced value iteration and faster algorithms for solving Markov decision processes. In *Proceedings of the Twenty-Ninth Annual ACM-SIAM Symposium on Discrete Algorithms*, pages 770–787, 2018.

[TM19] A. Tsiamis and P. Maragos. Sparsity in max-plus algebra and systems. *Discrete Event Dynamic Systems*, 29:163–189, 2019.

[TP19] A. Themelis and P. Patrinos. Supermann: A superlinearly convergent algorithm for finding fixed points of nonexpansive operators. *IEEE Transactions on Automatic Control*, 64(12):4875–4890, 2019.

[Tse90] Paul Tseng. Solving h-horizon, stationary Markov decision problems in time proportional to log (h). *Operations Research Letters*, 9(5):287–297, 1990.

[TTM20] Nikos Tsilivis, Anastasios Tsiamis, and Petros Maragos. Sparse approximate solutions to max-plus equations with application to multivariate convex regression. *arXiv preprint arXiv:2011.04468*, 2020.

[TWY20] X. Tang, H. Wang, and R. Yoshida. Tropical support vector machine and its applications to phylogenomics, 2020. arXiv:2003.00677.

[TY19] N. M. Tran and J. Yu. Product-mix auctions and tropical geometry. *Mathematics of Operations Research*, 44(4):1396–1411, November 2019.

[Vai90] Pravin M Vaidya. An algorithm for linear programming which requires $O(((m+n)n^2 + (m+n)^{1.5}n)L)$ arithmetic operations. *Mathematical Programming*, 47(1):175–201, 1990.

[VOW12] Martijn Van Otterlo and Marco Wiering. Reinforcement learning and markov decision processes. In *Reinforcement learning*, pages 3–42. Springer, 2012.

[Wan17] Mengdi Wang. Primal-dual π learning: Sample complexity and sublinear run time for ergodic Markov decision problems, 2017. arXiv:1710.06100.

[Whi63] Douglas J White. Dynamic programming, markov chains, and the method of successive approximations. *Journal of Mathematical Analysis and Applications*, 6(3):373–376, 1963.

[Whi83] P. Whittle. *Optimization over time. Vol. II*. Wiley Series in Probability and Mathematical Statistics: Applied Probability and Statistics. John Wiley & Sons Ltd., Chichester, 1983.

[Wil09] Byron K Williams. Markov decision processes in natural resources management: Observability and uncertainty. *Ecological Modelling*, 220(6):830–840, 2009.

[WN11] Homer F. Walker and Peng Ni. Anderson acceleration for fixed-point iterations. *SIAM Journal on Numerical Analysis*, 49(4):1715–1735, 2011.

[Ye11a] Yinyu Ye. The simplex and policy-iteration methods are strongly polynomial for the Markov decision problem with a fixed discount rate. *Mathematics of Operations Research*, 36(4):593–603, 2011.

[Ye11b] Yinyu Ye. The simplex and policy-iteration methods are strongly polynomial for the Markov decision problem with a fixed discount rate. *Mathematics of Operations Research*, 36(4):593–603, 2011.

[YZZ19] R. Yoshida, L. Zhang, and X. Zhang. Tropical principal component analysis and its application to phylogenetics. *Bulletin of Mathematical Biology*, 81:568–597, 2019.

[ZNL18] Liwen Zhang, Gregory Naitzat, and Lek-Heng Lim. Tropical geometry of deep neural networks. In *International Conference on Machine Learning*, pages 5824–5832. PMLR, 2018.

[ZOB20] Junzi Zhang, Brendan O'Donoghue, and Stephen Boyd. Globally convergent type-I Anderson acceleration for nonsmooth fixed-point iterations. *SIAM Journal on Optimization*, 30(4):3170–3197, 2020.

[ZP96] U. Zwick and M. Paterson. The complexity of mean payoff games on graphs. *Theoret. Comput. Sci.*, 158(1-2):343–359, 1996.

