SPRINGER BRIEFS IN MATHEMATICS

Sabine Burgdorf
Igor Klep
Janez Povh

Optimization of Polynomials in Non-Commuting Variables

Springer

SpringerBriefs in Mathematics

Series Editors

Nicola Bellomo
Michele Benzi
Palle Jorgensen
Tatsien Li
Roderick Melnik
Otmar Scherzer
Benjamin Steinberg
Lothar Reichel
Yuri Tschinkel
George Yin
Ping Zhang

SpringerBriefs in Mathematics showcases expositions in all areas of mathematics and applied mathematics. Manuscripts presenting new results or a single new result in a classical field, new field, or an emerging topic, applications, or bridges between new results and already published works, are encouraged. The series is intended for mathematicians and applied mathematicians.

More information about this series at http://www.springer.com/series/10030

Sabine Burgdorf • Igor Klep • Janez Povh

Optimization of Polynomials in Non-Commuting Variables

Springer

Sabine Burgdorf
Centrum Wiskunde & Informatica
Amsterdam, The Netherlands

Igor Klep
Department of Mathematics
The University of Auckland
Auckland, New Zealand

Janez Povh
Faculty of Information Studies
in Novo Mesto
Novo Mesto, Slovenia

ISSN 2191-8198 ISSN 2191-8201 (electronic)
SpringerBriefs in Mathematics
ISBN 978-3-319-33336-6 ISBN 978-3-319-33338-0 (eBook)
DOI 10.1007/978-3-319-33338-0

Library of Congress Control Number: 2016938173

Mathematics Subject Classification (2010): 90C22, 14P10, 90C26, 08B20, 13J30, 47A57

© The Author(s) 2016
This work is subject to copyright. All rights are reserved by the Publisher, whether the whole or part of the material is concerned, specifically the rights of translation, reprinting, reuse of illustrations, recitation, broadcasting, reproduction on microfilms or in any other physical way, and transmission or information storage and retrieval, electronic adaptation, computer software, or by similar or dissimilar methodology now known or hereafter developed.
The use of general descriptive names, registered names, trademarks, service marks, etc. in this publication does not imply, even in the absence of a specific statement, that such names are exempt from the relevant protective laws and regulations and therefore free for general use.
The publisher, the authors and the editors are safe to assume that the advice and information in this book are believed to be true and accurate at the date of publication. Neither the publisher nor the authors or the editors give a warranty, express or implied, with respect to the material contained herein or for any errors or omissions that may have been made.

Printed on acid-free paper

This Springer imprint is published by Springer Nature
The registered company is Springer International Publishing AG Switzerland

Introduction

Optimization problems involving polynomial data arise across many sciences, e.g., in control theory [Che10, HG05, Sch06], operations research [Sho90, Nie09], statistics and probability [Las09], combinatorics and graph theory [LS91, AL12], computer science [PM81], and elsewhere. They are however difficult to solve. For example, very simple instances of polynomial optimization problems (POPs) are known to be NP hard. Because of their importance, various algorithms have been devised to approximately solve POPs. Traditionally techniques have drawn from operations research, computer science, and numerical analysis. Since the boom in semidefinite programming (SDP) in the 1990s, newer techniques for solving POPs are based on sums of squares concepts taken from real algebraic geometry and inspired by moment theory from probability and functional analysis. There are now many excellent packages available for solving POPs based on these methods, such as GloptiPoly [HLL09], SOSTOOLS [PPSP05], SparsePOP [WKK$^+$09], or YALMIP [Löf04].

In this book, our focus is on polynomial optimization problems in matrix unknowns, i.e., non-commutative POPs or NCPOPs for short. Many applied problems, for example, those in all the textbook classics in control theory [SIG97], have matrices as variables, and the formulas naturally involve polynomials in matrices. These polynomials depend only on the system layout and do not change with the size of the matrices involved; such problems are "dimension-free." Analyzing them is in the realm of *free analysis* [KVV14] and *free real algebraic geometry* (free RAG) [BPT13].

The booming area of free analysis provides an analytic framework for dealing with quantities with the highest degree of non-commutativity, such as large (random) matrices. Free RAG is its branch that studies positivity of polynomials in freely non-commuting (nc) matrix variables. In recent years, free RAG has found many applications of which we mention only a small selection. NCPOPs are ubiquitous.

Pironio, Navascués, and Acín [NPA08, PNA10] give applications to quantum theory and quantum information science and also consider computational aspects of NCPOPs. In quantum theory, NCPOPs are used to produce upper bounds on the maximal violation of Bell inequalities [PV09]. These inequalities provide a method to investigate *entanglement*, one of the most peculiar features of quantum mechanics, which allows two parties to be correlated in a non-classical way. In the same spirit, in [DLTW08], the authors investigate the quantum moment problem and entangled multi-prover games using NCPOPs. NCPOPs can also be used in quantum chemistry to compute atomic and molecular ground state energies, etc. A famous open problem due to Tsirelson [JNP$^+$11, Fri12] asks whether every quantum mechanical system can be modeled in finite-dimensional spaces. Tsirelson's problem is equivalent to two big questions in operator algebras, Kirchberg's conjecture [Kir93] and Connes' embedding conjecture [Con76]. The latter of these has a natural reformulation as a question on NCPOPs [KS08a, BDKS14, Oza04]. Closely related to Tsirelson's problem is a question about the right model for non-local quantum correlations. Without details, there are two widely accepted models, one with a tensor product structure of operators and one where only commutativity between operators located at different sites is assumed. A variant of Tsirelson's problem would imply that both models describe the same set of quantum correlations. Whereas for the latter model one can apply sums of hermitian squares (SOHS) to check for positivity, as it is done by Pironio, Navascués, and Acín [NPA08, PNA10], the former model is more difficult due to the tensor product structure. Mančinska and Roberson [MR14] and Sikora and Varvitsiotis [SV15] showed that bipartite quantum correlations from the tensor model can be written as projection of an affine section of the completely positive semidefinite cone introduced by Laurent and Piovesan [LP15]. Formally, it is the cone of Gram matrices of tuples of positive semidefinite matrices. But this cone can also be derived by dualizing a certain cone of nc polynomials with positive trace, bringing NCPOPs back into the picture; see also [BLP15].

Helton et al. in [HMdOP08] survey applications and connections to control and systems engineering. Free RAG and NCPOPs are employed to enforce convexity in classes of dimension-free problems. Cimprič [Cim10] uses NCPOPs to investigate PDEs and eigenvalues of polynomial partial differential operators. Inspired by randomized algorithms in machine learning, Recht and Re [RR12] investigate the arithmetic-geometric mean inequality for matrices with the aid of NCPOPs.

Finally, we mention an application of NCPOPs to statistical physics. The Bessis-Moussa-Villani (BMV) conjecture [BMV75] (now a theorem of Stahl [Sta13]) arose from an attempt to simplify the calculation of partition functions of quantum mechanical systems. It states that for any two symmetric matrices A,B, where B is positive semidefinite, the function $t \mapsto \operatorname{tr}\left(e^{A-tB}\right)$ is the Laplace transform of a positive Borel measure with real support. This permits the calculation of explicit upper and lower bounds of energy levels in multiple particle systems. The BMV conjecture is intimately related with positivity of certain symmetric nc polynomials [KS08b, Bur11].

We developed `NCSOStools` [CKP11] as a consequence of this recent flurry of interest in free RAG. `NCSOStools` [CKP11] is an open-source Matlab toolbox for handling NCPOPs. It solves unconstrained and constrained NCPOPs, either optimizing for eigenvalues or trace of an nc polynomial objective function, by converting them to a standard SDP which is then solved using one of the existing solvers such as SeDuMi [Stu99], SDPT3 [TTT99], or SDPA [YFK03]. As a side product, our toolbox implements symbolic computation with nc variables in Matlab. This book presents the theoretical underpinnings needed for all the algorithms we implemented with examples computed in `NCSOStools` [CKP11].

Organization of the Book

The book is organized as follows. Chapter 1 collects all the background material from algebra, functional analysis, and mathematical optimization needed throughout the book. On the algebraic side, we introduce non-commutative polynomials, commutators, sums of hermitian squares, quadratic modules, and semialgebraic sets from free RAG. Then we discuss the nc moment problem and its solution via flatness and the Gelfand-Naimark-Segal (GNS) construction. Finally, the chapter concludes with a discussion of SDP.

Our basic tool to minimize the eigenvalues of an nc polynomial is based on SOHS. In fact, by Helton's sums of squares theorem [Hel02], an nc polynomial is positive semidefinite if and only if it is an SOHS. Chapter 2 explains how to test if a given nc polynomial is an SOHS. This is based on an appropriate variant of the Gram matrix method; an nc polynomial is an SOHS if and only if the associated SDP is feasible. What is new and in sharp contrast to the commutative case is the complexity of the constructed SDP. Namely, its order is linear in the size of the input data. This is obtained from a careful analysis of the nc Newton polytope and the so-called Newton chip method.

Observe that a matrix has nonnegative trace if and only if it is a sum of a positive semidefinite matrix (a hermitian square) and a trace zero matrix (a commutator). Motivated by this simple observation, we propose a sum of hermitian squares and commutators certificate for trace positivity of nc polynomials. These certificates are analyzed in Chap. 3. We provide tracial analogs of the Gram matrix method and the Newton polytope.

In Chap. 4, we turn to optimization of nc polynomials. We present unconstrained and constrained optimizations. Uncostrained optimization is a single SDP, while we give a Lasserre-type [Las01] relaxation scheme for constrained optimization. This includes a study of exactness based on the Curto-Fialkow [CF96, CF98] flatness results generalized to the non-commutative setting. Special attention is given to special cases of convex constraint sets, i.e., nc balls and nc polydiscs. Their constrained optimization reduces to a single SDP.

Finally, Chap. 5 presents tracial optimization of nc polynomials and tracial analogs of the results in Chap. 4.

References

[AL12] Anjos, M.F., Lasserre, J.B.: Handbook of Semidefinite, Conic and Polynomial Optimization: Theory, Algorithms, Software and Applications. International Series in Operational Research and Management Science, vol. 166. Springer, New York (2012)

[BMV75] Bessis, D., Moussa, P., Villani, M.: Monotonic converging variational approximations to the functional integrals in quantum statistical mechanics. J. Math. Phys. **16**(11), 2318–2325 (1975)

[BPT13] Blekherman, G., Parrilo, P.A., Thomas, R.R.: Semidefinite Optimization and Convex Algebraic Geometry, vol. 13. SIAM, Philadelphia (2013)

[Bur11] Burgdorf, S.: Sums of hermitian squares as an approach to the BMV conjecture. Linear Multilinear Algebra **59**(1), 1–9 (2011)

[BDKS14] Burgdorf, S., Dykema, K., Klep, I., Schweighofer, M.: Addendum to "Connes' embedding conjecture and sums of hermitian squares" [Adv. Math. 217 (4) (2008) 1816–1837]. Adv. Math. **252**, 805–811 (2014)

[BLP15] Burgdorf, S., Laurent, M., Piovesan, T.: On the closure of the completely positive semidefinite cone and linear approximations to quantum colorings. arXiv preprint. arXiv:1502.02842 (2015)

[CKP11] Cafuta, K., Klep, I., Povh, J.: NCSOStools: a computer algebra system for symbolic and numerical computation with noncommutative polynomials. Optim. Methods Softw. **26**(3), 363–380 (2011). Available from http://ncsostools.fis.unm.si/

[Che10] Chesi, G.: LMI techniques for optimization over polynomials in control: a survey. IEEE Trans. Autom. Control **55**(11), 2500–2510 (2010)

[Cim10] Cimprič, J.: A method for computing lowest eigenvalues of symmetric polynomial differential operators by semidefinite programming. J. Math. Anal. Appl. **369**(2), 443–452 (2010)

[Con76] Connes, A.: Classification of injective factors. Cases II_1, II_∞, III_λ, $\lambda \neq 1$. Ann. Math. (2) **104**(1), 73–115 (1976)

[CF96] Curto, R.E., Fialkow, L.A.: Solution of the truncated complex moment problem for flat data. Mem. Am. Math. Soc. **119**(568), x+52 (1996)

[CF98] Curto, R.E., Fialkow, L.A.: Flat extensions of positive moment matrices: recursively generated relations. Mem. Am. Math. Soc. **136**(648), x+56 (1998)

[DLTW08] Doherty, A.C., Liang, Y.-C., Toner, B., Wehner, S.: The quantum moment problem and bounds on entangled multi-prover games. In: 23rd Annual IEEE Conference on Computational Complexity, 2008. CCC'08, pp. 199–210. IEEE, LOS ALAMOS (2008)

[Fri12] Fritz, T.: Tsirelson's problem and Kirchberg's conjecture. Rev. Math. Phys. **24**(05), 1250012 (2012)

[Hel02] Helton, J.W.: "Positive" noncommutative polynomials are sums of squares. Ann. Math. (2) **156**(2), 675–694 (2002)

[HG05] Henrion, D., Garulli, A.: Positive Polynomials in Control, vol. 312. Springer, Heidelberg (2005)

[HLL09] Henrion, D., Lasserre, J.B., Löfberg, J.: GloptiPoly 3: moments, optimization and semidefinite programming. Optim. Methods Softw. **24**(4–5), 761–779 (2009). Available from http://www.laas.fr/~henrion/software/gloptipoly3/

[HMdOP08] Helton, J.W., McCullough, S., de Oliveira, M.C., Putinar, M.: Engineering Systems and Free Semi-Algebraic Geometry. In: Emerging Applications of Algebraic Geometry. The IMA Volumes in Mathematics and Its Applications, vol. 149, pp. 17–62. Springer, New York (2008)

[JNP$^+$11] Junge, M., Navascues, M., Palazuelos, C., Perez-Garcia, D., Scholz, V.B., Werner, R.F.: Connes' embedding problem and Tsirelson's problem. J. Math. Phys. **52**(1), 012102 (2011)

[KVV14] Kaliuzhnyi-Verbovetskyi, D.S., Vinnikov, V.: Foundations of Free Noncommutative Function Theory, vol. 199. American Mathematical Society, Providence (2014)

[Kir93] Kirchberg, E.: On non-semisplit extensions, tensor products and exactness of group C*-algebras. Invent. Math. **112**(1), 449–489 (1993)

[KS08a] Klep, I., Schweighofer, M.: Connes' embedding conjecture and sums of hermitian squares. Adv. Math. **217**(4), 1816–1837 (2008)

[KS08b] Klep, I., Schweighofer, M.: Sums of hermitian squares and the BMV conjecture. J. Stat. Phys **133**(4), 739–760 (2008)

[Las01] Lasserre, J.B.: Global optimization with polynomials and the problem of moments. SIAM J. Optim. **11**(3), 796–817 (2000/01)

[Las09] Lasserre, J.B.: Moments, Positive Polynomials and Their Application. Imperial College Press, London (2009)

[LP15] Laurent, M., Piovesan, T.: Conic approach to quantum graph parameters using linear optimization over the completely positive semidefinite cone. SIAM J. Optim. **25**(4), 2461–2493 (2015)

[Löf04] Löfberg, J.: YALMIP: a toolbox for modeling and optimization in MATLAB. In: Proceedings of the CACSD Conference, Taipei (2004). Available from http://control.ee.ethz.ch/~joloef/wiki/pmwiki.php

[LS91] Lovász, L., Schrijver, A.: Cones of matrices and set-functions and 0–1 optimization. SIAM J. Optim. **1**(2), 166–190 (1991)

[MR14] Mančinska, L., Roberson, D.E.: Note on the correspondence between quantum correlations and the completely positive semidefinite cone. Available from http://quantuminfo.quantumlah.org/memberpages/laura/corr.pdf (2014)

[NPA08] Navascués, M., Pironio, S., Acín, A.: A convergent hierarchy of semidefinite programs characterizing the set of quantum correlations. N. J. Phys. **10**(7), 073013 (2008)

[Nie09] Nie, J.: Sum of squares method for sensor network localization. Comput. Optim. Appl. **43**(2), 151–179 (2009)

[Oza04] Ozawa, N.: About the QWEP conjecture. Int. J. Math. **15**(05), 501–530 (2004)

[PV09] Pál, K.F., Vértesi, T.: Quantum bounds on Bell inequalities. Phys. Rev. A **79**, 022120 (2009)

[PM81] Paz, A., Moran, S.: Non deterministic polynomial optimization problems and their approximations. Theor. Comput. Sci. **15**(3), 251–277 (1981)

[PNA10] Pironio, S., Navascués, M., Acín, A.: Convergent relaxations of polynomial optimization problems with noncommuting variables. SIAM J. Optim. **20**(5), 2157–2180 (2010)

[PPSP05] Prajna, S., Papachristodoulou, A., Seiler, P., Parrilo, P.A.: SOSTOOLS and its control applications. In: Positive Polynomials in Control, vol. 312. Lecture Notes in Control and Informaion Science, pp. 273–292. Springer, Berlin (2005)

[RR12] Recht, B., Ré, C.: Beneath the valley of the noncommutative arithmetic-geometric mean inequality: conjectures, case-studies, and consequences. In: JMLR: Workshop and Conference Proceedings, pp. 11.1–11.24 (2012)

[Sch06] Scherer, C.W.: LMI relaxations in robust control. Eur. J. Control **12**(1), 3–29 (2006)

[Sho90] Shor, N.Z.: Dual quadratic estimates in polynomial and Boolean programming. Ann. Oper. Res. **25**(1), 163–168 (1990)

[SV15] Sikora, J., Varvitsiotis, A.: Linear conic formulations for two-party correlations and values of nonlocal games. arXiv preprint. arXiv:1506.07297 (2015)

[SIG97] Skelton, R.E., Iwasaki, T., Grigoriadis, D.E.: A Unified Algebraic Approach to Control Design. CRC Press, Boca Raton (1997)

[Sta13] Stahl, H.R.: Proof of the BMV conjecture. Acta Math. **211**(2), 255–290 (2013)

[Stu99] Sturm, J.F.: Using SeDuMi 1.02, a MATLAB toolbox for optimization over symmetric cones. Optim. Methods Softw. **11/12**(1–4), 625–653 (1999). Available from http://sedumi.ie.lehigh.edu/

[TTT99] Toh, K.C., Todd, M.J., Tütüncü, R.: SDPT3–a MATLAB software package for semidefinite programming, version 1.3. Optim. Methods Softw. **11/12**(1–4), 545–581 (1999). Available from http://www.math.nus.edu.sg/~mattohkc/sdpt3.html

[WKK$^+$09] Waki, H., Kim, S., Kojima, M., Muramatsu, M., Sugimoto, H.: Algorithm 883: sparsePOP—a sparse semidefinite programming relaxation of polynomial optimization problems. ACM Trans. Math. Softw. **35**(2), Art. 15, 13 (2009)

[YFK03] Yamashita, M., Fujisawa, K., Kojima, M.: Implementation and evaluation of SDPA 6.0 (semidefinite programming algorithm 6.0). Optim. Methods Softw. **18**(4), 491–505 (2003). Available from http://sdpa.sourceforge.net/

Contents

List of Figures

List of Tables

Chapter 1
Selected Results from Algebra and Mathematical Optimization

1.1 Positive Semidefinite Matrices

Positive semidefinite matrices will be used extensively throughout the book. Therefore we fix notation here and present some basic properties needed later on.

Definition 1.1. A matrix $A \in \mathbb{R}^{n\times n}$ is *symmetric* if $A^T = A$. We denote the vector space of all symmetric matrices of order n by $\mathbb{S}_n$. Further, we denote the set of k-tuples $\underline{A} = (A_1,\ldots,A_k)$ of symmetric matrices A_i of order n by $\mathbb{S}_n^k$; and set $\mathbb{S}^k = \bigcup_{n\geq 1}\mathbb{S}_n^k$ if we consider tuples of arbitrary order.

Definition 1.2. A matrix $A \in \mathbb{S}_n$ is *positive semidefinite* (*definite*) if $\mathbf{x}^T A\mathbf{x} \geq 0$ $(\mathbf{x}^T A\mathbf{x} > 0)$ for any nonzero $\mathbf{x} \in \mathbb{R}^n$. We denote the sets of positive semidefinite and positive definite matrices by $\mathbb{S}_n^+$ and $\mathbb{S}_n^{++}$, respectively. By $A \succeq B$ $(A \succ B)$ we denote that $A - B \in \mathbb{S}_n^+$ $(A - B \in \mathbb{S}_n^{++})$.

In the following theorems we quote the most important properties of positive (semi)definite matrices. By $\lambda_{\min}(\cdot)$ we denote the smallest eigenvalue of (symmetric) matrix. Proofs of the theorems can be found in, e.g., [HJ12, Sect. 7.2].

Theorem 1.3. *Let $A \in \mathbb{S}_n$. The following are equivalent:*

(i) $A \in \mathbb{S}_n^+$,
(ii) $\lambda_{\min}(A) \geq 0$,
(iii) *there exists a matrix B such that $A = B^T B$,*
(iv) $\det A_{II} \geq 0$ *for any principal submatrix $A_{II} = [a_{i,j}]_{i,j\in I}$, $I \subseteq \{1,\ldots,n\}$.*

Theorem 1.4. *For a matrix $A \in \mathbb{S}_n$ the following propositions are equivalent:*

(i) $A \in \mathbb{S}_n^{++}$,
(ii) A^{-1} *exists and* $A^{-1} \in \mathbb{S}_n^{++}$,

© The Author(s) 2016

S. Burgdorf et al., *Optimization of Polynomials in Non-Commuting Variables*,
SpringerBriefs in Mathematics, DOI 10.1007/978-3-319-33338-0_1

(iii) $\lambda_{\min}(A) > 0$,
(iv) *there exists a non-singular matrix* $B \in \mathbb{R}^{n\times n}$ *such that* $A = B^T B$,
(v) $\det A_k > 0$ *for any leading k-submatrix* $A_k = [a_{i,j}]_{1\leq i,j\leq k}$ *of* A.

The property (iv) of Theorem 1.3 implies the following:

Corollary 1.5. *If* $A \in \mathbb{S}_n^+$*, then for any* $1 \leq i \leq n$ *we have* $a_{i,i} \geq 0$.

This implies, in particular, that the trace of a positive semidefinite matrix is always nonnegative.

Definition 1.6. The (normalized) *trace* of a matrix $A \in \mathbb{S}_n$ is given as the sum of its diagonal entries divided by its order, i.e.,

$$\operatorname{tr} A = \frac{1}{n}\sum_{i=1}^{n} a_{i,i}. \tag{1.1}$$

Proposition 1.7. *Let* $B \in \mathbb{R}^{n\times n}$ *be an arbitrary non-singular matrix. A matrix* $A \in \mathbb{S}_n$ *is positive semidefinite (definite) if and only if the matrix* $B^T AB$ *is also positive semidefinite (definite).*

Proof. For $\mathbf{u} \in \mathbb{R}^n$ and $\mathbf{v} = B^{-1}\mathbf{u}$ we have $\mathbf{u}^T A\mathbf{u} = \mathbf{u}^T B^{-T} B^T ABB^{-1}\mathbf{u} = \mathbf{v}^T B^T AB\mathbf{v}$. Therefore $B^T AB \in \mathbb{S}_n^+$ ($\mathbb{S}_n^{++}$) if and only if $A \in \mathbb{S}_n^+$ ($\mathbb{S}_n^{++}$). ■

Remark 1.8. In the rest of the book we only need the simpler implication of the above proposition. If $A \in \mathbb{S}_n$ is positive semidefinite, then $B^T AB$ is also positive semidefinite for any B.

Lemma 1.9. *Let* $A \in \mathbb{S}_n^+$*. Then:*

(i) *There exists* $i \in \{1,2,\ldots,n\}$ *such that* $a_{i,i} = \max\{|a_{i,j}|\,;\, i,j \in \{1,2,\ldots,n\}\}$.
(ii) *If* $a_{i,i} = 0$ *for some i, then* $A(i,:) = 0$ *and* $A(:,i) = 0$.

Proof. If (i) is not true, then we can find indices $i,j \in \{1,\ldots,n\}$, $i < j$, such that $\max\{a_{i,i}, a_{j,j}\} < |a_{i,j}|$. On the other hand, if (ii) fails, then we can find $i \neq j$ such that $a_{i,i} = 0$ and $a_{i,j} \neq 0$. In both cases the following principal submatrix:

$$\begin{bmatrix} a_{i,i} & a_{i,j} \\ a_{j,i} & a_{j,j} \end{bmatrix}$$

has a negative determinant, contradicting the property (iv) from Theorem 1.3. ■

We can relate the positive semidefiniteness of a block matrix to the positive semidefiniteness of its blocks.

Theorem 1.10 (Schur Complement). *Let* $A \in \mathbb{S}_m^{++}$*,* $C \in \mathbb{S}_n$*, and* $B \in \mathbb{R}^{m\times n}$*. We have*

$$\begin{bmatrix} A & B \\ B^T & C \end{bmatrix} \succeq 0 \text{ if and only if } C - B^T A^{-1} B \succeq 0$$

and

$$\begin{bmatrix} A & B \\ B^T & C \end{bmatrix} \succ 0 \text{ if and only if } C - B^T A^{-1} B \succ 0.$$

Proof. See [HJ12, Theorem 7.7.6] ∎

The following theorem will be used later on when we consider flat matrices.

Proposition 1.11. *Write*

$$\tilde{A} = \begin{bmatrix} A & B \\ B^T & C \end{bmatrix}$$

with $A \in \mathbb{S}_m$, $C \in \mathbb{S}_n$, and $B \in \mathbb{R}^{m \times n}$. Then $\tilde{A} \succeq 0$ if and only if $A \succeq 0$, and there is some Z with

$$B = AZ \quad \textit{and} \quad C \succeq Z^T AZ.$$

Proof. Assume $A \succeq 0$. Given a Z with $B = AZ$ and $C - Z^T AZ \succeq 0$ we get

$$\begin{bmatrix} A & B \\ B^T & C \end{bmatrix} = \begin{bmatrix} I & 0 \\ Z^T & I \end{bmatrix} \begin{bmatrix} A & 0 \\ 0 & C - Z^T AZ \end{bmatrix} \begin{bmatrix} I & Z \\ 0 & I \end{bmatrix}. \tag{1.2}$$

Therefore $\tilde{A} \succeq 0$ by Proposition 1.7, since

$$\begin{bmatrix} I & Z \\ 0 & I \end{bmatrix}$$

is non-singular.