ECOLE
DOCTORALE
DE MATHEMATIQUES
HADAMARD

Titre : Jeux répétés à somme nulle: algorithmes accélérés et meilleure approximation tropicale

Mots clés : Jeux stochastiques, processus de décision Markoviens, accélération de Nesterov, itération sur les valeurs déflatée, régression linéaire tropicale, approximation tropicale de petit rang

Résumé : Dans cette thèse, nous développons des algorithmes accélérés pour les processus de décision Markoviens (MDP) et plus généralement pour les jeux stochastiques à somme nulle (SG). Nous abordons également les problèmes de meilleure approximation qui se posent en géométrie tropicale.
Dans la première partie de cette thèse, nous développons deux raffinements des algorithmes classiques d'itération sur les valeurs ou sur les politiques. Nous proposons d'abord une version accélérée de l'itération sur les valeurs (AVI) permettant de résoudre des problèmes de point fixe affines avec des matrices non auto-adjointes, ainsi qu'une version accélérée de l'itération sur les politiques (API) pour MDP, basée sur AVI. Nous introduisons également un algorithme accéléré de degré d, et montrons qu'il donne un taux de convergence multi-accéléré. Une autre contribution est une version déflatée de l'itération sur les valeurs (DVI) pour résoudre la version à paiement moyen des jeux stochastiques. Cette méthode permet de transformer un problème à paiement moyen en un problème escompté. En combinant cette méthode de déflation avec des techniques de réduction de la variance, nous obtenons un algorithme sous-linéaire résolvant les jeux stochastiques à paiement moyen.
Dans la deuxième partie de cette thèse, nous étudions différents problèmes de meilleure approximation tropicale. Nous résolvons d'abord un problème de régression linéaire tropicale consistant à trouver la meilleure approximation d'un ensemble de points par un hyperplan tropical. Nous montrons que la valeur de ce problème de régression coïncide avec le rayon maximal d'une boule de Hilbert incluse dans un polyèdre tropical, et que ce problème est équivalent en temps polynomial aux jeux à paiement moyen. Nous appliquons ces résultats à un problème inverse de la théorie des enchères. Nous étudions également un analogue tropical de l'approximation de petit rang pour les matrices. Ceci est motivé par les méthodes approchées en programmation dynamique, dans lesquelles la fonction valeur est approximée par un supremum de fonctions élémentaires. Nous établissons des propriétés générales de l'approximation tropicale de petit rang et identifions des classes particulières de problèmes d'approximation de petit rang qui peuvent être résolus en temps polynomial.

Title : Zero-sum repeated games: accelerated algorithms and tropical best approximation

Keywords : Stochastic games, Markov decision processes, Nesterov acceleration, deflated value iteration, tropical linear regression, tropical low-rank approximation

Abstract :
In this thesis, we develop accelerated algorithms for Markov decision processes (MDP) and more generally for zero-sum stochastic games (SG). We also address best approximation problems arising in tropical geometry.
In the first part of this thesis, we develop two refinements of the classical value or policy iteration algorithms. We first propose an accelerated version of value iteration (AVI) allowing to solve affine fixed point problems with non self-adjoint matrices, alongside with an accelerated version of policy iteration (API) for MDP, building on AVI. We also introduce an accelerated algorithm of degree d, and show that it yields a multiply accelerated rate of convergence. Another contribution is a deflated version of value iteration (DVI) to solve the mean payoff version of stochastic games. This method allows one to transform a mean payoff problem to a discounted one. Combining this deflation method with variance reduction techniques, we derive a sublinear algorithm solving mean payoff stochastic games.
In the second part of this thesis, we study tropical best approximation problems. We first solve a tropical linear regression problem consisting in finding the best approximation of a set of points by a tropical hyperplane. We show that the value of this regression problem coincides with the maximal radius of a Hilbert's ball included in a tropical polyhedron, and that this problem is polynomial-time equivalent to mean payoff games. We apply these results to an inverse problem from auction theory. We study also a tropical analogue of low-rank approximation for matrices. This is motivated by approximate methods in dynamic programming, in which the value function is approximated by a supremum of elementary functions. We establish general properties of tropical low-rank approximation, and identify classes of low-rank approximation problems that are polynomial-time solvable.

Institut Polytechnique de Paris
91120 Palaiseau, France