For the converse direction, we use that the columns of B are in the range of A if $\tilde{A} \succeq 0$. Indeed, by the positive semidefiniteness of A, which follows from $\tilde{A} \succeq 0$, we have that $\operatorname{ran} A = (\ker A)^\perp$. So it suffices to show that the columns of B belong to $(\ker A)^\perp$. For this let $\mathbf{x} \in \ker A$, then $\left[\begin{smallmatrix} \mathbf{x} \\ 0 \end{smallmatrix}\right] \in \ker \tilde{A}$ by the positive semidefiniteness of $\tilde{A}$ and thus

$$\begin{bmatrix} A & B \\ B^T & C \end{bmatrix} \begin{bmatrix} \mathbf{x} \\ 0 \end{bmatrix} = 0.$$

Hence $B^T \mathbf{x} = 0$ which implies the statement.

Knowing that columns of B belong to the range of A we find a Z such that $B = AZ$. Using this, the decomposition (1.2) together with $\tilde{A} \succeq 0$ implies that $C - Z^T AZ \succeq 0$. ∎

1.2 Words and Polynomials in Non-commuting Variables

This book is about real polynomials in non-commutative variables. We construct such polynomials in two steps. Starting with a finite alphabet $X_1, \dots, X_n$ (where we fix an integer $n \in \mathbb{N}$) we first generate all possible words in these letters with finite length by concatenating the letters. We add the empty word, denoted by 1. The set of words, obtained this way, is therefore a monoid, freely generated by letters $X_1, \dots, X_n$. We denote it by $\langle \underline{X} \rangle$, where we use notation $\underline{X} := (X_1, \dots, X_n)$. Note that the order of letters in words of $\langle \underline{X} \rangle$ is important, i.e., X_1X_2 is different from X_2X_1. If we consider only two variables, we often denote them as X, Y instead of X_1, X_2.

In the second step we construct all possible finite real linear combinations of words from $\langle \underline{X} \rangle$, which we call real polynomials in non-commutative variables, shorter *nc polynomials*. This set $\mathbb{R}\langle \underline{X} \rangle$ is therefore a free algebra with generating set $\{X_1, \dots, X_n\}$. Hence

$$\mathbb{R}\langle \underline{X} \rangle = \{\sum_{i=1}^{N} a_i w_i \mid N \in \mathbb{N},\ a_i \in \mathbb{R},\ w_i \in \langle \underline{X} \rangle\}.$$

An element of the form $a_w w$ where $a_w \in \mathbb{R} \setminus \{0\}$ and $w \in \langle \underline{X} \rangle$ is called a *monomial* and a_w its *coefficient*. Hence words (elements of $\langle \underline{X} \rangle$) are monomials whose coefficient is 1.

The length of the longest word in an nc polynomial $f \in \mathbb{R}\langle \underline{X} \rangle$ is the *degree* of f and is denoted by $\deg f$. We shall also consider the degree of f in X_i, $\deg_i f$. Similarly, the length of the shortest word appearing in $f \in \mathbb{R}\langle \underline{X} \rangle$ is called the *minimum degree* of f and denoted by $\operatorname{mindeg} f$. Likewise, $\operatorname{mindeg}_i f$ is introduced. If the variable X_i does not occur in some monomial in f, then $\operatorname{mindeg}_i f = 0$. For instance, if $f = X_1^3 + 2X_1X_2X_3 - X_1^2X_4^2$, then

$$\deg f = 4, \quad \deg_1 f = 3, \quad \deg_2 f = \deg_3 f = 1, \quad \deg_4 f = 2,$$

$$\operatorname{mindeg} f = 3, \quad \operatorname{mindeg}_1 f = 1, \quad \operatorname{mindeg}_2 f = \operatorname{mindeg}_3 f = \operatorname{mindeg}_4 f = 0.$$

The set of all words of degree $\leq d$ will be denoted by $\langle \underline{X} \rangle_d$. Likewise we denote nc polynomials of degree $\leq d$ by $\mathbb{R}\langle \underline{X} \rangle_d$. We let $\mathbf{W}_d$ denote the vector of all words of degree $\leq d$ (i.e., $\mathbf{W}_d$ is $\langle \underline{X} \rangle_d$ in lexicographic order). If an nc polynomial f involves only two variables, we use $\mathbb{R}\langle X, Y \rangle$ instead of $\mathbb{R}\langle X_1, X_2 \rangle$.

Remark 1.12. The dimension of $\mathbb{R}\langle \underline{X} \rangle_d$ equals the length of $\mathbf{W}_d$ (containing words in n letters), which is

$$\sigma(n,d) := \sum_{k=0}^{d} n^k = \frac{n^{d+1} - 1}{n - 1}.$$

Thus $\sigma(n,d)$ grows exponentially with the polynomial degree d. Since the number of letters n is usually obvious we simplify notation and use $\sigma(d)$ instead of $\sigma(n,d)$.

We equip $\mathbb{R}\langle \underline{X}\rangle$ with the *involution* $*$ that fixes $\mathbb{R}\cup\{X_1,\ldots,X_n\}$ point-wise and thus reverses words, e.g., $(X_1X_2^2X_3-2X_3^3)^* = X_3X_2^2X_1-2X_3^3$. Hence $\mathbb{R}\langle \underline{X}\rangle$ is the $*$-algebra freely generated by n symmetric letters. The involution extends naturally to matrices (in particular, to vectors) over $\mathbb{R}\langle \underline{X}\rangle$. For instance, if $\mathbf{V}=(v_i)$ is a (column) vector of nc polynomials $v_i\in\mathbb{R}\langle \underline{X}\rangle$, then $\mathbf{V}^*$ is the row vector with components v_i^*. We use $\mathbf{V}^T$ to denote the row vector with components v_i.

Let $\operatorname{Sym}\mathbb{R}\langle \underline{X}\rangle$ denote the set of all *symmetric elements*, that is,

$$\operatorname{Sym}\mathbb{R}\langle \underline{X}\rangle = \{f\in\mathbb{R}\langle \underline{X}\rangle \mid f=f^*\}.$$

Remark 1.13. Occasionally one needs to work with the free $*$-algebra $\mathbb{R}\langle \underline{X},\underline{X}^*\rangle$, i.e., the $*$-algebra freely generated by n (non-symmetric) nc variables $\underline{X}$, or with the mixed case where some of the variables are symmetric and some are not. All of the notions introduced above in the case of symmetric variables have natural counterparts in $\mathbb{R}\langle \underline{X},\underline{X}^*\rangle$. For clarity of exposition, we have restricted ourselves to $\mathbb{R}\langle \underline{X}\rangle$ but most of the results in the book can be easily adapted to $\mathbb{R}\langle \underline{X},\underline{X}^*\rangle$.

1.3 Sums of Hermitian Squares and Gram Matrices

In this section we consider the nc polynomials that are sums of hermitian squares (SOHS) and show how to decide for a given nc polynomial whether such a representation exists.

Definition 1.14. An nc polynomial of the form g^*g is called a *hermitian square*. We say that $f\in\mathbb{R}\langle \underline{X}\rangle$ is a *sum of hermitian squares* (SOHS) if there exist nc polynomials $g_1,\ldots,g_N\in\mathbb{R}\langle \underline{X}\rangle$ for some $N\in\mathbb{N}$ such that $f=\sum_{i=1}^N g_i^*g_i$. The set of SOHS polynomials will be denoted by Σ^2 and the set of SOHS polynomials of degree $\leq 2d$ by Σ^2_{2d}:

$$\Sigma^2 := \Big\{ \sum_{i=1}^N a_i^*a_i \mid N\in\mathbb{N},\ a_i\in\mathbb{R}\langle \underline{X}\rangle \Big\},$$

$$\Sigma^2_{2d} := \Big\{ f\in\Sigma^2 \mid \deg f\leq 2d \Big\}.$$

Clearly, $\Sigma^2,\Sigma^2_{2d}\subsetneq \operatorname{Sym}\mathbb{R}\langle \underline{X}\rangle$.

Example 1.15.

$$XY-YX\notin \operatorname{Sym}\mathbb{R}\langle X,Y\rangle,\quad XYX\in\operatorname{Sym}\mathbb{R}\langle X,Y\rangle\setminus\Sigma^2,$$

$$\begin{aligned}&1-2X+2X^2+XY+YX-X^2Y-YX^2+YX^2Y\\ &=(1-X+XY)^*(1-X+XY)+X^2\in\Sigma^2.\end{aligned}$$

The question whether $f \in \operatorname{Sym}\mathbb{R}\langle \underline{X} \rangle$ is a sum of hermitian squares can be answered through the procedure known as the *Gram matrix method*. The core of the method is given by the following proposition (cf. [Hel02, Sect. 2.2] or [MP05, Theorem 2.1]), the non-commutative version of the classical result due to Choi, Lam, and Reznick ([CLR95, Sect. 2]; see also [Par03, PW98]). The easy proof is included for the sake of completeness.

Proposition 1.16. *Suppose $f \in \operatorname{Sym}\mathbb{R}\langle \underline{X} \rangle_{2d}$. Then $f \in \Sigma^2$ if and only if there exists a positive semidefinite matrix G satisfying*

$$f = \mathbf{W}_d^* G \mathbf{W}_d, \tag{1.3}$$

where $\mathbf{W}_d$ is a vector consisting of all words in $\langle \underline{X} \rangle$ of degree $\leq d$.

Conversely, given such a positive semidefinite matrix G with rank r, one can construct nc polynomials $g_1, \ldots, g_r \in \mathbb{R}\langle \underline{X} \rangle$ of degree $\leq d$ such that

$$f = \sum_{i=1}^{r} g_i^* g_i. \tag{1.4}$$

The matrix G from Proposition 1.16 is called *a Gram matrix* for f.

Proof. If $f = \sum_i g_i^* g_i \in \Sigma^2$, then $\deg g_i \leq d$ for all i as the highest degree terms cannot cancel. Indeed, otherwise by extracting all the appropriate highest degree terms h_i with degree $> d$ from the g_i we would obtain $h_i \in \mathbb{R}\langle \underline{X} \rangle \setminus \{0\}$ satisfying

$$\sum_i h_i^* h_i = 0. \tag{1.5}$$

By substituting symmetric matrices for variables in (1.5), we see that each h_i vanishes for all these substitutions. But then the nonexistence of polynomial identities for arbitrary tuples of symmetric matrices (cf. [Row80, Sects. 2.5 and 1.4]) implies $h_i = 0$ for all i. Contradiction.

Hence we can write $g_i = G_i^T \mathbf{W}_d$, where G_i^T is the (row) vector consisting of the coefficients of g_i. Then $g_i^* g_i = \mathbf{W}_d^* G_i G_i^T \mathbf{W}_d$ and by setting $G := \sum_i G_i G_i^T$, (1.3) clearly holds.

Conversely, given a positive semidefinite $G \in \mathbb{R}^{N \times N}$ of rank r satisfying (1.3), write $G = \sum_{i=1}^r G_i G_i^T$ for $G_i \in \mathbb{R}^{N \times 1}$. Defining $g_i := G_i^T \mathbf{W}_d$ yields (1.4). ■

Proposition 1.16 implies straightforwardly the following corollary:

Corollary 1.17. *Let $f \in \operatorname{Sym}\mathbb{R}\langle \underline{X} \rangle_{2d}$. Then $f \in \Sigma_{2d}^2$ if and only if there exist at most $\sigma(d)$ polynomials $g_i \in \mathbb{R}\langle \underline{X} \rangle_d$ such that $f = \sum_i g_i^* g_i$.*

Example 1.18. Let

$$f = 1 - 2X + X^2 + X^4 + Y^2 + Y^4 - XY^3 + X^3Y + YX^3 - Y^3X + XY^2X + YX^2Y.$$

A Gram matrix for f is given by

$$G = \begin{bmatrix} 1 & -1 & 0 & 0 & 0 & 0 & 0 \\ -1 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & -1 \\ 0 & 0 & 0 & 0 & 0 & -1 & 1 \end{bmatrix},$$

if the word vector is $\mathbf{W}_2 = \begin{bmatrix} 1 \; X \; Y \; X^2 \; XY \; YX \; Y^2 \end{bmatrix}^T$. G is positive semidefinite as is easily seen from its characteristic polynomial or by observing that $G = C^T C$ for

$$C = \begin{bmatrix} 1 & -1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & -1 \end{bmatrix}.$$

From $C\mathbf{W}_2 = \begin{bmatrix} 1-X & Y & X^2+XY & YX-Y^2 \end{bmatrix}^T$ it follows that

$$f = (1-X)^2 + Y^2 + (X^2+XY)^*(X^2+XY) + (YX-Y^2)^*(YX-Y^2) \in \Sigma^2.$$

Note that in this example all monomials from $\mathbf{W}_2$ appear in the SOHS decomposition of f. Another Gram matrix for f is given by

$$\tilde{G} = \begin{bmatrix} 1 & -1 & 0 & \frac{1}{2} & 0 & 0 & 0 \\ -1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ \frac{1}{2} & 0 & 0 & 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & -1 \\ 0 & 0 & 0 & 0 & 0 & -1 & 1 \end{bmatrix}.$$

It is obviously *not* positive semidefinite, since $\tilde{G}_{2,2} = 0$, while $\tilde{G}_{2,1} \neq 0$, contradicting (ii) of Lemma 1.9. Hence it does not give rise to an SOHS decomposition.

Proposition 1.19. *Suppose $h \in \operatorname{Sym}\mathbb{R}\langle \underline{X} \rangle$ is homogeneous of degree $2d$ and let $\mathbf{V}_d$ be a vector consisting of all words in $\langle \underline{X} \rangle$ of degree exactly d. Then*

(i) *h has essentially a unique Gram matrix, i.e., there is a unique symmetric matrix G satisfying*

$$h = \mathbf{V}_d^* G \mathbf{V}_d. \tag{1.6}$$

(ii) *$h \in \Sigma^2$ if and only if G in (1.6) is positive semidefinite.*

Proof. (i) follows from the fact that every word of degree $2d$ can be written *uniquely* as a product of two words of degree d.

For (ii) suppose $h \in \Sigma^2$. In an SOHS decomposition of h we may leave out all monomials of degree $\neq d$ (the lowest, resp. highest degree terms cannot cancel), hence a desired positive semidefinite G exists (cf. proof of Proposition 1.16). The converse is obvious. ■

The sets Σ^2 and Σ^2_{2d} are convex cones. The latter cone is additionally closed in the finite dimensional vector space $\mathbb{R}\langle \underline{X}\rangle_{2d}$, as follows from the following proposition [MP05, Proposition 3.4]:

Proposition 1.20. *Σ^2_{2d} is a closed convex cone in $\mathbb{R}\langle \underline{X}\rangle_{2d}$.*

Proof. This is a variant of the analogous claim in the commutative setting. Endow $\mathbb{R}\langle \underline{X}\rangle_{2d}$ with a norm $\|\_\|$. Recall that $\sigma(d) = \dim \mathbb{R}\langle \underline{X}\rangle_{d}$. By Corollary 1.17 each element $f \in \Sigma^2_{2d}$ can be written as a sum of at most $\sigma(d)$ hermitian squares of degree at most $2d$ since highest degree terms cannot cancel. Hence the image of

$$\varphi : (\mathbb{R}\langle \underline{X}\rangle_{d})^{\sigma(d)} \to \mathbb{R}\langle \underline{X}\rangle_{2d}$$

$$(g_j)_{j=1}^{\sigma(d)} \mapsto \sum_{j=1}^{\sigma} g_j^* g_j$$

equals Σ^2_{2d}. In $(\mathbb{R}\langle \underline{X}\rangle_{d})^{\sigma}$ we define $\mathscr{V} := \{h = (h_i) \mid \|h\| = 1\}$. Note that $\mathscr{V}$ is compact, thus $\varphi(\mathscr{V}) \subseteq \Sigma^2_{2d}$ is compact as well. Since $0 \notin \mathscr{V}$, (there are no dimension-free polynomial identities for tuples of symmetric matrices, cf. [Row80]) we see that $0 \notin \varphi(\mathscr{V})$.

Now, consider a sequence $f_k \in \Sigma^2_{2d}$ $(k \geq 1)$ converging to $f \in \mathbb{R}\langle \underline{X}\rangle_{2d}$. Write $f_k = \lambda_k v_k$ for $\lambda_k \in \mathbb{R}_{\geq 0}$ and $v_k \in \varphi(\mathscr{V})$. Since $\varphi(\mathscr{V})$ is compact, there exists a subsequence of $(v_k)_k$ (again denoted as $(v_k)_k$) converging to $v \in \varphi(\mathscr{V})$. In particular, v is a sum of hermitian squares. As $0 \notin \varphi(\mathscr{V})$, we have $v \neq 0$ and thus we have convergence

$$\lambda_k = \frac{\|f_k\|}{\|v_k\|} \to \frac{\|f\|}{\|v\|} \text{ as } k \to \infty.$$

Thus f_k converges to $f := \frac{\|f\|}{\|v\|} v \in \Sigma^2_{2d}$. ■

1.4 Quadratic Modules and Semialgebraic Sets

In this section we introduce notation and basic results needed later for constrained optimization of nc polynomials. We start by a (usually finite) subset $S \subseteq \operatorname{Sym} \mathbb{R}\langle \underline{X}\rangle$ of symmetric nc polynomials and consider the quadratic module and semialgebraic set associated with S.

Let us shortly recap Carathèodory's theorem [Bar02, p. 10], which will be needed several times later on.

Theorem 1.21 (Carathèodory). *Let $C \subseteq \mathbb{R}^d$. Then every point $\mathbf{c} \in \operatorname{conv}(C) := \{\sum_k \alpha_k \mathbf{c}_k \mid \alpha_k \geq 0, \sum_k \alpha_k = 1\}$ can be represented as a convex combination of $d+1$ points from C.* ■

Definition 1.22. A subset $M \subseteq \operatorname{Sym}\mathbb{R}\langle\underline{X}\rangle$ is called a *quadratic module* if

$$1 \in M, \quad M+M \subseteq M \quad \text{and} \quad a^*Ma \subseteq M \text{ for all } a \in \mathbb{R}\langle\underline{X}\rangle.$$

Given a subset $S \subseteq \operatorname{Sym}\mathbb{R}\langle\underline{X}\rangle$, the quadratic module M_S generated by S is the smallest subset of $\operatorname{Sym}\mathbb{R}\langle\underline{X}\rangle$ containing all a^*sa for $s \in S \cup \{1\}$, $a \in \mathbb{R}\langle\underline{X}\rangle$, and being closed under addition:

$$M_S := \Big\{ \sum_{i=1}^{N} a_i^* s_i a_i \mid N \in \mathbb{N}, s_i \in S \cup \{1\}, a_i \in \mathbb{R}\langle\underline{X}\rangle \Big\}. \tag{1.7}$$

Similarly we define the *truncated quadratic module* of order $2d$ generated by S:

$$M_{S,2d} := \Big\{ \sum_{i=1}^{N} a_i^* s_i a_i \mid N \in \mathbb{N},\ s_i \in S \cup \{1\}, a_i \in \mathbb{R}\langle\underline{X}\rangle, \deg(a_i^* s_i a_i) \le 2d \Big\}.$$

By Theorem 1.21 one gets the uniform bound $N \le 1+\sigma(2d) = 1+\dim\mathbb{R}\langle\underline{X}\rangle_{2d}$.

Note that the sets of sums of hermitian squares Σ^2 and Σ^2_{2d} are special examples of M_S and $M_{S,2d}$, respectively, corresponding to $S = \varnothing$.

Definition 1.23. Fix a finite subset $S \subseteq \operatorname{Sym}\mathbb{R}\langle\underline{X}\rangle$. The *semialgebraic set* $\mathscr{D}_S$ associated with S is the class of tuples $\underline{A} = (A_1, \ldots, A_n) \in \mathbb{S}^n$ of real symmetric matrices of the same order making $s(\underline{A})$ positive semidefinite for every $s \in S$. In case we are considering only tuples of symmetric matrices of fixed order $k \in \mathbb{N}$ we shall use $\mathscr{D}_S(k) := \mathscr{D}_S \cap \mathbb{S}_k^n$.

We can extend this notion to the set of all bounded self-adjoint operators on a (possibly infinite dimensional) Hilbert space making $s(\underline{A})$ a positive semidefinite operator for every $s \in S$. We call this set *operator semialgebraic set* and denote it by $\mathscr{D}_S^\infty$.

Remark 1.24. Clearly, $\mathscr{D}_S \subseteq \mathscr{D}_S^\infty$. On the other hand, there are examples of finite $S \subseteq \operatorname{Sym}\mathbb{R}\langle\underline{X}\rangle$ with

$$\varnothing = \mathscr{D}_S \subsetneq \mathscr{D}_S^\infty.$$

For concrete examples, one can start with finitely presented groups that do not admit finite dimensional representations, and encode the defining relations of such groups. Alternately, employ the generalized Clifford algebras that admit infinite dimensional $*$-representations but no finite dimensional representations, e.g., algebras associated with Brändén's Vamos polynomial [Brä11, NT14].

The following is an obvious but important observation:

Proposition 1.25. *Let $S \subseteq \operatorname{Sym}\mathbb{R}\langle\underline{X}\rangle$. If $f \in M_S$, then $f|_{\mathscr{D}_S} \succeq 0$. Likewise $f|_{\mathscr{D}_S^\infty} \succeq 0$.*

Proof. Indeed, since $f \in M_S$ we have $f = \sum_{i=1}^{N} a_i^* s_i a_i$ for some $N \in \mathbb{N}$, $s_i \in S \cup \{1\}$ and $a_i \in \mathbb{R}\langle \underline{X} \rangle$. For any fixed $\underline{A} \in \mathscr{D}_S$ it follows by definition that $s_i(\underline{A}) \succeq 0$ for all $s_i \in S \cup \{1\}$. Therefore using Remark 1.8 we get

$$a_i^*(\underline{A}) s_i(\underline{A}) a_i(\underline{A}) = a_i(\underline{A})^T s_i(\underline{A}) a_i(\underline{A}) \succeq 0.$$

The proof for the second part is similar. ■

The converse of Proposition 1.25 is false in general, i.e., nonnegativity on an nc semialgebraic set does not imply the existence of a weighted sum of squares certificate. A weak converse holds for *positive* nc polynomials under a strong *boundedness* assumption, see Theorem 1.32 below.

We close this section by giving a last but crucial definition. Archimedeanity of a quadratic module is a condition we shall refer to frequently in this book.

Definition 1.26. A quadratic module M is *archimedean* if

$$\forall a \in \mathbb{R}\langle \underline{X} \rangle \ \exists N \in \mathbb{N} : \ N - a^* a \in M. \tag{1.8}$$

Note if a quadratic module M_S is archimedean, then $\mathscr{D}_S^\infty$ is bounded, i.e., there is an $N \in \mathbb{N}$ such that for every $\underline{A} \in \mathscr{D}_S^\infty$ we have $\|\underline{A}\| \leq N$. The converse is false (cf. [KS07]) in general. However, given $S \subseteq \operatorname{Sym} \mathbb{R}\langle \underline{X} \rangle$ with bounded $\mathscr{D}_S^\infty$, by adding the redundant constraint $g_0 := N - \sum_j X_j^2$ one can ensure the archimedeanity of the quadratic module $M_{S'}$ generated by $S' := S \cup \{g_0\}$ without changing the semialgebraic set $\mathscr{D}_S^\infty = \mathscr{D}_{S'}^\infty$.

1.5 Gelfand–Naimark–Segal's Construction

The Gelfand–Naimark–Segal theorem is a classical theorem in the well-developed theory of C^*-algebras which establishes a correspondence between *-representations of a C^*-algebra and positive linear functionals on it. It is proven by constructing a *-representation out of a positive linear functional. This process is known as the Gelfand–Naimark–Segal (GNS) construction. We will use the same construction for $\mathbb{R}\langle \underline{X} \rangle$ or $\mathbb{R}\langle \underline{X} \rangle_d$, which are not C^*-algebras. Therefore we have to consider a slightly different setup as in the classical GNS theorem, but the technique remains the same. The biggest technical modification is that we need to restrict ourselves to linear functionals on $\mathbb{R}\langle \underline{X} \rangle$ which are positive on an archimedean quadratic module.

Theorem 1.27. *Let $S \subseteq \operatorname{Sym} \mathbb{R}\langle \underline{X} \rangle$ be given such that its quadratic module M_S is archimedean. Let $L : \mathbb{R}\langle \underline{X} \rangle \to \mathbb{R}$ be a (nontrivial) linear functional with $L(M_S) \subseteq \mathbb{R}_{\geq 0}$. Then there exists a tuple $\underline{A} = (A_1, \ldots A_n) \in \mathscr{D}_S^\infty$ and a vector $\mathbf{v}$ such that for all $p \in \mathbb{R}\langle \underline{X} \rangle$:*

$$L(p) = \langle p(\underline{A}) \mathbf{v} \,|\, \mathbf{v} \rangle.$$

Proof. We consider $\mathbb{R}\langle \underline{X}\rangle$ as a vector space, acting on itself by left multiplication. The linear functional L induces the sesquilinear form

$$(p,q) \mapsto L(q^*p) \tag{1.9}$$

on $\mathbb{R}\langle \underline{X}\rangle$. This sesquilinear form is positive semidefinite since L is positive on sums of hermitian squares (cf. Lemma 1.44), hence the Cauchy–Schwarz inequality holds.

Let $\mathscr{N} = \{x \in \mathbb{R}\langle \underline{X}\rangle \mid (x,x) = 0\}$ denote the nullvectors corresponding to L. By the Cauchy–Schwarz inequality this set is a vector subspace of $\mathbb{R}\langle \underline{X}\rangle$. Indeed, for $p,q \in \mathscr{N}$ and $\lambda \in \mathbb{R}$ we have

$$0 \le (p+\lambda q, p+\lambda q) \le (p,p) + (q,q) + 2|\lambda||(p,q)| \le 2|\lambda|\sqrt{(p,p)}\sqrt{(q,q)} = 0.$$

Thus the sesquilinear form (1.9) induces an inner product on the quotient space $\mathbb{R}\langle \underline{X}\rangle/\mathscr{N}$. Set $\mathscr{H}$ to be the Hilbert space completion of $\mathbb{R}\langle \underline{X}\rangle/\mathscr{N}$. Since $1 \notin \mathscr{N}$, $\mathscr{H}$ is nontrivial. Furthermore, $\mathscr{H}$ is separable.

Since L is positive on the archimedean quadratic module M_S, there exists an $N \in \mathbb{N}$ such that $L(p^*(N - X_i^2)p) \ge 0$ for all $i = 1\dots,n$ (using that there exists an $N \in \mathbb{N}$ with $N - X_i^2 \in M_S$). Hence

$$0 \le (X_i p, X_i p) = L(p^* X_i^2 p) \le NL(p^*p) \tag{1.10}$$

This implies that $\mathscr{N}$ is a left ideal. Hence the left multiplication by X_j (i.e., $p \mapsto X_j p$) is well defined on $\mathbb{R}\langle \underline{X}\rangle/\mathscr{N}$ for all $j = 1,\dots n$. It is also bounded as Eq. (1.10) shows as well, and thus extends uniquely to all of $\mathscr{H}$. Now fix an orthonormal basis of $\mathscr{H}$ and let A_j denote the corresponding representative of the left multiplication by X_j in $\mathscr{B}(\mathscr{H})$ with respect to this basis. Since

$$(X_j p, p) = L((X_j p)^* p) = L(p^*(X_j p)) = (p, X_j p)$$

holds true for all $p \in \mathbb{R}\langle \underline{X}\rangle/\mathscr{N}$ and $\mathbb{R}\langle \underline{X}\rangle/\mathscr{N}$ is dense in $\mathscr{H}$, the operators A_j are self-adjoint, that is, $A_j^* = A_j$.

Now

$$L(p) = (p,1) = \langle p(\underline{A})\mathbf{v} \,|\, \mathbf{v}\rangle, \tag{1.11}$$

where $\mathbf{v}$ denotes the vector in $\mathscr{H}$ corresponding to the identity polynomial 1. We claim that $\underline{A} \in \mathscr{D}_S^\infty$. Let $g \in S$. By the density of $\mathbb{R}\langle \underline{X}\rangle/\mathscr{N}$ in $\mathscr{H}$ any vector $\mathbf{u} \in \mathscr{H}$ can be approximated to arbitrary precision by elements of $\mathbb{R}\langle \underline{X}\rangle/\mathscr{N}$. Hence it is sufficient to show that $\langle g(\underline{A})\mathbf{u} \,|\, \mathbf{u}\rangle \ge 0$ where we consider $\mathbf{u}$ as vector representative of $u \in \mathbb{R}\langle \underline{X}\rangle/\mathscr{N}$. By construction, for any such $\mathbf{u}$ exists a polynomial $p \in \mathbb{R}\langle \underline{X}\rangle$ such that $\mathbf{u} = p(\underline{A})\mathbf{v}$. Now

$$\begin{aligned}\langle g(\underline{A})\mathbf{u} \,|\, \mathbf{u}\rangle &= \langle g(\underline{A})p(\underline{A})\mathbf{v} \,|\, p(\underline{A})\mathbf{v}\rangle \\ &= \langle (gp)(\underline{A})\mathbf{v} \,|\, p(\underline{A})\mathbf{v}\rangle = L(p^*gp).\end{aligned}$$

This proves the claim since $p^*gp \in M_S$ and thus by assumption

$$L(p^*gp) \geq 0.$$

■

Remark 1.28. The general Gelfand–Naimark–Segal construction can be simplified if we have finite dimensional vector spaces since then the completion is not needed to obtain a Hilbert space. In this case, we can also replace the condition of positivity of L on the quadratic module, which has been used to obtain bounded operators, by the assumption that L is strictly positive on sums of hermitian squares. More concretely, we will use later on the following statement:

Let $L : \mathbb{R}\langle \underline{X} \rangle_{2d} \to \mathbb{R}$ be a linear functional strictly positive on $\Sigma^2_{2d} \setminus \{0\}$. Then there exists a tuple of matrices $\underline{A}$ of order at most $\sigma(2d)$ and a vector $\mathbf{v} \in \mathbb{R}^{\sigma(2d)}$ such that L is given by $L(p) = \langle p(\underline{A})\mathbf{v} \,|\, \mathbf{v} \rangle$ for $p \in \mathbb{R}\langle \underline{X} \rangle_{2d-2}$.

Proof. By our assumption, the linear functional L induces a (positive definite) inner product $(p,q) \mapsto L(q^*p)$ on the Hilbert space $\mathscr{H} = \mathbb{R}\langle \underline{X} \rangle_{2d}$. We define the left multiplication by X_j for $j = 1, \ldots, n$ by

$$\begin{aligned} \hat{X}_j : \mathbb{R}\langle \underline{X} \rangle_{2d} &\to \mathbb{R}\langle \underline{X} \rangle_{2d} \\ p &\mapsto \pi(X_j p), \end{aligned}$$

where $\pi : \mathbb{R}\langle \underline{X} \rangle_{2d+2} \to \mathbb{R}\langle \underline{X} \rangle_{2d}$ is the canonical projection of $\mathbb{R}\langle \underline{X} \rangle_{2d+2}$ to $\mathbb{R}\langle \underline{X} \rangle_{2d}$. Remark that $\hat{X}_j$ is defined as $\hat{X}_j(p) = X_j p$ as in the classical GNS construction if $p \in \mathbb{R}\langle \underline{X} \rangle_{2d-2}$. These multiplication maps are well defined and also symmetric. Letting A_j be the corresponding representative in $\mathscr{B}(\mathscr{H}) = \mathbb{R}^{\sigma(2d)\times\sigma(2d)}$, and $\mathbf{v} \in \mathbb{R}^{\sigma(2d)}$ be the vector corresponding to the identity polynomial 1, we get the desired representation of L for polynomials in $\mathbb{R}\langle \underline{X} \rangle_{2d-2}$. ■

1.6 Sums of Hermitian Squares and Positivity

A matrix is positive semidefinite if and only if it is a hermitian square (see Theorem 1.3). A similar statement holds true in the case of nc polynomials. It provides the theoretical underpinning for reformulation of the eigenvalue optimization of an nc polynomial into a semidefinite programming (SDP) problem.

Definition 1.29. An nc polynomial $f \in \mathbb{R}\langle \underline{X} \rangle$ is *positive semidefinite* if $f(A_1, \ldots, A_n)$ is positive semidefinite for all tuples $\underline{A} = (A_1, \ldots, A_n) \in \mathbb{S}^n$ of real symmetric matrices of the same order. Likewise we define positive semidefinite polynomials over a semialgebraic set $\mathscr{D}_S$.

The important characterization of positive semidefinite polynomials as exactly the sums of hermitian squares of polynomials has been given by Helton [Hel02], see also [McC01] or [MP05] for different proofs.

Theorem 1.30. *For $f \in \mathbb{R}\langle \underline{X} \rangle$ we have $f(\underline{A}) \succeq 0$ for all $\underline{A} \in \mathbb{S}^n$ if and only if f is a sum of hermitian squares.*

Proof. If f is a sum of hermitian squares, then obviously for all $\underline{A} \in \mathbb{S}^n$ the matrix $f(\underline{A}) \in \mathbb{S}^n$ is a sum of hermitian squares of matrices and thus positive semidefinite.

For the converse implication, assume $f \notin \Sigma^2$. Let $\deg f = 2d - 2$, then $f \notin \Sigma^2_{2d}$. Now, since Σ^2_{2d} is closed, we can apply the Minkowski separation theorem [Bar02, Theorem 1.3] on the finite dimensional vector space $\mathbb{R}\langle \underline{X} \rangle_{2d}$ to obtain a strictly separating linear functional, i.e., there exists a linear functional $L : \mathbb{R}\langle \underline{X} \rangle_{2d} \to \mathbb{R}$ such that $L(\Sigma^2_{2d}) \geq 0$ and $L(f) < 0$. By small changes (if needed) we can guarantee that L is strictly positive on $\Sigma^2_{2d} \setminus \{0\}$. The Gelfand–Naimark–Segal construction (see Remark 1.28) yields then a tuple of matrices $\underline{A} \in \mathbb{S}^n_{\sigma(2d)}$ and a vector $\mathbf{v}$ such that $L(f) = \langle f(\underline{A})\mathbf{v} \,|\, \mathbf{v} \rangle \geq 0$ contradicting $L(f) < 0$. ■

Remark 1.31. It suffices to consider Σ^2_{2d} instead of Σ^2 when $\deg f = 2d$ since highest degree terms do not cancel. This has been shown in the proof of Proposition 1.16.

For the constrained setting where one considers positive semidefiniteness of a polynomial over a given semialgebraic set $\mathscr{D}^\infty_S$ there is the following perfect generalization [HM04, Theorem 1.2] of Putinar's Positivstellensatz [Put93] for commutative polynomials.

Theorem 1.32. *Let $S \cup \{f\} \subseteq \operatorname{Sym} \mathbb{R}\langle \underline{X} \rangle$ and suppose that M_S is archimedean. If $f(\underline{A}) \succ 0$ for all $\underline{A} \in \mathscr{D}^\infty_S$, then $f \in M_S$.*

Proof. Since the argument is standard and resembles the proof of Theorem 1.30, we only present a sketch. Assume that $f \notin M_S$. By archimedeanity of M_S, there is a linear functional $L : \operatorname{Sym} \mathbb{R}\langle \underline{X} \rangle \to \mathbb{R}$ with $L(f) < 0$ and $L(M_S) \subseteq \mathbb{R}_{\geq 0}$. The GNS construction (Theorem 1.27) then yields a tuple $\underline{A} = (A_1 \ldots, A_n) \in \mathscr{D}^\infty_S$ of bounded self-adjoint operators and a vector $\mathbf{v}$ such that $0 > L(f) = \langle f(\underline{A})\mathbf{v} \,|\, \mathbf{v} \rangle$, contradicting the positive definiteness of f. ■

Remark 1.33. In general it does not suffice to test for positive definiteness of f on $\mathscr{D}_S$ (as opposed to $\mathscr{D}^\infty_S$) in Theorem 1.32; cf. Remark 1.24 above. However, if $\mathscr{D}_S$ is convex [HM04, Sect. 2], then it is by [HM12] a linear matrix inequality (LMI) domain $\mathscr{D}_L$. In this case every polynomial positive semidefinite on $\mathscr{D}_L$ admits a weighted sum of squares certificate with optimal degree bounds [HKM12].

1.7 Vanishing Nc Polynomials

We present some facts about vanishing nc polynomials which are needed later on. The following results are basically consequences of the standard theory of polynomial identities, cf. [Row80]. They all essentially boil down to the well-known fact that there are no nonzero polynomial identities that hold for all orders of (symmetric) matrices. In fact, it is enough to test on an ε-neighborhood of 0.

Definition 1.34. For $\varepsilon > 0$ we introduce

$$\begin{aligned}\mathscr{N}_\varepsilon &= \bigcup_{k\in\mathbb{N}} \Big\{ \underline{A} = (A_1,\dots,A_n) \in \mathbb{S}_k^n \mid \varepsilon^2 - \sum_{i=1}^n A_i^2 \succeq 0 \Big\} \\ &= \bigcup_{k\in\mathbb{N}} \Big\{ \underline{A} = (A_1,\dots,A_n) \in \mathbb{S}_k^n \mid \Big\| \big[A_1 \cdots A_n \big]^T \Big\| \le \varepsilon \Big\},\end{aligned} \tag{1.12}$$

the *nc* ε*-neighborhood of* 0. (Unless mentioned otherwise, all our norms are assumed to be operator norms, i.e., $\|A\| = \sup\{\|Ax\| \mid \|x\| = 1\}$.) We will also refer in the sequel to $\mathscr{N}_\varepsilon(N) = \mathbb{S}_N^n \bigcap \mathscr{N}_\varepsilon$.

Lemma 1.35. *If $f \in \mathbb{R}\langle \underline{X} \rangle$ is zero on $\mathscr{N}_\varepsilon$ for some $\varepsilon > 0$, then $f = 0$.*

Proof. This follows from the following: an nc polynomial of degree $< 2d$ that vanishes on all n-tuples of symmetric matrices $\underline{A} \in \mathscr{N}_\varepsilon(N)$, for some $N \ge d$, is zero (using the standard multilinearization trick and, e.g., [Row80, Sects. 2.5 and 1.4]). ■

Lemma 1.36. *Suppose $f \in \mathbb{R}\langle \underline{X} \rangle$ and let $\varepsilon > 0$. If $f(\underline{A})$ is singular for all $\underline{A} \in \mathscr{N}_\varepsilon$, then $f = 0$.*

Proof. Let $\underline{A} \in \mathbb{S}_k^n$ for some $k \in \mathbb{N}$ be arbitrary. Then $p(t) = \det f(t\underline{A})$ is a real polynomial in t. By assumption it vanishes on all small enough $t > 0$. Hence $p = 0$ as every polynomial of finite degree in one real variable has only finitely many zeros. This implies that $f(\underline{A})$ is singular for all $k \in \mathbb{N}$ and all $\underline{A} \in \mathbb{S}_k^n$.

Now consider the ring $GM_{2^\ell}(n)$ of n symmetric $2^\ell \times 2^\ell$ generic matrices. It is a PI ring and a domain, so admits a skew field of fractions $UD_{2^\ell}(n)$ [Pro76, PS76]. However, by the Cayley–Hamilton theorem, the image $\check{f}$ of f in $UD_{2^\ell}(n)$ is a zero divisor, so $\check{f} = 0$, i.e., f is a polynomial identity for symmetric $2^\ell \times 2^\ell$ matrices. Since ℓ was arbitrary, this yields $f = 0$. ■

In our subsequent analysis, we will need to deal with neighborhoods of non-scalar points $\underline{A}$. Given $\underline{A} \in \mathbb{S}_k^n$, let

$$\mathscr{B}(\underline{A}, \varepsilon) = \bigcup_{\ell\in\mathbb{N}} \{ \underline{B} \in \mathbb{S}_{k\ell}^n \mid \|\underline{B} - I_\ell \otimes \underline{A}\| \le \varepsilon \}$$

denote the *nc neighborhood* of $\underline{A}$. These are used to define topologies in free analysis [KVV14].

Proposition 1.37. *Suppose $f \in \mathbb{R}\langle \underline{X} \rangle$, $\varepsilon > 0$, and let $\underline{A} \in \mathbb{S}_{2^k}^n$. If $f(\underline{B})$ is singular for all $\ell \in \mathbb{N}$ and all $\underline{B} \in \mathscr{B}(\underline{A}, \varepsilon)(2^{k+\ell})$, then $f = 0$.*

Proof. For $\ell \in \mathbb{N}$ and $\underline{B} \in \mathbb{S}_{2^{k+\ell}}^n$ consider the univariate polynomial $\Phi_{\underline{B}}$ defined by

$$t \mapsto \det f(I_{2^\ell} \otimes \underline{A} + t\underline{B}).$$

By assumption, $\Phi_{\underline{B}}$ vanishes for all t of small absolute value. Hence by analyticity it vanishes everywhere. We can now proceed as in the proof of Lemma 1.36 to deduce f is a polynomial identity for symmetric matrices of all orders, whence $f = 0$. ■

The following technical proposition, which provides the closedness of a quadratic module M_S, is a variant of a Powers–Scheiderer result [PS01, Sect. 2].

Proposition 1.38. *Suppose* $S = \{g_1, \dots, g_t\} \subseteq \operatorname{Sym}\mathbb{R}\langle \underline{X} \rangle$ *is such that* $\mathscr{D}_S$ *contains an* ε*-neighborhood of* 0. *Then* $M_{S,2d}$ *is a closed convex cone in the finite dimensional real vector space* $\operatorname{Sym}\mathbb{R}\langle \underline{X} \rangle_{2d}$.

For the proof of this proposition we need to isolate a (possibly) non-scalar point and its neighborhood where all the g_j are positive definite:

Lemma 1.39. *Suppose* $0 \notin S = \{g_1, \dots, g_t\} \subseteq \operatorname{Sym}\mathbb{R}\langle \underline{X} \rangle$ *is such that* $\mathscr{D}_S$ *contains an* ε*-neighborhood of* 0. *Then there is an* $\underline{A} \in \mathbb{S}_{2^k}^n$ *and* $\bar{\varepsilon} > 0$ *such that all* g_j *are positive definite on* $\mathscr{B}(\underline{A}, \bar{\varepsilon})$.

Proof. By Proposition 1.37, we find $\delta_1 > 0$ and $\underline{A}^1 \in \mathscr{N}_{\delta_1}(2^{k_1})$ such that $g_1(\underline{A}^1) \succ 0$. Then there is an $\varepsilon_1 > 0$ such that $g_1(\underline{B}) \succ 0$ for all $\underline{B} \in \mathscr{B}(\underline{A}^1, \varepsilon_1)$.

Now g_2 is not singular everywhere on $\mathscr{B}(\underline{A}^1, \varepsilon_1)$ by Proposition 1.37. Hence we find $\underline{A}^2 \in \mathscr{B}(\underline{A}^1, \varepsilon_1)(2^{k_2})$ with $g_2(\underline{A}^2) \succ 0$, and a corresponding $\varepsilon_2 > 0$ with $g_2|_{\mathscr{B}(\underline{A}^2, \varepsilon_2)} \succ 0$. Without loss of generality, $\mathscr{B}(\underline{A}^2, \varepsilon_2) \subseteq \mathscr{B}(\underline{A}^1, \varepsilon_1)$. We repeat this procedure for $g_3, \dots, g_t$. Finally, setting $\underline{A} = \underline{A}^r$, $\bar{\varepsilon} = \varepsilon_r$ yields the desired conclusion. ■

Proof (Proposition 1.38). By Lemma 1.39, we find an $\bar{\varepsilon} > 0$ and $\underline{A} \in \mathbb{S}_k^n$ such that $g_j(\underline{B}) \succ 0$ for all j and all $\underline{B} \in \mathscr{B}(\underline{A}, \bar{\varepsilon})$. Using $\mathscr{B}(\underline{A}, \bar{\varepsilon})$ we norm $\mathbb{R}\langle \underline{X} \rangle_{2d}$ by

$$|||p||| := \sup\left\{ \|p(\underline{B})\| \mid \underline{B} \in \mathscr{B}(\underline{A}, \bar{\varepsilon}) \right\}.$$

Let $\delta > 0$ be a lower bound on all the $g_j(\underline{B})$ for $\underline{B} \in \mathscr{B}(\underline{A}, \bar{\varepsilon})$, i.e., $g_j(\underline{B}) - \delta I \succeq 0$ for all $\underline{B} \in \mathscr{B}(\underline{A}, \bar{\varepsilon})$.

Now the proof of the proposition follows a standard argument, and is essentially a consequence of Theorem 1.21. Suppose now $(p_m)_m$ is a sequence from $M_{S,2d}$ which converges to some $p \in \mathbb{R}\langle \underline{X} \rangle$ of degree at most $2d$. By Theorem 1.21 there is an ℓ (at most the dimension of $\mathbb{R}\langle \underline{X} \rangle_{2d}$ plus one) such that for each m there exist nc polynomials $r_{m,i} \in \mathbb{R}\langle \underline{X} \rangle_d$ and $t_{m,i,j} \in \mathbb{R}\langle \underline{X} \rangle_d$ such that

$$p_m = \sum_{i=1}^{\ell} r_{m,i}^* r_{m,i} + \sum_{j=1}^{r} \sum_{i=1}^{\ell} t_{m,i,j}^* g_i t_{m,i,j}.$$

Since $|||p_m||| \leq N^2$ for some $N > 0$, it follows that $|||r_{m,i}||| \leq N$ and likewise we get $\left|\left|\left| t_{m,i,j}^* g_j t_{m,i,j} \right|\right|\right| \leq N^2$. In view of the choice of ε, δ, we obtain $\left|\left|\left| t_{m,i,j} \right|\right|\right| \leq \frac{1}{\sqrt{\delta}} N$ for all i, m, j. Hence for each i, j, the sequences $(r_{m,i})$ and $(t_{m,i,j})$ are bounded in m. They thus have convergent subsequences. Tracking down these subsequential limits finishes the proof. ■

Proposition 1.38 allows us to deduce the following separation result:

Corollary 1.40. *Let $S \subseteq \operatorname{Sym}\mathbb{R}\langle \underline{X} \rangle$. Assume $\mathscr{D}_S$ contains an ε-neighborhood of* 0*, and $f \in \operatorname{Sym}\mathbb{R}\langle \underline{X} \rangle_{2d} \setminus M_{S,2d}$. Then there exists a linear functional $L : \mathbb{R}\langle \underline{X} \rangle_{2d} \to \mathbb{R}$ which is nonnegative on $M_{S,2d}$, strictly positive on nonzero elements of Σ^2_{2d}, and with $L(f) < 0$.*

Proof. The existence of a separating linear functional L follows from Proposition 1.38 and the Minkowski separation theorem [Bar02, Theorem 1.3]. If necessary, add a small multiple of a linear functional strictly positive on $\Sigma^2_d \setminus \{0\}$, and the proof is complete. ■

1.8 Hankel Matrices and Flatness

Following standard definitions and notation [BV04] the dual cone of Σ^2_{2d} is defined by

$$(\Sigma^2_{2d})^{\vee} := \{L : \mathbb{R}\langle \underline{X} \rangle_{2d} \to \mathbb{R} \mid L \text{ linear},\ L(f) = L(f^*),\ L(f) \geq 0\ \forall f \in \Sigma^2_{2d}\}.$$

In this section we will present an alternative representation of linear functionals in $(\Sigma^2_{2d})^{\vee}$ by positive semidefinite matrices, and define the concept of flatness for them.

Definition 1.41. Given $g \in \operatorname{Sym}\mathbb{R}\langle \underline{X} \rangle$ and a linear functional $L : \mathbb{R}\langle \underline{X} \rangle_{2d} \to \mathbb{R}$ we can associate to L two matrices:

(i) An *nc Hankel matrix* H_L, indexed by words $u, v \in \langle \underline{X} \rangle_d$, with

$$(H_L)_{u,v} = L(u^* v);$$

(ii) A *localizing matrix* $H^{\Uparrow}_{L,g}$ indexed by words $u, v \in \langle \underline{X} \rangle_{d - \lceil \deg(g)/2 \rceil}$ with

$$(H^{\Uparrow}_{L,g})_{u,v} = L(u^* g v).$$

We say that L is *unital* if $L(1) = 1$, and we call L *symmetric* if $L(f^*) = L(f)$ for all f in the domain of L.

Hankel matrices are closely related to the Hankel condition.

Definition 1.42. A matrix H indexed by words of length $\leq d$ satisfies the *nc Hankel condition* if and only if

$$H_{u_1,v_1} = H_{u_2,v_2} \quad \text{whenever} \quad u_1^* v_1 = u_2^* v_2. \tag{1.13}$$

Remark 1.43. Symmetric linear functionals on $\mathbb{R}\langle \underline{X} \rangle_{2d}$ and matrices from $\mathbb{S}_{\sigma(d)}$ satisfying the nc Hankel condition are in bijective correspondence. To each symmetric $L\colon \mathbb{R}\langle \underline{X}\rangle_{2d} \to \mathbb{R}$ we can assign $H_L \in \mathbb{S}_{\sigma(d)}$ defined by $(H_L)_{u,v} = L(u^*v)$ and vice versa. The following lemma relates positivity of L and positive semidefinitness of its Hankel matrix H_L.

Lemma 1.44. *Let $g \in \operatorname{Sym}\mathbb{R}\langle \underline{X}\rangle$ and let $L : \mathbb{R}\langle \underline{X}\rangle_{2d} \to \mathbb{R}$ be a symmetric linear functional. Then the following holds:*

(i) *$L(p^*p) \geq 0$ for all $p \in \mathbb{R}\langle \underline{X}\rangle_d$ (i.e., L is positive) if and only if $H_L \succeq 0$.*
(ii) *$L(p^*gp) \geq 0$ for all $p \in \mathbb{R}\langle \underline{X}\rangle_{d-\lceil \deg(g)/2\rceil}$ if and only if $H^{\Uparrow}_{L,g} \succeq 0$.*

Proof. For $p = \sum_w p_w w \in \mathbb{R}\langle \underline{X}\rangle_d$ let $\mathbf{p} \in \mathbb{R}^\sigma$ be the vector consisting of all coefficients p_w of p. Then the first statement follows from

$$L(p^*q) = \sum_{u,v} p_u q_v L(u^*v) = \sum_{u,v} p_u q_v (H_L)_{u,v} = \mathbf{p}^T H_L \mathbf{q}.$$

The second statement follows similarly by verifying

$$L(p^*gq) = \mathbf{p}^T H^{\Uparrow}_{L,g}\mathbf{q}.$$

■

If we summarize Proposition 1.16 and Lemma 1.44, we obtain the following identifications.

Corollary 1.45.

$$\Sigma^2_{2d} \rightleftharpoons \{G \in \mathbb{S}_{\sigma(d)} \mid G \succeq 0\}$$
$$(\Sigma^2_{2d})^\vee \rightleftharpoons \{H \in \mathbb{S}_{\sigma(d)} \mid H_{u_1,v_1} = H_{u_2,v_2} \ \textit{whenever}\ u_1^*v_1 = u_2^*v_2,\ H \succeq 0\}.$$

We mention that these identifications are not isomorphisms since, e.g., a sum of hermitian squares usually has more than one positive semidefinite Gram matrix G representing it.

Hankel matrices which satisfy the so-called flatness condition will play a crucial role later on.

Definition 1.46. Let $A \in \mathbb{R}^{s\times s}$ be a symmetric matrix. An *extension* of A is a symmetric matrix $\tilde{A} \in \mathbb{R}^{(s+\Delta)\times(s+\Delta)}$ of the form

$$\tilde{A} = \begin{bmatrix} A & B \\ B^T & C \end{bmatrix}$$

for some $B \in \mathbb{R}^{s\times\Delta}$ and $C \in \mathbb{R}^{\Delta\times\Delta}$.

By Proposition 1.11, $\tilde{A} \succeq 0$ if and only if $A \succeq 0$, and there is some Z with

$$B = AZ \quad \text{and} \quad C \succeq Z^T AZ. \tag{1.14}$$

This characterization is useful to define what we understand as a flat matrix.

Definition 1.47. An extension $\tilde{A}$ of A is *flat* if $\operatorname{rank} A = \operatorname{rank} \tilde{A}$, or, equivalently, if $B = AZ$ and $C = Z^T AZ$ for some matrix Z.

For a comprehensive study of flatness in functional analysis we refer the reader to [CF96, CF98]. We shall only need some basic properties for flat matrix extensions, see, for instance, [CF96, Lemma 5.2] or [CF98, Proposition 2.1].

Lemma 1.48. *Let $\tilde{A}$ be a flat extension of A as in Definition 1.47. Then the following statements hold.*

(i) $\ker \tilde{A} = \ker [A\ B]$*;*

(ii) $\mathbf{x} \in \ker A \implies [\mathbf{x}\ 0]^T \in \ker \tilde{A}$*;*

(iii) $A \succeq 0$ *if and only if* $\tilde{A} \succeq 0$.

Proof. We have $\operatorname{rank} \tilde{A} \geq \operatorname{rank} [A\ B] \geq \operatorname{rank} A$. Since $\operatorname{rank} A = \operatorname{rank} \tilde{A}$, equality holds, which implies (1.48). To show (1.48) let $\mathbf{x} \in \ker A$. Since $\tilde{A}$ is a flat extension of A there is a matrix Z such that $B = AZ$. Hence we have $B^T \mathbf{x} = 0$, which implies $[\mathbf{x}\ 0]^T \in \ker \tilde{A}$. For the last statement let $Z \in \mathbb{R}^{s \times \Delta}$ be given such that $B = AZ$ and $C = Z^T AZ$. Let $\mathbf{v} = \begin{bmatrix}\mathbf{a}\ \mathbf{b}\end{bmatrix}^T \in \mathbb{R}^{s+\Delta}$ with $\mathbf{a} \in \mathbb{R}^s$ and $\mathbf{b} \in \mathbb{R}^\Delta$ be given. Then one easily verifies that

$$\mathbf{v}^T \tilde{A} \mathbf{v} = (\mathbf{a} + Z\mathbf{b})^T A (\mathbf{a} + Z\mathbf{b}),$$

which implies (1.48). $\blacksquare$

If $\tilde{A} \succeq 0$ we can express its deviation from flatness, using the Frobenius norm $\|\text{␣}\|_F$, by

$$\mathtt{err_{flat}} := \frac{\|C - Z^T AZ\|_F}{1 + \|C\|_F + \|Z^T AZ\|_F}. \tag{1.15}$$

Here Z is as in (1.14); it is easy to see that $\mathtt{err_{flat}}$ is independent of the choice of Z.

Using the definition of flat matrices we define when a linear functional is considered to be flat.

Definition 1.49. Suppose $L : \mathbb{R}\langle \underline{X} \rangle_{2d+2\delta} \to \mathbb{R}$ is a linear functional with restriction $\check{L} : \mathbb{R}\langle \underline{X} \rangle_{2d} \to \mathbb{R}$. We associate to L and $\check{L}$ the Hankel matrices H_L and $H_{\check{L}}$, respectively, and get the block form

$$H_L = \begin{bmatrix} H_{\check{L}} & B \\ B^T & C \end{bmatrix}.$$

We say that L is *δ-flat*, or that L is a *flat extension* of $\check{L}$, if H_L is flat over $H_{\check{L}}$.

1.9 Commutators, Cyclic Equivalence, and Trace Zero Polynomials

We proceed with our analysis of vanishing nc polynomials. But this time we consider nc polynomials with vanishing trace. This will be useful in Chap. 5 where we perform trace optimization.

It is well known and easy to see that trace zero matrices are (sums of) commutators. To mimic this property for nc polynomials, we introduce cyclic equivalence [KS08a]:

Definition 1.50. An element of the form $[p,q] := pq - qp$ for $p,q \in \mathbb{R}\langle \underline{X} \rangle$ is called a *commutator*. Two nc polynomials $f,g \in \mathbb{R}\langle \underline{X} \rangle$ are called *cyclically equivalent* ($f \overset{\text{cyc}}{\sim} g$) if $f-g$ is a sum of commutators:

$$f - g = \sum_{i=1}^{k} [p_i, q_i] = \sum_{i=1}^{k} (p_i q_i - q_i p_i) \text{ for some } k \in \mathbb{N} \text{ and } p_i, q_i \in \mathbb{R}\langle \underline{X} \rangle.$$

It is clear that $\overset{\text{cyc}}{\sim}$ is an equivalence relation. The following proposition shows that it can be easily tested and motivates its name.

Proposition 1.51.

(i) *For $v, w \in \langle \underline{X} \rangle$, we have $v \overset{\text{cyc}}{\sim} w$ if and only if there are $v_1, v_2 \in \langle \underline{X} \rangle$ such that $v = v_1 v_2$ and $w = v_2 v_1$, i.e., if and only if w is a cyclic permutation of v.*

(ii) *Nc polynomials $f = \sum_w a_w w$ and $g = \sum_w b_w w$ (with $a_w, b_w \in \mathbb{R}$) are cyclically equivalent if and only if for each $v \in \langle \underline{X} \rangle$,*

$$\sum_{\substack{w \in \langle \underline{X} \rangle \\ w \overset{\text{cyc}}{\sim} v}} a_w = \sum_{\substack{w \in \langle \underline{X} \rangle \\ w \overset{\text{cyc}}{\sim} v}} b_w. \tag{1.16}$$

Proof. For the first case, $v \overset{\text{cyc}}{\sim} w$ if and only if $v - w$ is a commutator. Then the statement is obvious. For the second statement, by linearity of the commutator, we can split up any sum of commutators of polynomials into a sum of commutators of words. Hence we can split up $f-g$ into groups of cyclically equivalent words. Then the statement follows by comparing coefficients. ■

This notion is important for us because symmetric trace zero nc polynomials are exactly sums of commutators [KS08a, Theorem 2.1]. In other words, they are cyclically equivalent to the zero polynomial.

Definition 1.52. An nc polynomial $p \in \operatorname{Sym} \mathbb{R}\langle \underline{X} \rangle$ is a *trace zero* nc polynomial if

$$\operatorname{tr} p(\underline{A}) = 0 \text{ for all } \underline{A} \in \mathbb{S}^n.$$

As in Lemma 1.35 it suffices to test for vanishing trace on an ε-neighborhood $\mathcal{N}_\varepsilon$ of 0 (see Definition 1.34); see also [BK09] for an alternative proof.

Lemma 1.53. *If $f \in \operatorname{Sym}\mathbb{R}\langle \underline{X}\rangle$ has zero trace on $\mathscr{N}_\varepsilon$ for some $\varepsilon > 0$, then f is a sum of commutators, i.e., $f \overset{\text{cyc}}{\sim} 0$.*

Remark 1.54. The assumption that f is symmetric is crucial for the lemma. Indeed, the polynomial $X_1X_2X_3 - X_3X_2X_1$ is a trace zero polynomial but not a sum of commutators.

We can get an even stronger result for polynomials in a quadratic module.

Lemma 1.55. *Let $S = \{g_1, \dots, g_t\} \subseteq \operatorname{Sym}\mathbb{R}\langle \underline{X}\rangle$. Assume that the semialgebraic set $\mathscr{D}_S$ contains an ε-neighborhood of 0, and*

$$\sum_j h_j^* h_j + \sum_{i,j} r_{ij}^* g_i r_{ij} \overset{\text{cyc}}{\sim} 0. \tag{1.17}$$

Then $h_j = r_{ij} = 0$ for all i,j.

Proof. Let $\underline{A}$, ε be such that $g_i \succ 0$ on $\mathscr{B}(\underline{A}, \varepsilon)$ for all $i = 1, \dots, r$, see Lemma 1.39. For each $\underline{B} \in \mathscr{B}(\underline{A}, \varepsilon)$ we have

$$\sum_j \operatorname{tr} h_j(\underline{B})^* h_j(\underline{B}) + \sum_{i,j} \operatorname{tr} r_{ij}(\underline{B})^* g_i(\underline{B}) r_{ij}(\underline{B}) = 0$$

by (1.17). Hence $h_j(\underline{B}) = r_{ij}(\underline{B}) = 0$. Now apply Proposition 1.37. ■

1.10 Cyclic Quadratic Modules and Trace-Positivity

Using cyclic equivalence one can define a tracial variant of a quadratic module associated with a semialgebraic set. Let $S \subseteq \operatorname{Sym}\mathbb{R}\langle \underline{X}\rangle$ with corresponding quadratic module M_S and its truncated variant $M_{S,2d}$, see Sect. 1.4 for definitions.

Definition 1.56. We set

$$\begin{aligned} \Theta^2_{S,2d} &= \{f \in \operatorname{Sym}\mathbb{R}\langle \underline{X}\rangle \mid \exists g \in M_{S,2d} : f \overset{\text{cyc}}{\sim} g\} \\ \Theta^2_S &= \bigcup_{d\in\mathbb{N}} \Theta^2_{S,2d}, \end{aligned}$$

and call Θ^2_S the *cyclic quadratic module* generated by S, and $\Theta^2_{S,2d}$ the *truncated cyclic quadratic module* generated by S.

When $S = \{g_1, \dots, g_t\}$ is finite, every element f of $\Theta^2_{S,2d}$ is cyclically equivalent to an element of the form

$$\sum_{k=1}^{N} a_k^* a_k + \sum_{i=1}^{r}\sum_{j=1}^{N_i} b_{ij}^* g_i b_{ij} \in M_{S,2d} \tag{1.18}$$

for some $a_k, b_{ij} \in \mathbb{R}\langle \underline{X}\rangle$ with $\deg a_k \leq d$ and $\deg(b_{ij}^* g_i b_{ij}) \leq 2d$.

By Theorem 1.21 one gets again uniform bounds for the number of polynomials needed: $N, N_i \leq 1 + \sigma(2d) = 1 + \dim \mathbb{R}\langle \underline{X} \rangle_{2d}$.

Definition 1.57. If $S = \varnothing$, we call Θ_S^2 (shortly denoted by Θ^2) the cone of nc polynomials cyclically equivalent to a sum of hermitian squares.

The cyclic quadratic module $\Theta_{S,2d}^2$ is closed as $M_{S,2d}$ is closed.

Proposition 1.58. *Suppose $S = \{g_1, \ldots, g_t\} \subseteq \operatorname{Sym} \mathbb{R}\langle \underline{X} \rangle$ and assume $\mathscr{D}_S$ contains an ε-neighborhood of 0. Then $\Theta_{S,2d}^2$ is a closed convex cone in the finite dimensional real vector space $\mathbb{R}\langle \underline{X} \rangle_{2d}$. In particular, if $f \in \operatorname{Sym} \mathbb{R}\langle \underline{X} \rangle_{2d} \setminus \Theta_{S,2d}^2$, then there exists a tracial linear functional $L : \mathbb{R}\langle \underline{X} \rangle_{2d} \to \mathbb{R}$ which is nonnegative on $\Theta_{S,2d}^2$, positive on $\Sigma_{2d}^2 \setminus \{0\}$ with $L(f) < 0$.*

Proof. With Lemma 1.55 at hand, the proof of this corollary is the same as that of Proposition 1.38 so is omitted. ∎

Elements in the cyclic quadratic module model polynomials which are positive with respect to the trace. To facilitate our research of the trace, we need to consider a distinguished subset of $\mathscr{D}_S^\infty$ (see Definition 1.23) obtained by restricting our attention from the algebra of all bounded operators $B(\mathscr{H})$ on a Hilbert space $\mathscr{H}$ (which does not admit a trace if $\mathscr{H}$ is infinite dimensional) to finite von Neumann algebras [Tak03]. One usually distinguishes between finite von Neumann algebras of type I and type II, the first being full matrix algebras and the second its infinite dimensional analog.

Definition 1.59. Let $\mathscr{F}$ be a type II_1-von Neumann algebra [Tak03, Chap. 5], and let $\mathscr{D}_S^{\mathscr{F}}$ be the *$\mathscr{F}$-semialgebraic set* generated by S; that is, $\mathscr{D}_S^{\mathscr{F}}$ consists of all tuples $\underline{A} = (A_1, \ldots, A_n) \in \mathscr{F}^n$ making $s(\underline{A})$ a positive semidefinite operator for every $s \in S$. Then

$$\mathscr{D}_S^{\mathrm{II}_1} := \bigcup_{\mathscr{F}} \mathscr{D}_S^{\mathscr{F}},$$

where the union is over all type II_1-von Neumann algebras $\mathscr{F}$ with separable predual, is called the *von Neumann (vN) semialgebraic set* generated by S.

As for the positive semidefinite case we define what we mean by a positive polynomial when considering the trace.

Definition 1.60. We say an nc polynomial $f \in \mathbb{R}\langle \underline{X} \rangle$ is *trace-positive* if

$$\operatorname{tr} f(\underline{A}) \geq 0 \text{ for all tuples } \underline{A} = (A_1, \ldots, A_n) \in \mathbb{S}^n \tag{1.19}$$

of real symmetric matrices of the same order. We call it trace-positive on $\mathscr{D}_S$ if $\operatorname{tr} f(\underline{A}) \geq 0$ for all tuples $\underline{A} \in \mathscr{D}_S$; and similarly for $\mathscr{D}_S^{\mathrm{II}_1}$, where the matricial trace as in (1.1) is replaced by the canonical trace-function in the corresponding finite von Neumann algebra (which we will also denote by tr), see [Tak03].

Remark 1.61. There are inclusions

$$\mathscr{D}_S \subseteq \mathscr{D}_S^{\mathrm{II}_1} \subseteq \mathscr{D}_S^{\infty}; \tag{1.20}$$

here the first is obtained via embedding matrix algebras in the hyperfinite II_1-factor $\mathscr{R}$, and for the second inclusion simply consider a separable II_1-factor as a subalgebra of $B(\mathscr{H})$.

Whether the first inclusion in (1.20) is "dense" in the sense that a polynomial $f \in \mathbb{R}\langle \underline{X} \rangle$ is trace-positive on $\mathscr{D}_S$ iff f is trace-positive on $\mathscr{D}_S^{\mathrm{II}_1}$ is closely related to Connes' embedding conjecture [Con76, KS08a], a deep and important open problem in operator algebras. To sidestep this problem, we shall focus on values of nc polynomials on $\mathscr{D}_S^{\mathrm{II}_1}$ instead of $\mathscr{D}_S$.

As for the positive semidefinite case we have the immediate observation:

Proposition 1.62. *Let* $S \subseteq \operatorname{Sym}\mathbb{R}\langle \underline{X} \rangle$*. If* $f \in \Theta_S^2$*, then* $\operatorname{tr} f(\underline{A}) \geq 0$ *for* $\underline{A} \in \mathscr{D}_S$*. Likewise* $\operatorname{tr} f(\underline{A}) \geq 0$ *for* $\underline{A} \in \mathscr{D}_S^{\mathrm{II}_1}$.

However, unlike in the positive semidefinite case, there are trace-positive polynomials which are *not* members of Θ^2. Surprisingly, the situation is in perfect analogy to non-homogeneous polynomials in commuting variables. Besides the univariate case, trace-positive quadratics and bivariate quartics (i.e., polynomials in two variables of degree four) are always sums of hermitian squares and commutators [BK10], but in all other cases this is not true any more. The easiest example for a trace-positive nc polynomials which is not a member of Θ^2 is the non-commutative Motzkin polynomial [KS08a, Example 4.4]

$$X_1 X_2^4 X_1 + X_2 X_1^4 X_2 - 3X_1 X_2^2 X_1 + 1. \tag{1.21}$$

An example in three variables is the nc polynomial

$$X_1^2 X_2^2 + X_1^2 X_3^2 + X_2^2 X_3^2 - 4X_1 X_2 X_3,$$

see [Qua15, Theorem 3.4]. We also refer the reader to [KS08b, Example 3.5] for more sophisticated examples (of homogeneous polynomials) obtained by considering the BMV conjecture.

Nevertheless, the obvious certificate for trace-positivity—being a sum of hermitian squares and commutators—turns out to be useful in optimization. We also have a tracial version of Theorem 1.32. It provides the theoretical underpinning for the tracial version of Lasserre's relaxation scheme (presented in Sect. 5.3 below) used to minimize the trace of an nc polynomial.

Proposition 1.63. *Let* $S \cup \{f\} \subseteq \operatorname{Sym}\mathbb{R}\langle \underline{X} \rangle$ *and suppose that* M_S *is archimedean. Then the following are equivalent:*

(i) $\operatorname{tr} f(\underline{A}) \geq 0$ *for all* $\underline{A} \in \mathscr{D}_S^{\mathrm{II}_1}$*;*

(ii) *for all* $\varepsilon > 0$ *there exists* $g \in M_S$ *with* $f + \varepsilon \overset{\mathrm{cyc}}{\sim} g$.

Proof. Since the argument is standard, we only present a sketch of the proof. The implication (ii) $\Rightarrow$ (i) is obvious. For the converse, assume $\varepsilon > 0$ is such that the conclusion of (ii) fails. By archimedeanity of M_S, there is a tracial linear functional $L : \operatorname{Sym}\mathbb{R}\langle \underline{X} \rangle \to \mathbb{R}$ with $L(f+\varepsilon) \leq 0$, $L(M_S) \subseteq \mathbb{R}_{\geq 0}$. The Gelfand–Naimark–Segal construction as in Theorem 1.27 yields then bounded self-adjoint operators A_j. The double commutant of $\mathscr{A}$, the algebra generated by the bounded operators A_j, is then a finite von Neumann algebra with trace tr. But by the GNS construction we have $\operatorname{tr} f(\underline{A}) = L(f) \leq -\varepsilon < 0$, contradicting (i). ■

Note that assumption (i) implies trace-positivity of f on the hyperfinite II_1-factor $\mathscr{R}$ [Tak03] and hence on all finite type I von Neumann algebras, i.e., on all full matrix algebras over $\mathbb{R}$.

Remark 1.64. The dual cone $(\Theta_{2d}^2)^\vee$ consists of symmetric linear functionals which are nonnegative on Σ^2 and on commutators. We can associate with each linear functional $L \in (\Theta_{2d}^2)^\vee$ its Hankel matrix H_L, which is positive semidefinite and is given by $(H_L)_{u,v} = L(u^* v)$ for indexing words $u, v \in \langle \underline{X} \rangle_d$, exactly as for linear functionals $L \in (\Sigma_{2d}^2)^\vee$. Since L is nonnegative (actually it is zero) on commutators, the Hankel matrix H_L is invariant under cyclic equivalence, i.e., $(H_L)_{u,v} = (H_L)_{w,z}$ whenever $u^* v \overset{\text{cyc}}{\sim} w^* z$ for $u, v, w, z \in \langle \underline{X} \rangle_d$. In this case we call L *tracial*, and H_L is a *tracial Hankel matrix*.

1.11 Wedderburn Theorem

The Wedderburn theorem classifies simple finite dimensional k-algebras as matrix algebras over a division ring, see [Lam01, Chap. 1] for more details and proofs.

Definition 1.65. A ring R is simple if its only ideals are $\{0\}$ and R itself.

Theorem 1.66 (Wedderburn). *Let R be a simple finite dimensional k-algebra. Then $R \cong D^{n \times n}$ for some division ring D, where D is unique up to isomorphism.* ■

Furthermore, the division ring D is also a division algebra which implies that $D = k$ if k is algebraically closed. Theorem 1.66 is also useful in combination with the following theorem of Frobenius [Lam01, (13.12)].

Theorem 1.67 (Frobenius). *Any finite dimensional associative division algebra over $\mathbb{R}$ is isomorphic to either $\mathbb{R}, \mathbb{C}$ or $\mathbb{H}$, where $\mathbb{H}$ denotes the quaternions.* ■

This immediately leads to the fact that the only central simple algebras over $\mathbb{R}$ are the full matrix algebras over $\mathbb{R}$, $\mathbb{C}$, or $\mathbb{H}$. This can be used to characterize positive linear functionals on $*$-subalgebras of $\mathbb{R}^{s \times s}$ which are zero on commutators.

Proposition 1.68. *Let $\mathscr{A}$ be a $*$-subalgebra of $\mathbb{R}^{s \times s}$ for some $s \in \mathbb{N}$ and $L : \mathscr{A} \to \mathbb{R}$ be a positive linear functional with $L(pq - qp) = 0$ for all $p, q \in \mathscr{A}$. Then there exist full matrix algebras $\mathscr{A}^{(i)}$ over $\mathbb{R}$, $\mathbb{C}$, or $\mathbb{H}$, a $*$-isomorphism*

$$\mathscr{A} \to \bigoplus_{i=1}^{N} \mathscr{A}^{(i)}, \tag{1.22}$$

and $\lambda_1, \dots, \lambda_N \in \mathbb{R}_{\geq 0}$ *with* $\sum_i \lambda_i = 1$*, such that for all* $A \in \mathscr{A}$*,*

$$L(A) = \sum_{i=1}^{N} \lambda_i \operatorname{tr} A^{(i)},$$

where the $A^{(i)}$ *come from* $\bigoplus_i A^{(i)}$*, the image of* A *under the isomorphism* (1.22)*. The order of* (*the real representation of*) $\bigoplus_i A^{(i)}$ *is at most* s*.*

Proof. By orthogonal transformation one can derive that $\mathscr{A}$ has block diagonal form as in (1.22). Each of the blocks $\mathscr{A}^{(i)}$ acts irreducibly on a subspace of $\mathbb{R}^s$ and is thus a central simple algebra (with involution) over $\mathbb{R}$. Knowing that the $\mathscr{A}^{(i)}$ are full matrix algebras over $\mathbb{R}$, $\mathbb{C}$, or $\mathbb{H}$ one can use Galois theory to derive that the only possibility for $L|_{\mathscr{A}^{(i)}}$ is the matricial trace. ■

1.12 Curto–Fialkow's Theorems

Curto and Fialkow studied necessity and sufficiency conditions for the classical truncated moment problem, i.e., which linear functionals on $\mathbb{R}[\underline{x}]_{2d}$ can be expressed by integration over a positive measure. One main condition which guarantees such a moment representation is flatness of the corresponding Hankel matrix. We will present non-commutative versions of their theorems in this section. The proofs will use several results presented previously in this chapter, namely the Gelfand–Naimark–Segal construction and the Wedderburn theorem. For the definition of flatness we refer to Definition 1.49.

Theorem 1.69. *Let* $S = \{g_1, \dots, g_t\} \subseteq \operatorname{Sym}\mathbb{R}\langle \underline{X} \rangle$ *and set* $\delta = \max\{\lceil \deg(g_i)/2 \rceil, 1\}$*. Let* $L : \mathbb{R}\langle \underline{X} \rangle_{2d+2\delta} \to \mathbb{R}$ *be a unital linear functional satisfying* $L(M_{S,2d+2\delta}) \subseteq \mathbb{R}_{\geq 0}$*. If* L *is* δ*-flat, then there exist* $\underline{A} \in \mathscr{D}_S(r)$ *for some* $r \leq \sigma(d)$ *and a unit vector* $\mathbf{v}$ *such that*

$$L(f) = \langle f(\underline{A})\mathbf{v} \,|\, \mathbf{v} \rangle.$$

Remark 1.70. The 1 in the definition of δ is only taken into account if S is the empty set. As we will see in the proof, r can be chosen to be the rank of the Hankel matrix H_L of L.

Proof. The representation of the δ-flat linear functional follows from the finite dimensional variant of the Gelfand–Naimark–Segal construction. Let $\operatorname{rank} H_L = r$. Since the nc Hankel matrix H_L is positive semidefinite, we can find a Gram decomposition $H_L = [\langle \mathbf{u} \,|\, \mathbf{w} \rangle]_{u,w}$ with vectors $\mathbf{u}, \mathbf{w} \in \mathbb{R}^r$, where the labels are words of degree at most $d + \delta$. Using this decomposition we set

$$\mathscr{H} = \operatorname{span}\{\mathbf{w} \mid \deg w \leq d + \delta\}.$$

By the flatness assumption one gets that

$$\mathscr{H} = \operatorname{span}\{\mathbf{w} \mid \deg w \leq d+\delta\} = \operatorname{span}\{\mathbf{w} \mid \deg w \leq d\}. \tag{1.23}$$

On $\mathscr{H}$ we now perform the Gelfand–Naimark–Segal construction. The linear functional L defines an inner product on $\mathscr{H}$ via

$$(\mathbf{p},\mathbf{q}) \mapsto L(q^*p),$$

thus $\mathscr{H}$ is a finite dimensional Hilbert space. Hence we can directly consider the operators A_i representing the left multiplication by X_i on $\mathscr{H}$, i.e., $X_i\mathbf{w} = \mathbf{X_i w}$. Since by Eq. (1.23) we only need to consider words w with $\deg w \leq d$, the resulting word X_iw is of degree $d+1 \leq d+\delta$. Hence the A_i are well defined. The remaining part of the proof follows the proof of Theorem 1.27. With a similar line of reasoning one derives that the A_i are symmetric, $\underline{A} = (A_1,\dots,A_n) \in \mathscr{D}_S(r)$, and that with $\mathbf{v}$ being the vector representing 1 one gets the desired representation

$$L(f) = \langle f(\underline{A})\mathbf{v} \mid \mathbf{v}\rangle.$$

■

In the tracial case we also need to use the Wedderburn theorem as an additional ingredient in the proof. Recall that a linear functional $L : \mathbb{R}\langle\underline{X}\rangle \to \mathbb{R}$ is tracial if L is 0 on commutators.

Theorem 1.71. *Let $S = \{g_1,\dots,g_t\} \subseteq \operatorname{Sym}\mathbb{R}\langle\underline{X}\rangle$ and set $\delta = \max\{\lceil \deg(g_i)/2\rceil, 1\}$. Let $L : \mathbb{R}\langle\underline{X}\rangle_{2d+2\delta} \to \mathbb{R}$ be an unital tracial linear functional with $L(\Theta_S^2) \subseteq \mathbb{R}_{\geq 0}$. If L is δ-flat, then there are finitely many n-tuples $\underline{A}^{(j)}$ of symmetric matrices in $\mathscr{D}_S(N)$ for some $N < 4\sigma(d)$ and positive scalars $\lambda_j > 0$ with $\sum_j \lambda_j = 1$ such that for all $p \in \mathbb{R}\langle\underline{X}\rangle_{2d}$:*

$$L(p) = \sum_j \lambda_j \operatorname{tr} p(\underline{A}^{(j)}).$$

Again, we can replace the bound $4\sigma(d)$ by $4r$, where r is the rank of the tracial Hankel matrix of L, which is in turn bounded by $\sigma(d)$.

Proof. As in the previous theorem we perform the finite dimensional Gelfand–Naimark–Segal construction resulting in a tuple $\underline{A} = (A_1,\dots,A_n) \in \mathscr{D}_S(\sigma(d))$ and a unit vector $\mathbf{v}$ such that

$$L(p) = \langle p(\underline{A})\mathbf{v} \mid \mathbf{v}\rangle. \tag{1.24}$$

To get a tracial representation let $\mathscr{A}$ be the subalgebra generated by the symmetric matrices A_j. Since the hermitian square of a nonzero matrix is not nilpotent, $\mathscr{A}$ is semisimple. By Proposition 1.68, which is a consequence of the Wedderburn theorem 1.66, $\mathscr{A}$ can be (orthogonally) block diagonalized into

$$\mathscr{A} = \oplus_{j=1}^k \mathscr{A}_j. \tag{1.25}$$

where the $\mathscr{A}_i$ are simple full matrix algebras over $\mathbb{R}$, $\mathbb{C}$, or $\mathbb{H}$. With respect to the decomposition (1.25), $A_j = \oplus_{\ell=1}^k A_j^\ell$. Each A_j^ℓ is a symmetric matrix, and the tuple $\underline{A}^\ell \in \mathscr{D}_S$. Without loss of generality each A_j^ℓ is a real matrix; if one of the blocks $\mathscr{A}_j$ is a matrix algebra over $\mathbb{C}$ or $\mathbb{H}$, we embed it into the real matrix algebra (of twice the order for $\mathbb{C}$ and four times the order for $\mathbb{H}$).

By (1.24) we can consider L to be a tracial linear functional on $\mathscr{A}$. Then L induces tracial $\mathbb{R}$–linear functionals L_j on the simple $*$-algebras $\mathscr{A}_j$. If $p(\underline{A}) = \underline{B} = \oplus_{j=1}^k \underline{B}^j$, then

$$L(p) = \sum_j L_j(\underline{B}^j).$$

Each L_j is a positive multiple, say λ_j, of the usual trace [BK12, Lemma 3.11]. Thus

$$L(p) = \sum_j L_j(\underline{B}^j) = \sum_j \lambda_j \mathrm{tr}\, p(\underline{A}^j).$$

Since $L(1) = 1$, $\sum_j \lambda_j = 1$. ■

Implementation

We summarize the procedures described in the proofs of Theorems 1.27 and 1.69 into an algorithm that we call GNS construction.

Remark 1.72. Algorithm 1.1 returns exactly the n tuple of matrices and the vector from Theorem 1.69. Note that the matrices $\bar{A}_i$ represent linear mappings $w \mapsto X_i w$ in the basis $\mathscr{C}$, while A_i are these mappings in the standard orthonormal basis of $\mathbb{R}^r$.

The main ingredients of the proof of Theorem 1.71 are summarized in Algorithm 1.2.

Algorithm 1.1: GNS construction

```
Input: $H_L$, Hankel matrix of $L$ satisfying assumptions of Theorem 1.69
Find $\mathscr{C} = [\mathbf{w}_1, \ldots, \mathbf{w}_r]$ - matrix with linearly independent columns of $H_L$, corresponding to
    words $w_i$ with $\deg w_i \leq d$. Take $w_1 = 1$;
Let $H_{\hat{L}}$ be the principal submatrix of $H_L$ consisting of columns and rows corresponding to
    words $w_1, \ldots, w_r$;
Compute by Cholesky factorization $G$ such that $G^T G = H_{\hat{L}}$;
for $i = 1, \ldots, n$ do
    Let $\mathscr{C}_i = [\mathbf{X_i w_1}, \ldots \mathbf{X_i w_r}]$;
    Compute $\bar{A}_i$ as solution of the system $\mathscr{C}\bar{A}_i = \mathscr{C}_i$;
    Let $A_i = G\bar{A}_i G^{-1}$;
end
Compute $\mathbf{v} = G\mathbf{e}_1$;
Output: $(A_1, \ldots, A_n)$, $\mathbf{v}$;
```

Algorithm 1.2: GNS-Wedderburn construction

```
Input: $L$ satisfying assumptions of Theorem 1.71
1 Compute $\underline{A} = (A_1, \ldots, A_n)$ and $\mathbf{v}$ by Algorithm 1.1;
2 Let $\mathscr{A}$ be the algebra generated by $A_i$, $i = 1, \ldots, n$. Compute orthogonal $Q$
  such that $Q^T \mathscr{A} Q = \{\mathrm{Diag}(B_1, \ldots, B_k) \mid B_i \in \mathscr{A}_i\}$, where $\mathscr{A}_i$ are simple
  algebras over $\mathbb{R}$. Let $\hat{A}_i = Q^T A_i Q = \mathrm{Diag}(\hat{A}_i^1, \ldots, \hat{A}_i^k)$. Then
  $L(p) = \langle p(\underline{A})\mathbf{v}, \mathbf{v}\rangle = \langle p(\underline{\hat{A}})Q^T\mathbf{v}, Q^T\mathbf{v}\rangle = \sum_{j=1}^k \langle p(\underline{\hat{A}}^j)\hat{\mathbf{v}}^j, \hat{\mathbf{v}}^j\rangle$ where $\hat{\mathbf{v}}^j$ is the $j$th
  part of $Q^T\mathbf{v}$;
3 $L$ induces $L_j$ on $\mathscr{A}_j$ by $L_j(B_j) = \langle p(\underline{\hat{A}}_j)\hat{\mathbf{v}}_j, \hat{\mathbf{v}}_j\rangle$, where $B_j = p(\underline{\hat{A}}^j)$;
4 $L_j$ is tracial with $L_j(B_j) = \lambda_j \mathrm{tr}\,(B_j)$, where $\lambda_j = \mathrm{tr}\,(I_j) = \|\hat{\mathbf{v}}_j\|^2$ (note that
  $\|\hat{\mathbf{v}}\|^2 = 1$);
Output: $(\hat{A}_1, \ldots, \hat{A}_n)$, $(\lambda_1, \ldots, \lambda_k)$.
```

Remark 1.73. Note that we can compute the matrix Q in Step 2 by Algorithm 4.1 from [MKKK10]. The implementation of the GNS-Wedderburn construction in `NCSOStools` is indeed based on this algorithm.

1.13 Semidefinite Programming

SDP is a subfield of convex optimization concerned with the optimization of a linear objective function over the intersection of the cone of positive semidefinite matrices with an affine space. More precisely, given symmetric matrices $C, A_1, \ldots, A_m$ of the same order over $\mathbb{R}$ and a vector $\mathbf{b} \in \mathbb{R}^m$, we formulate a *semidefinite program in standard primal form* [in the sequel we refer to problems of this type by (PSDP)] as follows:

$$\begin{aligned} \inf \ & \langle C \mid G \rangle \\ \text{s.t.} \ & \langle A_i \mid G \rangle = b_i, \ i = 1, \ldots, m \\ & G \succeq 0. \end{aligned} \tag{PSDP}$$

Here $\langle \textvisiblespace \mid \textvisiblespace \rangle$ stands for the standard scalar product of matrices: $\langle A \mid B \rangle = \mathrm{tr}\, B^T A$. The dual problem to (PSDP) is the following *semidefinite program in standard dual form*

$$\begin{aligned} \sup \ & \langle \mathbf{b} \mid \mathbf{y} \rangle \\ \text{s.t.} \ & \textstyle\sum_i y_i A_i \preceq C. \end{aligned} \tag{DSDP}$$

Here $\mathbf{y} \in \mathbb{R}^m$ and the difference $C - \sum_i y_i A_i$ is usually denoted by Z. If $C = 0$, then (PSDP) is a SDP feasibility problem (in standard primal form):

$$\begin{aligned} & G \succeq 0, \\ \text{s.t.} \ & \langle A_i, G \rangle = b_i, \ i = 1, \ldots, m. \end{aligned} \tag{FSDP}$$

The primal–dual pair of SDP problems (PSDP)–(DSDP) are strongly related. These relations are collected in the duality theory. We state some of the most well-known results. They can be found, e.g., in [dK02, Hel00].

Theorem 1.74 (Weak Duality). *Let G be a feasible solution for* (DSDP) *and* $(\mathbf{y}, Z)$ *a feasible solution for* (DSDP)*. We have*

$$\mathbf{b}^T\mathbf{y} \leq \langle C \,|\, G \rangle$$

with equality holding if and only if $\langle G \,|\, Z \rangle = 0$.

Proof. This theorem shows that every feasible solution for (PSDP) gives an upper bound for the optimal value of (DSDP), and similarly every feasible solution for the (DSDP) yields a lower bound for the optimal value of (PSDP). We know that G and $Z = C - \sum_i y_i A_i$ are positive semidefinite, therefore $\langle G \,|\, Z \rangle \geq 0$. This implies that:

$$\begin{aligned} \langle C \,|\, G \rangle - \mathbf{b}^T\mathbf{y} &= \langle C \,|\, G \rangle - \sum_i \langle A_i \,|\, G \rangle y_i \\ &= \langle C - \sum_i y_i A_i \,|\, G \rangle = \langle Z \,|\, G \rangle \geq 0. \end{aligned}$$

■

We call the nonnegative quantity $\langle C \,|\, G \rangle - \mathbf{b}^T\mathbf{y}$ a *duality gap*. If we have primal and dual feasible solutions G and $(\mathbf{y}, Z)$ with zero duality gap, then obviously these solutions are optimal. The converse is not always true. It holds for the case of linear programming, but for SDP and more general conic programming problems we need additional assumptions for this conclusion to hold.

Let OPT_P and OPT_D be the optimal values of (PSDP) and (DSDP), respectively. From the weak duality theorem we know that $\mathrm{OPT}_P - \mathrm{OPT}_D \geq 0$. We call this difference the *optimal duality gap*. We say that (PSDP) and (DSDP) have the *strong duality property* if the optimal duality gap is zero. There are some sufficient conditions for the strong duality property that we are going to present.

Definition 1.75 (Strict Feasibility). The SDP (PSDP) is *strictly feasible* if there exists $G \succ 0$ such that $\langle A_i \,|\, G \rangle = b_i$, $\forall i$. The SDP (DSDP) is *strictly feasible* if there exist $\mathbf{y} \in \mathbb{R}^m$ and $Z \succ 0$ such that $\sum_i y_i A_i + Z = C$.

The strict feasibility condition is also known as the *Slater condition*.

Theorem 1.76 (Strong Duality). *If the primal problem* (PSDP) *is strictly feasible, we have either*

(i) *an infeasible dual problem* (DSDP) *if the primal problem* (PSDP) *is unbounded, i.e.,* $\mathrm{OPT}_P = \mathrm{OPT}_D = -\infty$, *or*
(ii) *a feasible dual problem* (DSDP) *if the primal problem* (PSDP) *is bounded. In this case the dual optimal value* OPT_D *is finite, attained and* $\mathrm{OPT}_D = \mathrm{OPT}_P$.

Proof. See [Hel00, Theorem 2.2.5] or [dK02, Theorem 2.2]. ■

Corollary 1.77. *If both the primal problem* (PSDP) *and the dual problem* (DSDP) *are strictly feasible, then we have a zero optimal duality gap and both optimal values are attained.*

Corollary 1.78. *If both the primal problem* (PSDP) *and the dual problem* (DSDP) *are strictly feasible, then the equations*

$$\left.\begin{array}{r}\langle A_i \,|\, G\rangle = b_i, \quad G \succeq 0\\ \sum_i y_i A_i + Z = C, \quad Z \succeq 0\\ \langle X \,|\, G\rangle = 0 \qquad\qquad\end{array}\right\} \tag{1.26}$$

are necessary and sufficient optimality conditions for the (PSDP) *and* (DSDP)*, i.e., a triple* $(G, \mathbf{y}, Z) \in \mathbb{S}_n^+ \times \mathbb{R}^m \times \mathbb{S}_n^+$ *is primal and dual optimal if and only if it satisfies (1.26).*

Most of the methods used to solve SDP problems actually solve (1.26) iteratively using different first or second order methods.

SDP is serving in algebraic geometry mainly as a tool to extract certificates that a given polynomial belongs to a set under consideration. These certificates are typically numerical since they are extracted from the (numerical) optimal solutions of the related SDPs. Often numerical certificates are not sufficient, and one needs to extract a *proof* for a given statement (e.g., an nc polynomial is or is not in Θ^2). Therefore in the rest of this section a particular emphasis is given to the extraction of *rational* certificates if the input data is rational.

Consider a feasibility SDP in primal form (FSDP) and assume the input data A_i, b_i is rational for $i = 1, \ldots, m$. If the problem is feasible, does there exist a *rational* solution? If so, can one use a combination of numerical and symbolic computation to produce one?

Example 1.79. Some caution is necessary, as a feasible SDP of the form (FSDP) needs not admit a rational solution. For a simple concrete example, note that

$$\begin{bmatrix} 2 & x \\ x & 1 \end{bmatrix} \oplus \begin{bmatrix} x & 1 & 0 \\ 1 & x & 1 \\ 0 & 1 & x \end{bmatrix} \succeq 0 \quad \Leftrightarrow \quad x = \sqrt{2}.$$

In fact there are commutative polynomials with rational coefficients that are sums of squares of polynomials over the reals, but not over the rationals (see [Sch12]). Adapting an example of Scheiderer, we obtain an nc polynomial with rational coefficients that is cyclically equivalent to a sum of hermitian squares of nc polynomials over the reals, but not over the rationals:

$$f = 1 + X^3 + X^4 - \frac{3}{2}XY - \frac{3}{2}YX - 4XYX + 2Y^2 + Y^3 + \frac{1}{2}XY^3 + \frac{1}{2}Y^3X + Y^4.$$

This is a dehomogenized and symmetrized non-commutative version of the (commutative) polynomial from [Sch12, Theorem 2.1] (setting $x_0 = 1$, $x_1 = X$ and $x_2 = Y$). So f is not cyclically equivalent to a sum of hermitian squares with rational coefficients. By [Sch12, Theorem 2.1], $f|_{\mathbb{R}^2} \geq 0$. Together with the fact that f is cyclically sorted, [KS08a, Proposition 4.2] implies that f is trace-positive. Since f is of degree 4 in two variables it is a sum of hermitian squares with commutators [BK12, BCKP13] (with real coefficients).

On the other hand, if (FSDP) admits a feasible *positive definite* solution, then it admits a (positive definite) *rational* feasible solution. More exactly, we have the following:

Theorem 1.80 (Peyrl and Parrilo [PP08]). *If an approximate feasible point* G_0 *for* (FSDP) *satisfies*

$$\delta := \lambda_{\min}(G_0) > \|(\langle A_i, G_0\rangle - b_i)_i\| =: \varepsilon, \tag{1.27}$$

then a (*positive definite*) *rational feasible point* G *exists. It can be obtained from* G_0 *in the following two steps* (*cf. Fig. 1.1*)*:*

(1) *compute a rational approximation* $\tilde{G}$ *of* G_0 *with* $\tau := \|\tilde{G} - G_0\|$ *satisfying*

$$\tau^2 + \varepsilon^2 < \delta^2;$$

(2) *project* $\tilde{G}$ *onto the affine subspace* $\mathscr{L}$ *given by the equations* $\langle A_i, G\rangle = b_i$ *to obtain* G*.*

Note that the results in [PP08] are stated for SDPs arising from sum of squares problems, but their results carry over verbatim to the setting of (the seemingly more) general SDPs. The rationalization scheme based on this Peyrl–Parrilo technique has been implemented in `NCSOStools`; see Example 3.25 for a demonstration.

Not all is lost, however, if the SDP solver gives a *singular* feasible point G_0 for (FSDP). Suppose that $\mathbf{z}$ is a *rational* nullvector for G_0. Let P be a change of basis matrix containing $\mathbf{z}$ as a first column and a (rational) orthogonal basis for the orthogonal complement $\{\mathbf{z}\}^\perp$ as its remaining columns. Then

$$P^T G_0 P = \begin{bmatrix} 0 & 0 \\ 0 & \hat{G}_0 \end{bmatrix},$$

i.e.,

$$G_0 = P^{-T} \begin{bmatrix} 0 & 0 \\ 0 & \hat{G}_0 \end{bmatrix} P^{-1}$$

for some symmetric $\hat{G}_0$. Hence

$$b_i = \langle A_i, G_0\rangle = \operatorname{tr}(A_i G_0) = \operatorname{tr}\left(A_i P^{-T} \begin{bmatrix} 0 & 0 \\ 0 & \hat{G}_0 \end{bmatrix} P^{-1}\right) = \operatorname{tr}\left(P^{-1} A_i P^{-T} \begin{bmatrix} 0 & 0 \\ 0 & \hat{G}_0 \end{bmatrix}\right).$$

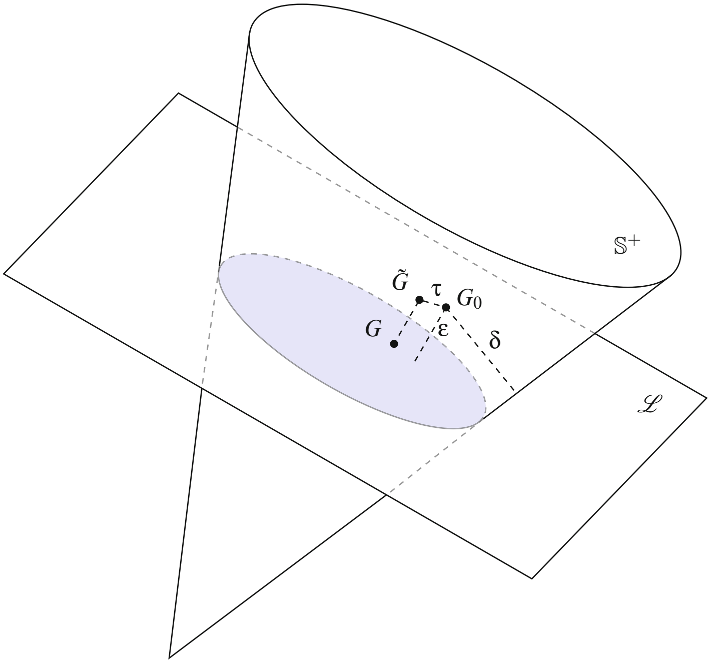

Fig. 1.1 Rounding and projecting to obtain a rational solution

So if

$$P^{-1}A_iP^{-T} = \begin{bmatrix} a_i & \mathbf{c}_i^T \\ \mathbf{c}_i & \hat{A}_i \end{bmatrix},$$

then $\hat{A}_i$ is a symmetric matrix with rational entries and

$$b_i = \operatorname{tr}\left(\begin{bmatrix} a_i & \mathbf{c}_i^T \\ \mathbf{c}_i & \hat{A}_i \end{bmatrix}\begin{bmatrix} 0 & 0 \\ 0 & \hat{G}_0 \end{bmatrix}\right) = \operatorname{tr}(\hat{A}_i\hat{G}_0) = \langle \hat{A}_i, \hat{G}_0 \rangle.$$

We have established a variant of the facial reduction [BW81] which applies whenever the original SDP is given by rational data and has a singular feasible point with a rational nullvector:

Theorem 1.81. *Let (FSDP), $\hat{A}_i$, and $\hat{G}_0$ be as above. Consider the feasibility SDP*

$$\begin{gathered} \hat{G} \succeq 0 \\ s.\,t.\ \langle \hat{A}_i, \hat{G} \rangle = b_i,\ i = 1, \ldots, m \end{gathered} \tag{FSDP'}$$

(i) (FSDP') *is feasible if and only if* (FSDP) *is feasible.*
(ii) (FSDP') *admits a rational solution if and only if* (FSDP) *does.*

The importance of SDP was spurred by the development of practically efficient methods to obtain (weakly) optimal solutions. More precisely, given an $\varepsilon > 0$ we can obtain by interior point methods an ε-optimal solution with polynomially many iterations, where each iteration takes polynomially many real number operations, provided that both (PSDP) and (DSDP) have non-empty interiors of feasible sets and we have good initial points. The variables appearing in these polynomial bounds are the order s of the matrix variable, the number m of linear constraints in (PSDP) and $\log \varepsilon$ (cf. [WSV00, Chap. 10.4.4]).

Note, however, that the complexity to obtain exact solutions of (PSDP) or (DSDP) is still a fundamental open question in semidefinite optimization [PK97]. The difficulties arise from the fact that semidefinite programs with rational input data may have an irrational optimal value or an optimal solution which is doubly exponential, hence has exponential length in any numerical system coding. Ramana [Ram97] proved that the decision problem whether there exists a feasible solution of (PSDP) or (DSDP)—the so-called SDP feasibility problem FSDP—is neither in NP nor in co-NP unless NP = co-NP, if we consider the Turing machine complexity models, and FSDP is in NP $\cap$ co-NP, if we consider the real number model. For more details about the complexity bounds for linear, SDP, and other convex quadratic programming problems we refer the reader to [BTN01].

There exist several open source packages which can efficiently find ε-optimal solutions in practice for most of the problems. If the problem is of medium size (i.e., $s \leq 1.000$ and $m \leq 10.000$), these packages are based on interior point methods, while packages for larger semidefinite programs use some variant of the first order methods (see [Mit15] for a comprehensive list of state-of-the-art SDP solvers and also [PRW06, MPRW09]). Nevertheless, once $s \geq 3.000$ or $m \geq 250.000$, the problem must share some special property otherwise state-of-the art solvers will fail to solve it for complexity reasons.

References

[Bar02] Barvinok, A.: A Course in Convexity. Graduate Studies in Mathematics, vol. 54. American Mathematical Society, Providence (2002)

[BTN01] Ben-Tal, A., Nemirovski, A.S.: Lectures on Modern Convex Optimization. MPS/SIAM Series on Optimization. Society for Industrial and Applied Mathematics (SIAM), Philadelphia (2001)

[BV04] Boyd, S., Vandenberghe, L.: Convex Optimization. Cambridge University Press, New York (2004)

[BW81] Borwein, J.M., Wolkowicz, H.: Facial reduction for a cone-convex programming problem. J. Aust. Math. Soc. Ser. A **30**(3), 369–380 (1980/1981)

[Brä11] Brändén, P.: Obstructions to determinantal representability. Adv. Math. **226**(2), 1202–1212 (2011)

[BK09] Brešar, M., Klep, I.: Noncommutative polynomials, Lie skew-ideals and tracial Nullstellensätze. Math. Res. Lett. **16**(4), 605–626 (2009)

[BK10] Burgdorf, S., Klep, I.: Trace-positive polynomials and the quartic tracial moment problem. C. R. Math. **348**(13–14), 721–726 (2010)

[BK12] Burgdorf, S., Klep, I.: The truncated tracial moment problem. J. Oper. Theory **68**(1), 141–163 (2012)

[BCKP13] Burgdorf, S., Cafuta, K., Klep, I., Povh, J.: The tracial moment problem and trace-optimization of polynomials. Math. Program. **137**(1–2), 557–578 (2013)

[CLR95] Choi, M.-D., Lam, T.Y., Reznick, B.: Sums of squares of real polynomials. In: *K*-Theory and Algebraic Geometry: Connections with Quadratic Forms and Division Algebras (Santa Barbara, CA, 1992). Proceedings of Symposia in Pure Mathematics, vol. 58, pp. 103–126. American Mathematical Society, Providence (1995)

[Con76] Connes, A.: Classification of injective factors. Cases II_1, II_∞, III_λ, $\lambda \neq 1$. Ann. Math. (2) **104**(1), 73–115 (1976)

[CF96] Curto, R.E., Fialkow, L.A.: Solution of the truncated complex moment problem for flat data. Mem. Am. Math. Soc. **119**(568), x+52 (1996)

[CF98] Curto, R.E., Fialkow, L.A.: Flat extensions of positive moment matrices: recursively generated relations. Mem. Am. Math. Soc. **136**(648), x+56 (1998)

[dK02] de Klerk, E.: Aspects of Semidefinite Programming. Applied Optimization, vol. 65. Kluwer Academic, Dordrecht (2002)

[Hel00] Helmberg, C.: Semidefinite Programming for Combinatorial Optimization. Konrad-Zuse-Zentrum für Informationstechnik, Berlin (2000)

[Hel02] Helton, J.W.: “Positive” noncommutative polynomials are sums of squares. Ann. Math. (2) **156**(2), 675–694 (2002)

[HM04] Helton, J.W., McCullough, S.: A Positivstellensatz for non-commutative polynomials. Trans. Am. Math. Soc. **356**(9), 3721–3737 (2004)

[HM12] Helton, J.W., McCullough, S.: Every convex free basic semi-algebraic set has an LMI representation. Ann. Math. (2) **176**(2), 979–1013 (2012)

[HKM12] Helton, J.W., Klep, I., McCullough, S.: The convex Positivstellensatz in a free algebra. Adv. Math. **231**(1), 516–534 (2012)

[HJ12] Horn, R.A., Johnson, C.R.: Matrix Analysis. Cambridge University Press, Cambridge (2012)

[KVV14] Kaliuzhnyi-Verbovetskyi, D.S., Vinnikov, V.: Foundations of Free Noncommutative Function Theory, vol. 199. American Mathematical Society, Providence (2014)

[KS07] Klep, I., Schweighofer, M.: A nichtnegativstellensatz for polynomials in noncommuting variables. Isr. J. Math. **161**, 17–27 (2007)

[KS08a] Klep, I., Schweighofer, M.: Connes’ embedding conjecture and sums of hermitian squares. Adv. Math. **217**(4), 1816–1837 (2008)

[KS08b] Klep, I., Schweighofer, M.: Sums of hermitian squares and the BMV conjecture. J. Stat. Phys. **133**(4), 739–760 (2008)

[Lam01] Lam, T.Y.: A First Course in Noncommutative Rings. Graduate Texts in Mathematics, vol. 131, 2nd edn. Springer, New York (2001)

[MPRW09] Malick, J., Povh, J, Rendl, F., Wiegele, A.: Regularization methods for semidefinite programming. SIAM J. Optim. **20**(1), 336–356 (2009)

[McC01] McCullough, S.: Factorization of operator-valued polynomials in several non-commuting variables. Linear Algebra Appl. **326**(1–3), 193–203 (2001)

[MP05] McCullough, S., Putinar, M.: Noncommutative sums of squares. Pac. J. Math. **218**(1), 167–171 (2005)

[Mit15] Mittelman, H.D.: http://plato.asu.edu/sub/pns.html (2015)

[MKKK10] Murota, K., Kanno, Y., Kojima, M., Kojima, S.: A numerical algorithm for block-diagonal decomposition of matrix $*$-algebras with application to semidefinite programming. Jpn. J. Ind. Appl. Math. **27**(1), 125–160 (2010)

[NT14] Netzer, T., Thom, A.: Hyperbolic polynomials and generalized clifford algebras. Discrete Comput. Geom. **51**, 802–814 (2014)

[Par03] Parrilo, P.A.: Semidefinite programming relaxations for semialgebraic problems. Math. Program. **96**(2, Ser. B), 293–320 (2003)

[PP08] Peyrl, H., Parrilo, P.A.: Computing sum of squares decompositions with rational coefficients. Theor. Comput. Sci. **409**(2), 269–281 (2008)

[PK97] Porkolab, L., Khachiyan, L.: On the complexity of semidefinite programs. J. Glob. Optim. **10**(4), 351–365 (1997)

[PRW06] Povh, J., Rendl, F., Wiegele, A.: A boundary point method to solve semidefinite programs. Computing **78**, 277–286 (2006)

[PS01] Powers, V., Scheiderer, C.: The moment problem for non-compact semialgebraic sets. Adv. Geom. **1**(1), 71–88 (2001)

[PW98] Powers, V., Wörmann, T.: An algorithm for sums of squares of real polynomials. J. Pure Appl. Algebra **127**(1), 99–104 (1998)

[Pro76] Procesi, C.: The invariant theory of $n \times n$ matrices. Adv. Math. **19**(3), 306–381 (1976)

[PS76] Procesi, C., Schacher, M.: A non-commutative real Nullstellensatz and Hilbert's 17th problem. Ann. Math. **104**(3), 395–406 (1976)

[Put93] Putinar, M.: Positive polynomials on compact semi-algebraic sets. Ind. Univ. Math. J. **42**(3), 969–984 (1993)

[Qua15] Quarez, R.: Trace-positive non-commutative polynomials. Proc. Am. Math. Soc. **143**(8), 3357–3370 (2015)

[Ram97] Ramana, M.V.: An exact duality theory for semidefinite programming and its complexity implications. Math. Program. **77**(2, Ser. B), 129–162 (1997)

[Row80] Rowen, L.H.: Polynomial Identities in Ring Theory. Pure and Applied Mathematics, vol. 84. Academic, New York (1980)

[Sch12] Scheiderer, C.: Sums of squares of polynomials with rational coefficients. arXiv:1209.2976, to appear in J. Eur. Math. Soc. (2012)

[Tak03] Takesaki, M.: Theory of Operator Algebras. III. Encyclopaedia of Mathematical Sciences, vol. 127. Operator Algebras and Non-commutative Geometry, 8. Springer, Berlin (2003)

[WSV00] Wolkowicz, H., Saigal, R., Vandenberghe, L.: Handbook of Semidefinite Programming. Kluwer, Boston (2000)

Chapter 2
Detecting Sums of Hermitian Squares

2.1 Introduction

The central question of this chapter is how to find out whether a given nc polynomial is a sum of hermitian squares (SOHS). We rely on Sect. 1.3, where we explained basic relations between SOHS polynomials and positive semidefinite Gram matrices. In this chapter we will enclose these results into the Gram matrix method and refine it with the Newton chip method.

2.2 The Gram Matrix Method

Recall from Sect. 1.3 that an nc polynomial $f \in \mathbb{R}\langle \underline{X} \rangle_{2d}$ is SOHS if and only if we can find a positive semidefinite Gram matrix associated with f, i.e., a positive semidefinite matrix G satisfying $\mathbf{W}_d^* G \mathbf{W}_d = f$, where $\mathbf{W}_d$ is the vector of all words of degree $\leq d$. This is a semidefinite feasibility problem in the matrix variable G. The constraints $\langle A_i \,|\, G \rangle = b_i$ are implied by the fact that for each monomial $w \in \mathbf{W}_{2d}$ we have

$$\sum_{\substack{u,v \in \mathbf{W}_d \\ u^* v = w}} G_{u,v} = a_w, \tag{2.1}$$

where a_w is the coefficient of w in f.

Problems like this can be (in theory) solved *exactly* using quantifier elimination [BPR06] as has been suggested in the commutative case by Powers and Wörmann [PW98]. However, this only works for problems of small size, so a *numerical* approach is needed in practice. Thus we turn to numerical methods to solve semidefinite programming problems.

© The Author(s) 2016

S. Burgdorf et al., *Optimization of Polynomials in Non-Commuting Variables*,
SpringerBriefs in Mathematics, DOI 10.1007/978-3-319-33338-0_2

Sums of hermitian squares are symmetric so we consider only $f \in \operatorname{Sym}\mathbb{R}\langle \underline{X} \rangle$. Two symmetric polynomials are equal if and only if all of their "symmetrized coefficients" (i.e., $a_w + a_{w^*}$) coincide, hence Eqs. (2.1) can be rewritten as

$$\sum_{\substack{u,v \in \mathbf{W}_d \\ u^*v=w}} G_{u,v} + \sum_{\substack{u,v \in \mathbf{W}_d \\ v^*u=w^*}} G_{v,u} = a_w + a_{w^*} \quad \forall w \in \mathbf{W}_{2d}, \tag{2.2}$$

or equivalently,

$$\langle A_w \,|\, G \rangle = a_w + a_{w^*} \quad \forall w \in \mathbf{W}_{2d}, \tag{2.3}$$

where A_w is the symmetric matrix defined by

$$(A_w)_{u,v} = \begin{cases} 2; \text{ if } \; u^*v = w, \; w^* = w, \\ 1; \text{ if } \; u^*v \in \{w, w^*\}, \; w^* \neq w, \\ 0; \text{ otherwise.} \end{cases}$$

Note that in this formulation the constraints obtained from w and w^* are the same so we keep only one of them. As we are interested in an arbitrary positive semidefinite G satisfying constraints (2.3), we can choose the objective function freely. However, in practice one prefers solutions of small rank leading to shorter SOHS decompositions. Hence we minimize the trace, a commonly used heuristic for matrix rank minimization (cf. [RFP10]). Therefore our SDP in primal form is as follows:

$$\begin{aligned} \inf \quad & \langle I \,|\, G \rangle \\ \text{s.t.} \; & \langle A_w \,|\, G \rangle = a_w + a_{w^*} \;\; \forall w \in \mathbf{W}_{2d} \\ & G \succeq 0. \end{aligned} \tag{SOHS$_{\text{SDP}}$}$$

Summing up, the Gram matrix method can be presented in Algorithm 2.1.

Algorithm 2.1: The Gram matrix method for finding SOHS decompositions

```
Input: $f \in \operatorname{Sym}\mathbb{R}\langle \underline{X} \rangle$ with $\deg f \leq 2d, f = \sum_{w \in \langle \underline{X} \rangle} a_w w,$, where $a_w \in \mathbb{R}$;
$\mathscr{G} = \varnothing$;
Construct $\mathbf{W}_d$;
Construct data $A_w, \mathbf{b}, C$ corresponding to (SOHS$_{\text{SDP}}$);
Solve (SOHS$_{\text{SDP}}$);
if (SOHS$_{\text{SDP}}$) is not feasible then
    $f \notin \Sigma^2$. Stop;
end
Take an optimal solution $G$ and compute the Cholesky decomposition
$G = R^*R$;
$\mathscr{G} = \{g_i\}$, where $g_i$ denotes the $i$th component of $R\mathbf{W}_d$;
Output: $\mathscr{G}$;
```

Remark 2.1. The order of G in (SOHS$_{\text{SDP}}$) is the length of $\mathbf{W}_d$, which is $\sigma = \frac{n^{d+1}-1}{n-1}$, as shown in Remark 1.12. Since $\sigma = \sigma(n,d)$ grows exponentially with the polynomial degree d it easily exceeds the size manageable by the state-of-the-art SDP solvers, which is widely accepted to be of order 1000. This implies, for example, that the above algorithm can only handle nc polynomials in two variables if they are of degree < 10. Therefore it is very important to find an improvement of the Gram matrix method which will be able to work with much larger nc polynomials. This will be done in the rest of the chapter.

Example 2.2. Let

$$f = X^2 - X^{10}Y^{20}X^{11} - X^{11}Y^{20}X^{10} + X^{10}Y^{20}X^{20}Y^{20}X^{10}. \tag{2.4}$$

The order of a Gram matrix G for f is $\sigma(10) = \sigma(2,10) = 2^{41} - 1$ and is too big for today's SDP solvers. Therefore any implementation of Algorithm 2.1 will get stuck. On the other hand, it is easy to see that

$$f = (X - X^{10}Y^{20}X^{10})^*(X - X^{10}Y^{20}X^{10}) \in \Sigma^2.$$

The polynomial f is sparse and an improved SDP for testing whether (sparse) polynomials are sums of hermitian squares will be given below.

The complexity of solving an SDP is also determined by the number of Eq. (2.3), which we denote by m. There are exactly

$$m = \operatorname{card}\{w \in \mathbf{W}_{2d} \mid w^* = w\} + \frac{1}{2}\operatorname{card}\{w \in \mathbf{W}_{2d} \mid w^* \neq w\}$$

such equations in (SOHS$_{\text{SDP}}$). Since $\mathbf{W}_d$ contains all words in $\langle \underline{X} \rangle$ of degree $\leq d$, we have $m > \frac{1}{2}\sigma(2d) = \frac{n^{2d+1}-1}{2(n-1)}$.

For each $w \in \mathbf{W}_{2d}$ there are t different pairs (u_i, v_i) such that $w = u_i^* v_i$, where $t = \deg w + 1$ if $\deg w \leq d$, and $t = 2d + 1 - \deg w$ if $\deg w \geq d + 1$. Note that $t \leq d + 1$. Therefore the matrices A_i defining constraints (2.3) have order $\sigma(d)$ and every matrix A_i has at most $d + 1$ nonzero entries if it corresponds to a symmetric monomial of f, and has at most $2(d+1)$ nonzero entries otherwise. Hence the matrices A_i are sparse. They are also pairwise orthogonal with respect to the standard scalar product on matrices $\langle X \,|\, Y \rangle = \operatorname{tr} X^T Y$, and have disjoint supports, as we now proceed to show:

Theorem 2.3. *Let $\{A_i \mid i = 1, \ldots, m\}$ be the matrices constructed in Step 3 of Algorithm 2.1 [i.e., matrices satisfying (2.3)]. If $(A_i)_{u,v} \neq 0$, then $(A_j)_{u,v} = 0$ for all $j \neq i$. In particular, $\langle A_i \,|\, A_j \rangle = 0$ for $i \neq j$.*

Proof. The equations in the SDP underlying the SOHS decomposition represent the constraints that the monomials in $\mathbf{W}_{2d}$ must have coefficients prescribed by the polynomial f. Let us fix $i \neq j$. The matrices A_i and A_j correspond to some monomials $p_1^* q_1$ and $p_2^* q_2$ ($p_i, q_i \in \mathbf{W}_d$), respectively, and $p_1^* q_1 \neq p_2^* q_2$. If A_i and A_j both have a nonzero entry at position (u, v), then $p_1^* q_1 = u^* v = p_2^* q_2$, a contradiction. ■

Remark 2.4. Sparsity and orthogonality of the constraints imply that the state-of-the-art SDP solvers can handle about 100,000 such constraints (see, e.g., [MPRW09]), if the order of the matrix variable is about 1000. The boundary point method introduced in [PRW06] and analyzed in [MPRW09] has turned out to perform best for semidefinite programs of this type. It is able to use the *orthogonality* of the matrices A_i (though *not* the disjointness of their supports). In the computationally most expensive steps—solving a linear system—the system matrix becomes diagonal, so solving the system amounts to dividing by the corresponding diagonal entries.

Since $\mathbf{W}_d$ contains all words in $\langle \underline{X} \rangle$ of degree $\leq d$, we have, e.g., for $n = 2$, $d = 10$ that $m = \sigma(20) = \sigma(2,20) = 2{,}097{,}150$ and this is clearly out of reach for *all* current SDP solvers. Nevertheless, we show in the sequel that one can replace the vector $\mathbf{W}_d$ in Step 2 of Algorithm 2.1 by a vector $\mathbf{W}$ which is usually much smaller and has at most kd words, where k is the number of symmetric monomials in f and $2d = \deg f$. Hence the order of the matrix variable G and the number of linear constraints m end up being much smaller in general.

2.3 Newton Chip Method

We present a modification of (Step 1 of) the Gram matrix method (Algorithm 2.1) by implementing the appropriate non-commutative analogue of the classical Newton polytope method [Rez78], which we call the *Newton chip method* and present it as Algorithm 2.2.

Definition 2.5. Let us define the *right chip function* $\operatorname{rc} : \langle \underline{X} \rangle \times \mathbb{N}_0 \to \langle \underline{X} \rangle$ by

$$\operatorname{rc}(w_1 \cdots w_n, i) := \begin{cases} w_{n-i+1} w_{n-i+2} \cdots w_n & \text{if } 1 \leq i \leq n; \\ w_1 \cdots w_n & \text{if } i > n; \\ 1 & \text{if } i = 0. \end{cases}$$

Example 2.6. Given the word $w = X_1 X_2 X_1 X_2^2 X_1 \in \langle \underline{X} \rangle$ we have $\operatorname{rc}(w,4) = X_1 X_2^2 X_1$, $\operatorname{rc}(w,6) = w$ and $\operatorname{rc}(w,0) = 1$.

We introduce the Newton chip method, presented as Algorithm 2.2. It substantially reduces the word vector needed in the Gram matrix method.

Theorem 2.7. *Suppose $f \in \operatorname{Sym}\mathbb{R}\langle \underline{X} \rangle$. Then $f \in \Sigma^2$ if and only if there exists a positive semidefinite matrix G satisfying*

$$f = \mathbf{W}^* G \mathbf{W},$$

where $\mathbf{W}$ is the output given by the Newton chip method (Algorithm 2.2).

Proof. Suppose $f \in \Sigma^2$. In every SOHS decomposition

$$f = \sum_i g_i^* g_i,$$

only words from $\mathscr{D}$ (constructed in Step 4) are used, i.e., $g_i \in \operatorname{span}\mathscr{D}$ for every i. This follows from the fact that the lowest and highest degree terms cannot cancel (cf. proof of Proposition 1.16). Let $\mathscr{W} := \bigcup_i \mathscr{W}_{g_i}$ be the union of the supports of the g_i. We shall prove that $\mathscr{W} \subseteq W$. For this, let us introduce a partial ordering on $\langle \underline{X} \rangle$:

$$w_1 \preceq w_2 \ \Leftrightarrow\ \exists i \in \mathbb{N}_0 : \operatorname{rc}(w_2, i) = w_1.$$

Note: $w_1 \preceq w_2$ if and only if there is a $v \in \langle \underline{X} \rangle$ with $w_2 = vw_1$.

CLAIM. For every $w \in \mathscr{W}$ there exists $u \in \langle \underline{X} \rangle$: $w \preceq u \preceq u^*u \in \mathscr{W}_f$.

Proof. Clearly, w^*w is a word that appears in the representation of $g_i^*g_i$ which one naturally gets by multiplying out without simplifying, for some i. If $w^*w \notin \mathscr{W}_f$, then there are $w_1, w_2 \in \mathscr{W} \setminus \{w\}$ with $w_1^*w_2 = w^*w$ (appearing with a negative coefficient so as to cancel the w^*w term). Then $w \preceq w_1$ or $w \preceq w_2$, without loss of generality, $w \preceq w_1$. Continuing the same line of reasoning, but starting with $w_1^*w_1$, we eventually arrive at $w_\ell \in \mathscr{W}$ with $w_\ell^*w_\ell \in \mathscr{W}_f$ and $w \preceq w_1 \preceq \cdots \preceq w_\ell$. Thus $w \preceq w_\ell \preceq w_\ell^*w_\ell \in \mathscr{W}_f$, concluding the proof of the claim.

The theorem follows now. Since $u^*u \in \mathscr{W}_f$ and w is a right chip of u we have $w \in \mathbf{W}$. ■

Algorithm 2.2: The Newton chip method

```
    Input: $f \in \operatorname{Sym}\mathbb{R}\langle \underline{X} \rangle$ with $\deg f \le 2d, f = \sum_{w\in\langle \underline{X}\rangle} a_w w,$, where $a_w \in \mathbb{R}$;
 1  Define the support of $f$ as $\mathscr{W}_f := \{w \in \langle \underline{X} \rangle \mid a_w \neq 0\}$;
 2  $W := \varnothing$;
 3  Let $m_i := \frac{\operatorname{mindeg}_i f}{2}$, $M_i := \frac{\deg_i f}{2}$, $m := \frac{\operatorname{mindeg} f}{2}$, $M := \frac{\deg f}{2}$;
 4  The set of admissible words is defined as

        $\mathscr{D} := \{w \in \langle \underline{X} \rangle \mid m_i \le \deg_i w \le M_i \text{ for all } i, m \le \deg w \le M\}$;

    for every $w^*w \in \mathscr{W}_f$ do
 5      for $0 \le i \le \deg w$ do
 6          if $\operatorname{rc}(w,i) \in \mathscr{D}$ then
 7              $W := W \cup \{\operatorname{rc}(w,i)\}$;
 8          end
 9      end
10  end
11  Sort $W$ in a lexicographic order and transform it into the vector $\mathbf{W}$;
    Output: $\mathbf{W}$;
```

Example 2.8 (Example 2.2 Continued). The polynomial f from Example 2.2 has two hermitian squares: X^2 and $X^{10}Y^{20}X^{20}Y^{20}X^{10}$. The first hermitian square contributes via the Newton chip method only one right chip: X; while the second hermitian square $X^{10}Y^{20}X^{20}Y^{20}X^{10}$ contributes to $\mathbf{W}$ the following words: X, $X^2, \ldots, X^{10}$ as well as $YX^{10}, Y^2X^{10}, \ldots, Y^{20}X^{10}, XY^{20}X^{10}, \ldots, X^{10}Y^{20}X^{10}$.

Applying the Newton chip method to f therefore yields $\mathbf{W}$ which is a vector in the lexicographic order and is equal to

$$\mathbf{W} = \left[X \;\; X^2 \cdots X^{10} \;\; YX^{10} \;\; \cdots Y^{20}X^{10} \;\; XY^{20}X^{10} \;\; \cdots X^{10}Y^{20}X^{10}\right]^T$$

of length 40. Problems of this size are easily handled by today's SDP solvers. Nevertheless we provide a further strengthening of our Newton chip method reducing the number of words needed in this example to 2 (see Sect. 2.4).

2.4 Augmented Newton Chip Method

The following simple observation is often crucial to reduce the size of $\mathbf{W}$ returned by the Newton chip method.

Lemma 2.9. *Suppose* $\mathbf{W}$ *is the vector of words returned by the Newton chip method. If there exists a word* $u \in \mathbf{W}$ *such that the constraint in* (SOHS$_{\text{SDP}}$) *corresponding to* u^*u *can be written as*

$$\langle A_{u^*u} \,|\, G \rangle = 0$$

and A_{u^*u} *is a diagonal matrix* (*i.e.,* $(A_{u^*u})_{u,u} = 2$ *and* A_{u^*u} *is* 0 *elsewhere*)*, then we can eliminate* u *from* $\mathbf{W}$ *and likewise delete this equation from the semidefinite program.*

Proof. Indeed, such a constraint implies that $G_{u,u} = 0$ for the given $u \in \mathbf{W}$, hence the uth row and column of G must be zero, since G is positive semidefinite. So we can decrease the order of (SOHS$_{\text{SDP}}$) by deleting the uth row and column from G and by deleting this constraint. ■

Lemma 2.9 applies if and only if there exists a constraint $\langle A_w \,|\, G \rangle = 0$, where $w = u^*u$ for some $u \in \mathbf{W}$ and $w \neq v^*z$ for all $v, z \in \mathbf{W}$, $v \neq z$. Therefore we augment the Newton chip method (Algorithm 2.2) by new steps, as shown in Algorithm 2.3.

Algorithm 2.3: The Augmented Newton chip method

```
   Input: $f \in \mathrm{Sym}\,\mathbb{R}\langle \underline{X} \rangle$ with $\deg f \leq 2d, f = \sum_{w \in \langle \underline{X} \rangle} a_w w,$, where $a_w \in \mathbb{R}$;
 1 Compute $\mathbf{W}$ by the Newton chip method (Algorithm 2.2);
 2 while exists $u \in \mathbf{W}$ such that $a_{u^*u} = 0$ and $u^*u \neq v^*z$ for every pair $v, z \in \mathbf{W}$,
   $v \neq z$ do
 3     delete $u$ from $\mathbf{W}$;
 4 end
   Output: $\mathbf{W}$;
```

Note that in Step 2 there might exist some word $u \in \mathbf{W}$ which does not satisfy the condition initially but after deleting another u' from $\mathbf{W}$ it does. We demonstrate Algorithm 2.3 in the following example:

Example 2.10 (Example 2.2 Continued). By applying the Augmented Newton chip method to f from (2.4) we reduce the vector $\mathbf{W}$ significantly. Note that after Step 1, $\mathbf{W}$ also contains the words X^8, X^9, X^{10}. Although X^{18} does not appear in f, we cannot delete X^9 from $\mathbf{W}$ immediately since $X^{18} = (X^9)^*X^9 = (X^8)^*X^{10}$. But we can delete X^{10} since X^{20} does not appear in f and $(X^{10})^*X^{10}$ is the unique decomposition of X^{20} inside $\mathbf{W}$. After deleting X^{10} from $\mathbf{W}$ we realize that $(X^9)^*X^9$ becomes the unique decomposition of X^{18}, hence we can eliminate X^9 too. Eventually the Augmented Newton chip method returns

$$\mathbf{W} = \begin{bmatrix} X & X^{10}Y^{20}X^{10} \end{bmatrix}^T,$$

which is exactly the minimum vector needed for the SOHS decomposition of f.

2.5 Implementation

2.5.1 On the Gram Matrix Method

The Gram matrix method (Algorithm 2.1) consists of two main parts: (1) constructing the matrices corresponding to ($\mathrm{SOHS}_{\mathrm{SDP}}$)—Step 3 and (2) solving the constructed SDP in Step 4. Step 3 is straightforward, running the Augmented Newton chip method (Algorithm 2.3) gives the desired vector of relevant words. There are no numerical problems, no convergence issues, Algorithm 2.3 always terminates with the desired vector $\mathbf{W}$.

The second main part is more subtle. Solving an instance of SDP in practice always involves algorithms that are highly numerical: algorithms to compute spectral decompositions, solutions of systems of linear equations, inverses of matrices, etc. Methods for solving SDP, especially interior point methods [dK02, Ter96, WSV00], but also some first order methods [MPRW09, PRW06], typically assume strictly feasible solutions on the primal and the dual side, which imply the strong duality property and the attainability of optimums on both sides. Moreover, this assumption also guarantees that most of the methods will converge to a primal–dual ε-optimal solution; see also Sect. 1.13.

As the following example demonstrates, the Slater condition is not necessarily satisfied on the primal side in our class of ($\mathrm{SOHS}_{\mathrm{SDP}}$) problems.

Example 2.11. Let $f = (XY + X^2)^*(XY + X^2)$. It is homogeneous, and the Augmented Newton chip method gives

$$\mathbf{W} = \begin{bmatrix} X^2 \\ XY \end{bmatrix}.$$

There exists a unique symmetric Gram matrix

$$G = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}$$

for f such that $f = \mathbf{W}^* G \mathbf{W}$. Clearly G, a rank 1 matrix, is the only feasible solution of (SOHS$_{\text{SDP}}$), hence the corresponding SDP has no strictly feasible solution on the primal side.

If we take the objective function in our primal SDP (SOHS$_{\text{SDP}}$) to be equal to $\langle I \mid G \rangle$, then the pair $y = 0$, $Z = I$ is always strictly feasible for the dual problem of (SOHS$_{\text{SDP}}$) and thus we *do* have the strong duality property.

Hence, when the given nc polynomial is in Σ^2, the corresponding semidefinite program (SOHS$_{\text{SDP}}$) is feasible, and the optimal value is attained. If there is no strictly feasible solution, then numerical difficulties might arise but state-of-the-art SDP solvers such as SeDuMi [Stu99], SDPT3 [TTT99], SDPA [YFK03], or MOSEK [ApS15] are able to overcome them in most of the instances. When the given nc polynomial is not in Σ^2, then the semidefinite problem (SOHS$_{\text{SDP}}$) is infeasible and this might cause numerical problems as well. However, state-of-the-art SDP solvers are generally robust and can reliably detect infeasibility for most practical problems; for more details see [dKRT98, PT09].

2.5.2 Software Package `NCSOStools`

The software package `NCSOStools` [CKP11] was developed to help researchers working in the area of non-commutative polynomials. `NCSOStools` [CKP11] is an open source Matlab toolbox for solving SOHS related problems using semidefinite programming. It also implements symbolic computation with non-commuting variables in Matlab.

There is a small overlap in features with Helton's `NCAlgebra` package for Mathematica [HMdOS15]. However, `NCSOStools` [CKP11] performs basic manipulations with non-commuting variables and is mainly oriented to detect several variants of constrained and unconstrained positivity of nc polynomials, while `NCAlgebra` is a fully fledged add-on for symbolic computation with polynomials, matrices, and rational functions in non-commuting variables.

When we started writing `NCSOStools` we decided to use Matlab as a main framework since we solve the underlying SDP instances by existing open source solvers like SeDuMi [Stu99], SDPT3 [TTT99], or SDPA [YFK03] and these solvers can be very easily run within Matlab.

Readers interested in solving sums of squares problems for commuting polynomials are referred to one of the many great existing packages, such as SOSTOOLS [PPSP05], SparsePOP [WKK$^+$09], GloptiPoly [HLL09], or YALMIP [Löf04].

Detecting sums of hermitian squares by the Gram matrix method and using the (Augmented) Newton chip method can be done within `NCSOStools` by calling `NCsos`.

Example 2.12 (Example 2.11 Continued). We declare the polynomial f that we started considering in Example 2.11 within NCSOStools by

```
NCvars x y
>> f=(x*y+x^2)'*(x*y+x^2)
```

By calling

```
>> [IsSohs,Gr,W,sohs,g,SDP_data,L] = NCsos(f)
```

we obtain that f is SOHS (`IsSohs=1`), the vector given by the Augmented Newton chip methods (`W`) and the corresponding Gram matrix `Gr`:

```
W =
    'x*x'
    'x*y'

Gr =
    1.0000    1.0000
    1.0000    1.0000
```

Likewise we obtain the SOHS decomposition of f

```
sohs =
       x^2+x*y
   2.2e-07*x*y
```

which means that the SOHS decomposition for f is

$$f = (X^2 + XY)^*(X^2 + XY) + (2.2 \cdot 10^{-7} XY)^*(2.2 \cdot 10^{-7} XY).$$

This is ε correct for $\varepsilon = 10^{-13}$, i.e., if we leave cut off all monomials with coefficients less than 10^{-13} we obtain f. We can control precision using the parameter `pars.precision`. All monomials in `sohs` having coefficient smaller than `pars.precision` are ignored. Therefore by running

```
>> pars.precision=1e-6;
>> [IsSohs,Gr,W,sohs,g,SDP_data,L] = NCsos(f,pars);
```

we obtain the exact value for a SOHS decomposition of f, i.e., f is exactly a SOHS of elements from `sohs`.

The data describing the semidefinite program (SOHS_{SDP}) is given in `SDP_data` while the optimal matrix for the dual problem to (SOHS_{SDP}) is given in `L`. In `g` we return sum of squares of entries from `sohs` with monomials having coefficient larger than 10^{-8} which is an internal parameter.

References

[BPR06] Basu, S., Pollack, R., Roy, M.-F.: Algorithms in Real Algebraic Geometry. Algorithms and Computation in Mathematics, vol. 10, 2nd edn. Springer, Berlin (2006)

[CKP11] Cafuta, K., Klep, I., Povh, J.: NCSOStools: a computer algebra system for symbolic and numerical computation with noncommutative polynomials. Optim. Methods Softw. **26**(3), 363–380 (2011). Available from http://ncsostools.fis.unm.si/

[dK02] de Klerk, E.: Aspects of Semidefinite Programming. Applied Optimization, vol. 65. Kluwer Academic, Dordrecht (2002)

[dKRT98] de Klerk, E., Roos, C., Terlaky, T.: Infeasible-start semidefinite programming algorithms via self-dual embeddings. In: Topics in Semidefinite and Interior-Point Methods (Toronto, ON, 1996). Fields Institute Communications, vol. 18, pp. 215–236. American Mathematical Society, Providence (1998)

[HMdOS15] Helton, J.W., Miller, R.L., de Oliveira, M.C., Stankus, M.: NCAlgebra: a mathematica package for doing non commuting algebra. Available from http://www.math.ucsd.edu/~ncalg/ (2015)

[HLL09] Henrion, D., Lasserre, J.B., Löfberg, J.: GloptiPoly 3: moments, optimization and semidefinite programming. Optim. Methods Softw. **24**(4–5), 761–779 (2009). Available from http://www.laas.fr/~henrion/software/gloptipoly3/

[Löf04] Löfberg, J.: YALMIP: a toolbox for modeling and optimization in MATLAB. In: Proceedings of the CACSD Conference, Taipei. Available from http://control.ee.ethz.ch/~joloef/wiki/pmwiki.php (2004)

[MPRW09] Malick, J., Povh, J., Rendl, F., Wiegele, A.: Regularization methods for semidefinite programming. SIAM J. Optim. **20**(1), 336–356 (2009)

[ApS15] MOSEK ApS: The MOSEK optimization toolbox for MATLAB manual. Version 7.1 (Revision 28) (2015)

[PT09] Pólik, I., Terlaky, T.: New stopping criteria for detecting infeasibility in conic optimization. Optim. Lett. **3**(2), 187–198 (2009)

[PRW06] Povh, J., Rendl, F., Wiegele, A.: A boundary point method to solve semidefinite programs. Computing **78**, 277–286 (2006)

[PW98] Powers, V., Wörmann, T.: An algorithm for sums of squares of real polynomials. J. Pure Appl. Algebra **127**(1), 99–104 (1998)

[PPSP05] Prajna, S., Papachristodoulou, A., Seiler, P., Parrilo, P.A.: SOSTOOLS and its control applications. In: Positive Polynomials in Control. Lecture Notes in Control and Information Science, vol. 312, pp. 273–292. Springer, Berlin (2005)

[RFP10] Recht, B., Fazel, M., Parrilo, P.A.: Guaranteed minimum-rank solutions of linear matrix equations via nuclear norm minimization. SIAM Rev. **52**(3), 471–501 (2010)

[Rez78] Reznick, B.: Extremal PSD forms with few terms. Duke Math. J. **45**(2), 363–374 (1978)

[Stu99] Sturm, J.F.: Using SeDuMi 1.02, a MATLAB toolbox for optimization over symmetric cones. Optim. Methods Softw. **11/12**(1–4), 625–653 (1999). Available from http://sedumi.ie.lehigh.edu/

[Ter96] Terlaky, T. (ed.): Interior Point Methods of Mathematical Programming. Applied Optimization, vol. 5. Kluwer Academic, Dordrecht (1996)

[TTT99] Toh, K.C., Todd, M.J., Tütüncü, R.: SDPT3–a MATLAB software package for semidefinite programming, version 1.3. Optim. Methods Softw. **11/12**(1–4), 545–581 (1999). Available from http://www.math.nus.edu.sg/~mattohkc/sdpt3.html

[WKK$^+$09] Waki, H., Kim, S., Kojima, M., Muramatsu, M., Sugimoto, H.: Algorithm 883: sparsePOP—a sparse semidefinite programming relaxation of polynomial optimization problems. ACM Trans. Math. Softw. **35**(2), Art. 15, 13 (2009)

[WSV00] Wolkowicz, H., Saigal, R., Vandenberghe, L.: Handbook of Semidefinite Programming. Kluwer, Boston (2000)

[YFK03] Yamashita, M., Fujisawa, K., Kojima, M.: Implementation and evaluation of SDPA 6.0 (semidefinite programming algorithm 6.0). Optim. Methods Softw. **18**(4), 491–505 (2003). Available from http://sdpa.sourceforge.net/

Chapter 3
Cyclic Equivalence to Sums of Hermitian Squares

3.1 Introduction

When we move focus from positive semidefinite non-commutative polynomials to trace-positive non-commutative polynomials we naturally meet cyclic equivalence to hermitian squares, see Definitions 1.57 and 1.60. In this chapter we will consider the question whether an nc polynomial is cyclically equivalent to SOHS, i.e., whether it is a member of the cone Θ^2, which is a sufficient condition for trace-positivity. A special attention will be given to algorithmic aspects of detecting members in Θ^2. We present a tracial version of the Gram matrix method based on the tracial version of the Newton chip method which by using semidefinite programming efficiently answers the question if a given nc polynomial is or is not cyclically equivalent to a sum of hermitian squares.

Example 3.1. Consider $f = X^2Y^2 + XY^2X + XYXY + YX^2Y + YXYX + Y^2X^2$. This nc polynomial belongs to Θ^2 as can be seen from

$$\begin{aligned} f &= (XYXY + YXYX + XY^2X + YX^2Y) + 2XY^2X + [Y^2X,X] + [X,XY^2] \\ &= (XY+YX)^*(XY+YX) + 2(YX)^*(YX) + [Y^2X,X] + [X,XY^2] \\ &\overset{\text{cyc}}{\sim} (XY+YX)^*(XY+YX) + 2(YX)^*(YX). \end{aligned}$$

In particular, $\operatorname{tr} f(A,B) \geq 0$ for all symmetric matrices A, B but in general $f(A,B)$ is not positive semidefinite.

Testing whether a given $f \in \mathbb{R}\langle \underline{X} \rangle$ is an element of Θ^2 can be done efficiently by using semidefinite programming as first observed in [KS08, Sect. 3], see also [BCKP13]. The method behind this is a variant of the Gram matrix method and arises as a natural extension of the results for sums of hermitian squares (cf. Sect. 2.2) or for polynomials in commuting variables [CLR95, Sect. 2]; see

© The Author(s) 2016

S. Burgdorf et al., *Optimization of Polynomials in Non-Commuting Variables*,
SpringerBriefs in Mathematics, DOI 10.1007/978-3-319-33338-0_3

also [Par03]. In this chapter we present the improved tracial Gram matrix method which is based on a tracial version of the classical Newton polytope used to reduce the size of the underlying semidefinite programming problem. The concrete formulation is a bit technical but the core idea is straightforward and goes as follows. Define the Newton polytope of an nc polynomial f as the Newton polytope of an appropriate interpretation of f as a polynomial in commuting variables. Now apply the Newton polytope method and then lift the obtained set of monomials in commuting variables to a set of monomials in non-commuting variables.

3.2 The Cyclic Degree

Our viewpoint focuses on the dual description of the tracial version of the Newton polytope, described by the so-called cyclic-α-degree. This viewpoint clarifies the chosen interpretation of an nc polynomial as a polynomial in commuting variables which is used in the algorithm.

We will need to consider the monoid $[\underline{x}]$ in *commuting* variables $\underline{x} = (x_1, \ldots, x_n)$ and its semigroup algebra $\mathbb{R}[\underline{x}]$ of polynomials. Its monomials are written in the form $\underline{x}^{\mathbf{d}} = x_1^{d_1} \cdots x_n^{d_n}$ for $\mathbf{d} = (d_1, \ldots, d_n) \in \mathbb{N}_0^n$. There is a natural mapping $\langle \underline{X} \rangle \to [\underline{x}]$. For a given word $w \in \langle \underline{X} \rangle$ its image under this mapping is of the form $\underline{x}^{\mathbf{d}_w}$, where $d_{w,i}$ denotes how many times X_i appears in w. It is called the *commutative collapse* of w. Similarly, we introduce the commutative collapse of a set of words $V \subseteq \mathbb{R}\langle \underline{X} \rangle$. For $f = \sum_w a_w w \in \mathbb{R}\langle \underline{X} \rangle$ we define the set

$$\mathrm{cc}(f) := \{\underline{x}^{\mathbf{d}_w} \in [\underline{x}] \mid a_w \neq 0\}.$$

We generalize the degree of an nc polynomial as follows:

Definition 3.2. Given $\underline{\alpha} = (\alpha_1, \ldots, \alpha_n) \in \mathbb{R}^n$ we define the *$\underline{\alpha}$-degree* $\deg_{\underline{\alpha}}$ of a word $w \in \langle \underline{X} \rangle$ as the standard scalar product between $\underline{\alpha}$ and the exponent of the commutative collapse $\underline{x}^{\mathbf{d}_w}$ of w, i.e.,

$$\deg_{\underline{\alpha}} w := \sum_{i=1}^{n} \alpha_i d_{w,i} = \langle \underline{\alpha} \,|\, \mathbf{d}_w \rangle.$$

We also set $\deg_{\underline{\alpha}} 0 := -\infty$.

Note that for all $\underline{\alpha} \in \mathbb{R}^n$, we have

$$u \overset{\mathrm{cyc}}{\sim} v \Rightarrow \deg_{\underline{\alpha}} u = \deg_{\underline{\alpha}} v \text{ and } \deg_{\underline{\alpha}}(uv) = \deg_{\underline{\alpha}} u + \deg_{\underline{\alpha}} v.$$

This notion extends naturally to the $\underline{\alpha}$-degree of arbitrary nc polynomial $f = \sum_w a_w w \in \mathbb{R}\langle \underline{X} \rangle$:

$$\deg_{\underline{\alpha}} f := \max_{a_w \neq 0} \deg_{\underline{\alpha}} w.$$

Remark 3.3. As special cases, note that $\deg f$ corresponds to the $\underline{\alpha}$ with all ones and $\deg_i f$ (the degree in variable X_i) corresponds to the standard unit vectors $\mathbf{e}_i$.

Two cyclically equivalent nc polynomials in general do not have the same $\underline{\alpha}$-degree. We therefore modify the definition to obtain the more robust *cyclic-$\underline{\alpha}$-degree* $\operatorname{cdeg}_{\underline{\alpha}}$:

$$\begin{aligned}\operatorname{cdeg}_{\underline{\alpha}} f &:= \min_{g \overset{\text{cyc}}{\sim} f} \deg_{\underline{\alpha}} g \\ \operatorname{cdeg} f &:= \operatorname{cdeg}_{(1,\dots,1)} f.\end{aligned}$$

For instance, for $f = X_1^2X_2^2X_1^2 + X_2^4X_3^4 - X_3^4X_2^4 + X_1X_2 - X_2X_1 \overset{\text{cyc}}{\sim} X_1^4X_2^2$ we have

$$\deg_{(1,1,3)} f = 16\ ,\ \operatorname{cdeg}_{(1,1,3)} f = 6.$$

Definition 3.4. Let $w \in \mathbb{R}\langle \underline{X} \rangle$. The canonical representative $[w]$ of w is the first with respect to the lexicographic order among words cyclically equivalent to w. For an nc polynomial $f = \sum_w a_w w \in \operatorname{Sym}\mathbb{R}\langle \underline{X} \rangle$ we define the *canonical representative* $[f]$ of f as follows:

$$[f] := \sum_{w \in \mathbb{R}\langle \underline{X} \rangle} a_w [w] \in \mathbb{R}\langle \underline{X} \rangle.$$

That is, $[f]$ contains only canonical representatives of words from f, and the coefficient of $[w]$ in $[f]$ is

$$\sum_{u \overset{\text{cyc}}{\sim} w} a_u.$$

For example, if $f = 2Y^2X^2 - XY^2X + XY - YX$, then $[f] = X^2Y^2$.

The next proposition shows that the cyclic-α-degree is compatible with the equivalence relation $\overset{\text{cyc}}{\sim}$ and equals the degree of the canonical representative.

Proposition 3.5.

(i) *For any polynomials $f, g \in \mathbb{R}\langle \underline{X} \rangle$, we have $f \overset{\text{cyc}}{\sim} g$ if and only if $[f] = [g]$.*
(ii) *For all $\alpha \in \mathbb{R}^n$ and $f \in \mathbb{R}\langle \underline{X} \rangle$ we have* $\operatorname{cdeg}_{\underline{\alpha}} f = \deg_{\underline{\alpha}} [f]$.

Proof. Property (i) is obvious by Remark 1.51, part (b). Let us consider (ii). Since $f \overset{\text{cyc}}{\sim} [f]$, $\operatorname{cdeg}_{\underline{\alpha}} f \le \deg_{\underline{\alpha}} [f]$. Suppose there exists $g \overset{\text{cyc}}{\sim} f$ with $\deg_{\underline{\alpha}_0} g < \deg_{\underline{\alpha}_0} [f]$ for some $\underline{\alpha}_0 \in \mathbb{R}^n$. There is a word $[w]$ with $\deg_{\underline{\alpha}_0} [w] = \deg_{\underline{\alpha}_0} [f]$, and the coefficient of $[w]$ in $[f]$ is nonzero. But by the first part of the proposition the same is true for g, hence $\deg_{\underline{\alpha}_0} g \ge \deg_{\underline{\alpha}_0} [f]$, which is a contradiction. ■

3.3 The Tracial Newton Polytope

Given a polynomial $f \in \mathbb{R}[\underline{x}]$ (in commuting variables) the *Newton polytope* $N(f)$ consists of all integer lattice points in the convex hull of the degrees $\mathbf{d} = (d_1, \dots, d_n)$ of monomials appearing in f, considered as vectors in $\mathbb{R}^n$ (see, e.g., [Rez78] for details). That is, for $f = \sum_{\mathbf{d}} a_{\mathbf{d}} \underline{x}^{\mathbf{d}} \in \mathbb{R}[\underline{x}]$,

$$N(f) := \mathbb{Z}^n \cap \operatorname{conv}\big(\{\mathbf{d} \in \mathbb{Z}^n \mid a_{\mathbf{d}} \neq 0\}\big).$$

We will also refer to the set $\frac{1}{2}N(f) := \{\mathbf{d} \in \mathbb{Z}^n \mid 2\mathbf{d} \in N(f)\}$. Alternatively, by dualization, one can describe the Newton polytope via the α-degree, namely

$$N(f) = \mathbb{Z}^n \cap \operatorname{conv}\{\mathbf{d} \in \mathbb{Z}^n \mid \deg_{\underline{\alpha}} \underline{x}^{\mathbf{d}} \leq \deg_{\underline{\alpha}} f \quad \text{for all } \underline{\alpha} \in \mathbb{R}^n\}.$$

Similarly, $N(S)$ and $\frac{1}{2}N(S)$ are defined, where S is a set of monomials in commuting variables. By dualization one immediately derives the following lemma:

Lemma 3.6. *Let $f \in \mathbb{R}\langle \underline{X} \rangle$. A word $w \in \langle \underline{X} \rangle$ with commutative collapse $\underline{x}^{\mathbf{d}_w}$ satisfies $\deg_{\underline{\alpha}} w \leq \operatorname{cdeg}_{\underline{\alpha}} f$ for all $\underline{\alpha} \in \mathbb{R}^n$ if and only if $\mathbf{d}_w$ is contained in the convex hull of the vectors $\{\mathbf{d}_v \mid v \in \operatorname{cc}([f])\}$.*

In other words, the *tracial Newton polytope* of an nc polynomial $f \in \mathbb{R}\langle \underline{X} \rangle$ is given by the classical Newton polytope for the commutative collapse of the canonical representative $[f]$ of f. Hence a word $w \in \langle \underline{X} \rangle$ should be included in the sum of hermitian squares and commutators factorization for a given noncommutative polynomial f if and only if the exponent $\mathbf{d}_w$ of its commutative collapse is contained in one half times the Newton polytope of the commutative collapse of $[f]$. In fact, this will be shown in Theorem 3.10, where we present the augmented tracial Gram matrix method.

Example 3.7. Let $f = 1 - XY^3 + Y^3X + 2Y^2 - 4X^5$. Then $[f] = 1 + 2Y^2 - 4X^5$,

$$\operatorname{cc}(f) = \{1, xy^3, y^2, x^5\} \subseteq [x,y] \text{ and } \quad \operatorname{cc}([f]) = \{1, y^2, x^5\} \subseteq [x,y],$$

where $[x,y]$ is the monoid generated by commuting variables x, y. Furthermore we have

$$\begin{aligned} N(\operatorname{cc}([f])) &= \mathbb{Z}^2 \cap \operatorname{conv}\big(\{(0,0),\, (0,2),\, (5,0)\}\big) \\ &= \big\{(0,0),\, (1,0),\, (2,0),\, (3,0),\, (4,0),\, (5,0),\, (0,1),\, (1,1),\, (2,1),\, (0,2)\big\}. \end{aligned}$$

We note that by taking the canonical representative $[f]$ instead of f itself we get a unique Newton polytope for f which is also the smallest Newton polytope among all Newton polytopes of possible interpretations of f in $\mathbb{R}[\underline{x}]$ (Fig. 3.1).

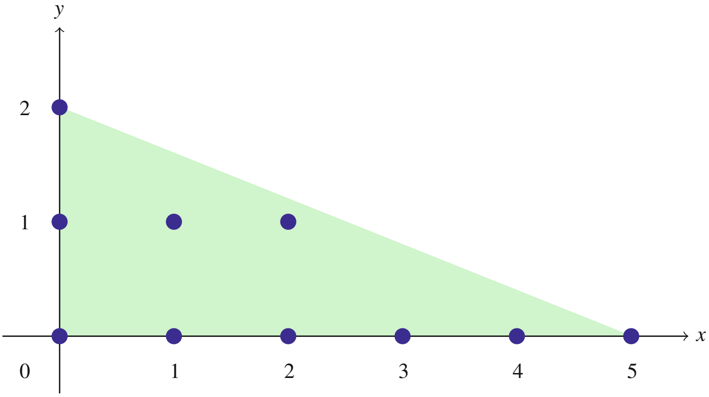

Fig. 3.1 The Newton polytope of $f = 1 - XY^3 + Y^3X + 2Y^2 - 4X^5$

3.4 The Tracial Gram Matrix Method

In this section we present the improved tracial Gram matrix method based on the tracial Newton polytope. That is, to construct a tracial Gram matrix for an nc polynomial $f \in \mathbb{R}\langle \underline{X} \rangle$ we will only consider words $w \in \langle \underline{X} \rangle$ whose exponent $\mathbf{d}_w$ of its commutative collapse is contained in one half times the tracial Newton polytope of f. This will be expressed by the cyclic-α-degree using the following corollary, which is an immediate consequence of Proposition 3.5 (ii) and Lemma 3.6.

Corollary 3.8. *Let $f \in \mathbb{R}\langle \underline{X} \rangle$ be an nc polynomial. Then*

$$\mathrm{cc}(\mathbf{W}) = \left\{ \underline{x}^{\mathbf{d}} \mid \mathbf{d} \in \frac{1}{2} N(\mathrm{cc}([f])) \right\} \tag{3.1}$$

for the vector $\mathbf{W}$ *consisting of all words $w \in \langle \underline{X} \rangle$ satisfying* $2\deg_{\underline{\alpha}} w \leq \mathrm{cdeg}_{\underline{\alpha}} f$ *for all $\underline{\alpha} \in \mathbb{R}^n$.*

Example 3.9. For $f = 1 - XY^3 + Y^3X + 2Y^2 - 4X^5$ from Example 3.7 we have $\frac{1}{2}N(\mathrm{cc}([f])) = \big\{(0,0), (0,1), (1,0), (2,0)\big\}$. One easily verifies $\mathbf{W} = \begin{bmatrix} 1 & Y & X & X^2 \end{bmatrix}^T$ and hence (3.1) holds.

Theorem 3.10. *Suppose $f \in \mathbb{R}\langle \underline{X} \rangle$. Then $f \in \Theta^2$ if and only if there exists a positive semidefinite matrix G such that*

$$f \overset{\mathrm{cyc}}{\sim} \mathbf{W}^* G \mathbf{W}, \tag{3.2}$$

where $\mathbf{W}$ *is a vector consisting of all words $w \in \langle \underline{X} \rangle$ satisfying*

$$2\deg_{\underline{\alpha}} w \leq \mathrm{cdeg}_{\underline{\alpha}} f \quad \textit{for all } \underline{\alpha} \in \mathbb{R}^n. \tag{3.3}$$

Furthermore, given such a positive semidefinite matrix G of rank r, one can construct nc polynomials $g_1, \ldots, g_r \in \mathbb{R}\langle \underline{X} \rangle$ with $f \overset{\mathrm{cyc}}{\sim} \sum_{i=1}^r g_i^* g_i$.

Corollary 3.11. *If $f \in \operatorname{Sym}\mathbb{R}\langle \underline{X}\rangle$ with* $\operatorname{cdeg} f = 2d$, *then* $f \in \Theta^2 \iff f \in \Theta^2_{2d}$.

Proof. Indeed, if $f \in \Theta^2$, then $f \overset{\mathrm{cyc}}{\sim} \mathbf{W}^* G\mathbf{W}$ with $\mathbf{W}$ and G from the theorem above. Suppose $w \in \mathbf{W}$. Then w satisfies (3.3), hence for $\alpha = (1,\dots,1)$ we have

$$2\deg_{\underline{\alpha}} w = 2\deg w \le \operatorname{cdeg}_{\underline{\alpha}} f = 2d,$$

i.e., $w \in \mathbf{W}_d$. Therefore $f \in \Theta^2_{2d}$. The converse is obvious. ■

For the proof of Theorem 3.10 we need one last ingredient, namely that the cyclic-α-degree of a sum of hermitian squares is equal to its α-degree.

Lemma 3.12. *If* $f \overset{\mathrm{cyc}}{\sim} g = \sum_i g_i^* g_i$, *then* $\operatorname{cdeg}_{\underline{\alpha}} f = \deg_{\underline{\alpha}} g$.

Proof. If $g = 0$, then the lemma is true for trivial reasons. Otherwise, by definition, $\operatorname{cdeg}_{\underline{\alpha}} f \le \deg_{\underline{\alpha}} g$ for all $\underline{\alpha} \in \mathbb{R}^n$. Suppose there exists a vector $\underline{\alpha}_0 \in \mathbb{R}^n$ with $\operatorname{cdeg}_{\underline{\alpha}_0} f < \deg_{\underline{\alpha}_0} g$. For $[f] \overset{\mathrm{cyc}}{\sim} f$ we have $\operatorname{cdeg}_{\underline{\alpha}_0} f = \deg_{\underline{\alpha}_0} [f] < \deg_{\underline{\alpha}_0} g =: 2\Delta \ne 0$. Let p_i be the homogeneous part of g_i with $\underline{\alpha}_0$-degree equal to Δ and $r_i = g_i - p_i$. Then $\deg_{\underline{\alpha}_0} r_i < \Delta$ and

$$\begin{aligned}[f] \overset{\mathrm{cyc}}{\sim} \sum g_i^* g_i &= \sum (p_i + r_i)^*(p_i + r_i) \\ &= \sum p_i^* p_i + \sum p_i^* r_i + \sum r_i^* p_i + \sum r_i^* r_i. \end{aligned} \tag{3.4}$$

Since each word w in $p_i^* r_i$, $r_i^* p_i$, and $r_i^* r_i$ has $\deg_{\underline{\alpha}_0} w < 2\Delta$, none of these can be cyclically equivalent to a nontrivial word in $p_i^* p_i$, where each nontrivial word in $p_i^* p_i$ has $\underline{\alpha}_0$-degree equal to $2\Delta \ne 0$ (note that for each i, either $p_i^* p_i \overset{\mathrm{cyc}}{\nsim} 0$ or $p_i = 0$ due to Lemma 1.55). Similarly, by assumption there is no word in $[f]$ with $\underline{\alpha}_0$-degree equal to 2Δ. Thus

$$0 \overset{\mathrm{cyc}}{\sim} \sum p_i^* p_i, \quad [f] \overset{\mathrm{cyc}}{\sim} \sum p_i^* r_i + \sum r_i^* p_i + \sum r_i^* r_i.$$

However, Lemma 1.55 implies that $p_i = 0$ for all i contradicting that $\deg_{\underline{\alpha}_0} g = 2\Delta$. ■

Proof (of Theorem 3.10). If $f \overset{\mathrm{cyc}}{\sim} g = \sum_i g_i^* g_i \in \Sigma^2$, then $\deg_{\underline{\alpha}} g = \operatorname{cdeg}_{\underline{\alpha}} f$ for all $\underline{\alpha} \in \mathbb{R}^n$, as follows from Lemma 3.12. Therefore,

$$2\deg_{\underline{\alpha}} g_i \le \deg_{\underline{\alpha}} g = \operatorname{cdeg}_{\underline{\alpha}} f$$

for all i and for all $\underline{\alpha} \in \mathbb{R}^n$, hence g_i contains only words satisfying (3.3). Write $g_i = G_i^T \mathbf{W}$, where G_i^T is the (row) vector consisting of the coefficients of g_i. Then $g_i^* g_i = \mathbf{W}^* G_i G_i^T \mathbf{W}$ and, by setting $G := \sum_i G_i G_i^T$, property (3.2) clearly holds.

The inverse of this claim is obvious. Given a positive semidefinite $G \in \mathbb{R}^{N\times N}$ of rank r satisfying (3.2), write $G = \sum_{i=1}^{r} G_i G_i^T$ for $G_i \in \mathbb{R}^{N\times 1}$. Defining $g_i := G_i^T \mathbf{W}$ yields $f \overset{\text{cyc}}{\sim} \sum_{i=1}^{r} g_i^* g_i$. ■

A matrix G satisfying (3.2) is called a *tracial Gram matrix* for f, which motivates the name of the method. For an nc polynomial $f \in \mathbb{R}\langle \underline{X} \rangle$ the tracial Gram matrix is in general *not* unique, hence determining whether $f \in \Theta^2$ amounts to finding *a* positive semidefinite tracial Gram matrix from the affine set of all tracial Gram matrices for f. Problems like this can in theory be solved *exactly* using quantifier elimination, but this only works for problems of small size. Therefore we follow as in the (usual) Gram matrix method a numerical approach in practice. Thus we turn to semidefinite programming, which has become a standard tool in mathematical optimization in the last two decades. The reader not familiar with this topic is referred to Sect. 1.13 and to the references therein.

Following Theorem 3.10 we must determine whether there exists a positive semidefinite matrix G such that $f \overset{\text{cyc}}{\sim} \mathbf{W}^* G \mathbf{W}$. This is a semidefinite feasibility problem in the matrix variable G, where the constraints $\langle A_i \,|\, G \rangle = b_i$ are essentially Eqs. (1.16).

Example 3.13. Let

$$\begin{aligned} f &= 2XY^2XYX + 4XYX^2YX + XY^4X + 2YXY^2X^2 \\ &= (Y^2X + 2XYX)^*(Y^2X + 2XYX) - 2XYXY^2X + 2YXY^2X^2 \\ &\overset{\text{cyc}}{\sim} (Y^2X + 2XYX)^*(Y^2X + 2XYX). \end{aligned}$$

If we take $\mathbf{W} = \begin{bmatrix} XYX & Y^2X \end{bmatrix}^T$, then a tracial (positive semidefinite) Gram matrix G for f is obtained as a solution to the following semidefinite program (SDP):

$$\begin{array}{rl} & \inf \ \langle C \,|\, G \rangle \\ & \text{s. t.} \\ XYX^2YX: & G_{1,1} = 4 \\ XYXY^2X: & G_{1,2} = 2 \\ XY^2XYX: & G_{2,1} = 2 \\ XY^4X: & G_{2,2} = 1 \\ & G \succeq 0. \end{array}$$

Remark 3.14. The matrix C in Example 3.13 is arbitrary. One can use $C = I$, a commonly used heuristic for matrix rank minimization [RFP10]. Often, however, a solution of *high-rank* is desired (this is the case when we want to extract *rational* certificates, a topic discussed in Sect. 1.13 and in Examples 3.25 and 3.26 below, see also [CKP15]). Then $C = 0$ is used, since under a strict feasibility assumption the interior point methods yield solutions in the relative interior of the optimal face, which is in our case the whole feasibility set. If strict complementary is additionally provided, the interior point methods lead to the analytic center of the feasibility set [HdKR02]. Even though these assumptions do not always hold for the instances of SDPs we construct, in our experiments the choice $C = 0$ in the objective function almost always gave a solution of higher rank than the choice $C = I$.

Remark 3.15. As we restrict our attention to nc polynomials which are cyclically equivalent to symmetric nc polynomials (the others are clearly not in Θ^2), we may always merge the equations corresponding to a particular word and its involution, e.g., in Example 3.13 we can replace the second and the third equation with a single constraint $G_{1,2} + G_{2,1} = 4$.

We formalize the lesson from Remark 3.15 as follows:

Lemma 3.16. *If $f = \sum_w a_w w \in \Theta^2$, then for every $v \in \langle \underline{X} \rangle$*

$$\sum_{w \overset{\text{cyc}}{\sim} v} a_w = \sum_{w \overset{\text{cyc}}{\sim} v^*} a_w. \tag{3.5}$$

Corollary 3.17. *Given $f \in \mathbb{R}\langle \underline{X} \rangle$ we have*

(i) *If f does not satisfy* (3.5)*, then $f \notin \Theta^2$.*
(ii) *If f satisfies* (3.5)*, then we can determine whether $f \in \Theta^2$ by solving the following SDP with only symmetric constraints:*

$$\begin{array}{lll}\inf & \langle C \,|\, G \rangle & \\ s.\,t. & \displaystyle\sum_{\substack{p,q,\; p^*q \overset{\text{cyc}}{\sim} v \\ \vee\; p^*q \overset{\text{cyc}}{\sim} v^*}} G_{p,q} = \sum_{w \overset{\text{cyc}}{\sim} v} (a_w + a_{w^*}), & \forall v \in \mathbf{W} \\ & G \succeq 0, & \end{array} \tag{CSOHS$_{\text{SDP}}$}$$

where $\mathbf{W}$ is the set of all words from $\langle \underline{X} \rangle$ needed to construct a Θ^2-certificate, i.e., the set from Theorem 3.10.

Thus we are left with the construction of $\mathbf{W}$, which is related to linear programming problems, as is implied by the following lemma:

Lemma 3.18. *Verifying whether $w \in \langle \underline{X} \rangle$ satisfies* (3.3) *is a linear programming problem.*

Proof. Indeed, let $f = \sum a_v v \in \mathbb{R}\langle \underline{X} \rangle$ of degree $\leq 2d$ be given and let $w \in \langle \underline{X} \rangle$ be a word for which we want to verify (3.3). Then the following is true:

$$\begin{aligned} & 2\deg_{\underline{\alpha}} w \;\leq\; \mathrm{cdeg}_{\underline{\alpha}} f \quad \text{for all } \underline{\alpha} \in \mathbb{R}^n \\ \Leftrightarrow\; & 2\deg_{\underline{\alpha}} w \;\leq\; \deg_{\underline{\alpha}} [f] \quad \text{for all } \underline{\alpha} \in \mathbb{R}^n \\ \Leftrightarrow\; & 2\langle \underline{\alpha} \,|\, \underline{d}_w \rangle \;\leq\; \max_{v \in \mathrm{cc}([f])} \{\langle \underline{\alpha} \,|\, \underline{d}_v \rangle\} \quad \text{for all } \underline{\alpha} \in \mathbb{R}^n \\ \Leftrightarrow\; & 0 \;\leq\; \inf_{\underline{\alpha} \in \mathbb{R}^n} \max_{v \in \mathrm{cc}([f])} \{\langle \underline{\alpha} \,|\, \underline{d}_v - 2\underline{d}_w \rangle\} \\ \Leftrightarrow\; & 0 \;\leq\; \inf\{t \mid \langle \underline{\alpha} \,|\, \underline{d}_v - 2\underline{d}_w \rangle \;\leq\; t,\; \forall v \in \mathrm{cc}([f]),\; \underline{\alpha} \in \mathbb{R}^n\}. \end{aligned}$$

Algorithm 3.1: The tracial Gram matrix method for finding Θ^2-certificates

```
    Input: $f \in \mathbb{R}\langle \underline{X} \rangle$ with $f = \sum_{w \in \langle \underline{X} \rangle} a_w w$, where $a_w \in \mathbb{R}$;
1   If $f$ does not satisfy (3.5), then $f \notin \Theta^2$. Stop;
2   Construct $\mathbf{W}$;
3   Construct data $A_v, \mathbf{b}, C$ corresponding to (CSOHS$_{\text{SDP}}$);
4   Solve (CSOHS$_{\text{SDP}}$) to obtain $G$. If it is not feasible, then $f \notin \Theta^2$. Stop;
5   Compute a decomposition $G = R^T R$;
    Output: Sum of hermitian squares cyclically equivalent to $f$: $f \overset{\text{cyc}}{\sim} \sum_i g_i^* g_i$, where $g_i$
            denotes the $i$th component of $R\mathbf{W}$;
```

Verifying the last inequality can be done in two steps: (i) solve the linear programming problem

$$\begin{aligned} t_{\text{opt}} = \inf\ & t \\ \text{s.t.}\ & \langle \underline{\alpha} \,|\, \underline{d}_v - 2\underline{d}_w \rangle \ \leq\ t, \quad \forall v \in \operatorname{cc}([f]) \\ & \underline{\alpha} \ \in\ \mathbb{R}^n, \end{aligned} \tag{3.6}$$

and (ii) check if $t_{\text{opt}} \geq 0$. ■

By composing results from this section we obtain an algorithm to determine whether a given nc polynomial is cyclically equivalent to a sum of hermitian squares. We call it the *tracial Gram matrix method* and describe it in Algorithm 3.1:

In Step 5 we can take different decompositions, e.g., a Cholesky decomposition (which is not unique if G is not positive definite), the eigenvalue decomposition, etc.

The implementation of Step 2 of the tracial Gram matrix method requires, according to Lemma 3.18, solving a small linear programming problem (3.6) for each candidate w for the set $\mathbf{W}$. Each linear program has $n+1$ variables with $\operatorname{card}(\operatorname{cc}([f]))$ linear inequalities. Solving such linear programs can be done easily for the problems we are interested in (note that due to limitations related to SDP solvers and to symbolic operations over lists of monomials we usually consider only nc polynomials f with $n+d \leq 20$ and with not many words). If f is an nc polynomial in 2 variables and has 10.000 monomials, then we obtain a linear program (LP) in 3 variables with at most 10.000 constraints. Nowadays LP solvers solve such problems easily (within a second); see [Mit03] for a comparison of the state-of-the-art LP solvers and [MPRW09] for a list of efficient alternative methods to solve semidefinite programs.

If $f \in \mathbb{R}\langle \underline{X} \rangle$ is a polynomial in n variables with $\deg f = 2d$, then it is enough to consider at Step 2 only words $w \in \langle \underline{X} \rangle$ such that $[w]$ has degree at most d. Since there are $\binom{n+d}{d}$ different $[w]$ of this type, Step 2 might be still time consuming.

Algorithm 3.2: The Newton cyclic chip method

```
  Input: $f \in \mathbb{R}\langle \underline{X} \rangle$ with $\deg f \leq 2d, f = \sum_{w \in \langle \underline{X} \rangle} a_w w$, where $a_w \in \mathbb{R}$;
1 Let $\mathbf{V}_d$ be the vector of all monomials in $[\underline{x}]$ with degree $\leq d$;
2 $W := \varnothing$;
3 for every $w \in \mathbf{V}_d$ do
4     Solve (3.6) to obtain $t_{\text{opt}}$;
5     if $t_{opt} \geq 0$ then
6         $W = W \cup \{$all (non-commutative) permutations of $w\}$;
7     end
8 end
9 Sort $W$ in a lexicographic order and transform it into the vector $\mathbf{W}$;
  Output: $\mathbf{W}$;
```

We present the details of the implementation of Step 2 of Algorithm 3.1 in Algorithm 3.2 below (the *Newton cyclic chip method*).

Remark 3.19. As mentioned above we need to run the **for** loop in Algorithm 3.2 $\binom{n+d}{d}$-times. For each monomial w which satisfies the condition in Step 5 we add at most $d!$ words to $\mathbf{W}$ in Step 6. Nevertheless, the length of the constructed $\mathbf{W}$ is usually much smaller than the number of all words $w \in \langle \underline{X} \rangle$ of degree $\leq d$. On the other hand, it is often much larger than the number of words obtained by the Newton chip method (see Algorithm 2.2) developed for the sum of hermitian squares decomposition.

3.5 Implementation

Coding the tracial Gram matrix method together with the Newton cyclic chip method needs to be done carefully due to several potential bottlenecks. Obviously the most expensive part of the Gram matrix method is solving ($\text{CSOHS}_{\text{SDP}}$) in Step 4. Its complexity is determined by the order of the matrix variable G and the number of linear equations. Both parameters are strongly related to the vector $\mathbf{W}$ from Step 2. Indeed, the order of G is exactly the length $|\mathbf{W}|$ and the number of linear equations is at least $\frac{|\mathbf{W}|^2}{(d+1)(2d-1)!}$. This follows from the fact that for each product $u^*v,\ u,v \in \mathbf{W}$ there are at most $d+1$ pairs u_i, v_i such that $u_i^* v_i = u^*v$ and at most $(2d-1)!$ cyclically equivalent products.

The vector $\mathbf{W}$ constructed by the Newton cyclic chip method is in general the best possible and is the default procedure used by `NCcycSos` in our package `NCSOStools` [CKP11]. If we know in advance that it is enough to consider in ($\text{CSOHS}_{\text{SDP}}$) only constraints corresponding to words from a (short) vector $\mathbf{V}$, then we can use this $\mathbf{V}$ as an input to Algorithm 3.1 and skip Step 2 of Algorithm 3.1.

Remark 3.20. In a special case we can construct a further reduced vector $\mathbf{W}$. Namely, if we know that for a representation $f \overset{\text{cyc}}{\sim} g \in \Sigma^2$ we have that $\sum_{w \overset{\text{cyc}}{\sim} v^*v} g_w \neq 0$

for all hermitian squares v^*v appearing in g, then we can construct $\mathbf{W}$ by a slight generalization of the Newton chip method from [KP10]. In this case we take the right chips satisfying (3.3) of all hermitian squares which are cyclically equivalent to words from f instead of all words $w \in \langle \underline{X} \rangle$ satisfying (3.3). This works, e.g., for the BMV polynomials (see Sect. 3.5.2) but does not work for, e.g.,

$$f = 1 - 4XYX + 2X^2 + X^2Y^4X^2 \overset{\mathrm{cyc}}{\sim} 2(XY - X)(YX - X) + (X^2Y^2 - 1)(Y^2X^2 - 1).$$

In fact, the hermitian square $2XY^2X$ cancels with $-X^2Y^2$ and $-Y^2X^2$ and we don't get the necessary words XY and YX in $\mathbf{W}$ by applying the enhancement from this remark.

We point out that in general the semidefinite program (CSOHS$_{\mathrm{SDP}}$) might have no strictly feasible points. Absence of (primal) strictly feasible points might cause numerical difficulties while solving (CSOHS$_{\mathrm{SDP}}$). However, as in Sect. 2.5, we can enforce strong duality which is crucial for all SDP solvers by setting the matrix C in (CSOHS$_{\mathrm{SDP}}$) equal to I (actually any full rank matrix will do). Another source of numerical problems is the infeasibility of (CSOHS$_{\mathrm{SDP}}$), which is the case when $f \notin \Theta^2$. We point out that SDP solvers which are supported by `NCSOStools` have easily overcome these difficulties on all tested instances.

Our implementation of the Newton cyclic chip method is augmented by an additional test used to further reduce the length of $\mathbf{W}$. Indeed, if $w \in \mathbf{W}$ satisfies the following properties:

(1) if $u^*v \overset{\mathrm{cyc}}{\sim} w^*w$ for some $u, v \in \mathbf{W}$, then $u = v$ (i.e., any product cyclically equivalent to w^*w is a hermitian square);
(2) neither w^*w nor any other product cyclically equivalent to w^*w appears in f,

then we can delete w from $\mathbf{W}$, and also all u with $u^*u \overset{\mathrm{cyc}}{\sim} w^*w$. This test is implemented in the script `NCcycSos` and is run before solving (CSOHS$_{\mathrm{SDP}}$). It amounts to finding (iteratively) all equations of the type $\langle A_w | G \rangle = 0$ with A_w nonnegative and diagonal.

3.5.1 *Detecting Members of Θ^2 by* `NCSOStools`

We implemented the tracial Gram matrix method based on the tracial Newton polytope in the `NCcycSos` function which is a part of the `NCSOStools` package. It is demonstrated in the following examples:

Example 3.21. Consider the nc polynomial $f = 4X^2Y^{10} + 2XY^2XY^4 + 4XY^6 + Y^2$ of degree 12. There are 127 words in 2 variables of degree ≤ 6 ($\sigma(2,6) = 127$). Using the Newton cyclic chip method (Algorithm 3.2) we get only 16 monomials and after the additional test mentioned at the end of the previous subsection, we are reduced to only 12 words in $\mathbf{W}$ as we can see with the aid of `NCSOStools` [CKP11]:

```
>> NCvars x y
>> f = 4*x^2*y^10 + 2*x*y^2*x*y^4 + 4*x*y^6 + y^2;
>> pars.precision = 1e-3;
>> [IsCycEq,G,W,sohs,g] = NCcycSos(f, pars)
```

We obtain a numerical confirmation that f is in Θ^2 (`IsCycEq=1`). Moreover, the vector given by tracial Newton chip method is

```
W =
      'y'
      'y*x*y*y'
      'y*y*x*y'
      'y*x*y*y*y'
      'y*y*x*y*y'
      'y*y*y*x*y'
      'x*y*y*y*y*y'
      'y*x*y*y*y*y'
      'y*y*x*y*y*y'
      'y*y*y*x*y*y'
      'y*y*y*y*x*y'
      'y*y*y*y*y*x'
```

We also obtain the vector

```
sohs =
         x*y^5+y+y^5*x
         y*x*y^2
         y^2*x*y
         x*y^5-y^5*x
```

containing nc polynomials g_i with $f \overset{\text{cyc}}{\sim} \sum_i g_i^* g_i =$ `g`.

Example 3.22. The nc polynomial $f = 1 + X^6 + Y^6 + X^3Y^3 + X^5Y - YX^5$ is cyclically equivalent to SOHS since $f \overset{\text{cyc}}{\sim} 1 + \frac{3}{4}X^6 + (\frac{1}{2}X^3 + Y^3)^*(\frac{1}{2}X^3 + Y^3)$. Running `NCcycSos` we obtain a numerical certificate for $f \in \Theta^2$, but we also see that in this case the Newton cyclic chip method does not yield any reduction. Indeed, the vector **W** returned by `NCcycSos` contains all possible words of length at most 3 and there are 15 of them.

3.5.2 BMV Polynomials

In an attempt to simplify the calculation of partition functions of quantum mechanical systems Bessis, Moussa, and Villani [BMV75] conjectured in 1975 that for any two symmetric matrices A, B, where B is positive semidefinite, the function $t \mapsto \operatorname{tr}(e^{A-tB})$ is the Laplace transform of a positive Borel measure with real support.

The conjecture in its original form has been proved recently by Stahl [Sta13]. This permits the calculation of explicit upper and lower bounds of energy levels in multiple particle systems, arising from Padé approximants.

Nevertheless, due to an algebraic reformulation of Lieb and Seiringer, the conjecture/statement still remains interesting. In their 2004 paper [LS04], Lieb and Seiringer have given the following purely algebraic reformulation:

Theorem 3.23. *The BMV conjecture is equivalent to the following statement:*

For all positive semidefinite matrices A and B and all $m \in \mathbb{N}$, the polynomial $p(t) := \operatorname{tr}((A+tB)^m) \in \mathbb{R}[t]$ has only nonnegative coefficients.

The coefficient of t^k in $p(t)$ for a given m is the trace of $S_{m,k}(A,B)$, where $S_{m,k}(A,B)$ is the sum of all words of length m in the letters A and B in which B appears exactly k times. For example,

$$S_{4,2}(A,B) = A^2B^2 + ABAB + AB^2A + BABA + B^2A^2 + BA^2B.$$

$S_{m,k}(X,Y)$ can thus be considered as an nc polynomial for $m \geq k$; it is the sum of all words in two variables X,Y of degree m with degree k in Y. Note that Theorem 3.23 considers polynomials in positive semidefinite matrices, while our definition of trace-positivity relates to symmetric matrices—see Definition 1.60. Since any positive semidefinite matrix is a square of some symmetric matrix, trace-positivity in the BMV conjecture is equivalent to trace-positivity of the polynomials $S_{m,k}$ in squared matrix variables. Thus by the proof of Stahl [Sta13] one derives that for each pair (m,k) with $m \geq k$ we have

$$\operatorname{tr} S_{m,k}(X^2,Y^2) \geq 0.$$

The question which BMV polynomials $S_{m,k}(X^2,Y^2)$ are cyclically equivalent to SOHS was an attempt to prove the BMV conjecture. It was completely resolved (in the positive and in the negative) by Burgdorf [Bur11], Hägele [Häg07], Klep and Schweighofer [KS08], Dykema et al. [CDTA10], Landweber and Speer [LS09], and Cafuta et al. [CKP10].

We demonstrate how `NCSOStools` can be used to show the main result of [KS08] establishing $S_{14,6}(X^2,Y^2) \in \Theta^2$. Together with some easier cases and a result of Hillar [Hil07] this implies the BMV conjecture for $m \leq 13$ (which was unsolved at that time).

Example 3.24. Consider the polynomial $f = S_{14,6}(X^2,Y^2)$. To prove (numerically) that $f \in \Theta^2$ with the aid of `NCSOStools` we proceed as follows:

(1) We define two non-commuting variables:

```
>> NCvars x y;
```

(2) Our polynomial f is constructed using `BMVq(14,6)`.

```
>> f=BMVq(14,6);
```

For a numerical test whether $f \in \Theta^2$, we first construct a small monomial vector **V** [KS08, Proposition 3.3] to be used in the Gram matrix method.

```
>> [v1,v2,v3]=BMVsets(14,6); V=[v1;v2;v3];
>> params.obj = 0; params.V=V;
>> [IsCycEq,G,W,sohs,g] = NCcycSos(f, params);
```

This yields a *floating point* positive definite 70×70 Gram matrix `G`. The rest of the output: `IsCycEq` $= 1$ since f is (numerically) in Θ^2; `W` is equal to `V`, `sohs` is a vector of polynomials g_i with $f \overset{\text{cyc}}{\sim} \sum_i g_i^* g_i =$ `g`.

To obtain an *exact* Θ^2-certificate, we can round and project the obtained solution `G` to get a positive semidefinite matrix with rational entries that satisfy liner constraints without any error (see Sect. 1.13 and Example 3.25 for details).

Example 3.25. In this example we demonstrate how to extract a rational certificate by the round and project method, described in Theorem 1.80. Let us consider $f = S_{10,2}(X,Y)$, i.e., the sum of all words of degree 10 in the nc variables X and Y in which Y appears exactly twice. To prove that $f \in \Theta^2$ with the aid of `NCSOStools`, proceed as follows:

(1) Define two non-commuting variables:

```
>> NCvars x y
```

(2) Our nc polynomial f is constructed using `BMV(10,2)`. For a numerical test whether $f \in \Theta^2$, run

```
>> p.obj = 0;
>> f = BMV(10,2);
>> [IsCycEq,G0,W,sohs,g,SDP_data] = NCcycSos(f,p);
```

Using the SDP solver SDPT3, this yields a *floating point* Gram matrix G_0

$$G_0 = \begin{bmatrix} 5.0000 & 2.5000 & -1.8851 & 0.8230 & -0.0899 \\ 2.5000 & 8.7702 & 1.6770 & -2.7313 & 0.8230 \\ -1.8851 & 1.6770 & 10.6424 & 1.6770 & -1.8851 \\ 0.8230 & -2.7313 & 1.6770 & 8.7702 & 2.5000 \\ -0.0899 & 0.8230 & -1.8851 & 2.5000 & 5.0000 \end{bmatrix}$$

for the word vector

$$\mathbf{W} = \begin{bmatrix} X^4Y & X^3YX & X^2YX^2 & XYX^3 & YX^4 \end{bmatrix}^T.$$

The rest of the output: `IsCycEq` $= 1$ since f is (numerically) an element of Θ^2; `sohs` is a vector of nc polynomials g_i with $f \overset{\text{cyc}}{\sim} \sum_i g_i^* g_i =$ `g`; `SDP_data` is the SDP data for ($\text{CSOHS}_{\text{SDP}}$) constructed from f.

(3) To round and project the obtained floating point solution `G0` following Theorem 1.80, feed `G0`, and `SDP_data` into `RprojRldlt`:

```
>> [G,L,D,P,err]=RprojRldlt(G0,SDP_data,true)
```

This produces a rational Gram matrix `G` for f with respect to $\mathbf{W}$ and its LDU decomposition $PLDL^TP^T$, where P is a permutation matrix, L lower unitriangular, and D a diagonal matrix with positive entries. We caution the reader that `L,D`, and `G` are cells, each containing numerators and denominators separately as a matrix. Finally, the obtained rational sum of hermitian squares certificate for $f = S_{10,2}(X,Y)$ is

$$f \overset{\text{cyc}}{\sim} \sum_{i=1}^{5} \lambda_i g_i^* g_i$$

for

$$\begin{aligned}
g_1 &= X^2YX^2 + \frac{7}{44}X^3YX + \frac{7}{44}XYX^3 - \frac{2}{11}X^4Y - \frac{2}{11}YX^4 \\
g_2 &= X^3YX - \frac{577}{1535}XYX^3 + \frac{408}{1535}X^4Y + \frac{188}{1535}YX^4 \\
g_3 &= XYX^3 + \frac{11909}{45984}X^4Y + \frac{7613}{15328}YX^4 \\
g_4 &= X^4Y - \frac{296301}{647065}YX^4 \\
g_5 &= YX^4
\end{aligned}$$

and

$$\lambda_1 = 11, \quad \lambda_2 = \frac{1535}{176}, \quad \lambda_3 = \frac{11496}{1535}, \quad \lambda_4 = \frac{647065}{183936}, \quad \lambda_5 = \frac{1242629}{647065}.$$

Example 3.26. Let us consider $f_{\text{Mot}} = XY^4X + YX^4Y - 3XY^2X + 1$, a non-commutative version of the well-known Motzkin polynomial from (1.21), and $f = f_{\text{Mot}}(X^3, Y^3) = X^3Y^{12}X^3 + Y^3X^{12}Y^3 - 3X^3Y^6X^3 + 1$. To prove that $f \in \Theta^2$ with the aid of `NCSOStools`, proceed as follows:

(1) Define two non-commuting variables and the nc polynomial f:

```
>> NCvars x y
>> f = x^3*y^12*x^3+y^3*x^12*y^3-3*x^3*y^6*x^3 + 1;
```

(2) Define a custom vector of monomials $\mathbf{W}$

```
>> W = {''; 'x*y*y'; 'x*x*y'; 'x*x*y*y*y*y';
 'x*x*x*x*y*y'; 'x*x*x*y*y*y*y*y*y';
           'x*x*x*x*y*y*y*y*y';
 'x*x*x*x*x*y*y*y*y'; 'x*x*x*x*x*x*y*y*y'};
```

(3) For a numerical test whether $f \in \Theta^2$, run

```
>> param.V = W;
>> [IsCycEq,G0,W,sohs,g,SDP_data] = NCcycSos(f,param);
```

This yields a floating point Gram matrix G_0 that is *singular*.

(4) Try to round and project the obtained floating point solution `G0`, feed `G0`, and `SDP_data` into `RprojRldlt`:

```
>> [G,L,D,P,err] = RprojRldlt(G0,SDP_data)
```

This exits with an error, since, unlike in Example 3.25, the rounding and projecting alone does not yield a rational feasible point.

(5) Instead, let us reexamine G_0. A detailed look at the matrix reveals three null vectors. We thus run our interactive procedure which aids the computer in reducing the size of the SDP as in Theorem 1.81.

```
>> [G,SDP_data] = fac_reduct(f,param)
```

This leads the computer to return a floating point feasible point $G_0 \in \mathbb{R}^{9\times 9}$ and the data for this SDP, `SDP_data`. It also stays in interactive mode and the user can inspect the matrix and enter the null vector z to be used in the dimension reduction. We feed in three nullvectors as a matrix of three columns:

```
K>> z = [0 -1 0; -1 0 0; 0 0 1; 0 -1 0; 0 -1 0;
-1 0 0; 0 0 1; -1 0 0; 0 0 1];
K >> return
```

Inside the interactive routine this enables the computer to produce a positive definite feasible $\hat{G}_0 \in \mathbb{R}^{6\times 6}$. Hence we exit the interactive routine.

```
K>> stop = 1; return
```

Now, `NCSOStools` [CKP11] uses $\hat{G}_0$ to produce a rational positive semidefinite Gram matrix G for f, which proves $f \in \Theta^2$. As in Example 3.25, the solution G returned by the `fac_reduct` is a cell containing two 9×9 matrices with numerators and denominators of the rational entries of G. The reader can verify that $f \overset{\text{cyc}}{\sim} \mathbf{W}^* G \mathbf{W}$ exactly by doing rational arithmetic or approximately by computing floating point approximation for G and using floating point arithmetic.

(6) To compute the LDU decomposition $PLDL^TP^T$ for the rational Gram matrix `G` of f with respect to $\mathbf{W}$ (where `G,L,D` are cells, each containing numerators and denominators separately as a matrix) run

```
>> [L,D,P] = Rldlt(G)
```

The obtained rational sum of hermitian squares certificate for $f_{\text{Mot}}(X^3,Y^3)$ is then

$$f_{\text{Mot}}(X^3,Y^3) \overset{\text{cyc}}{\sim} \sum_{i=1}^{6} \lambda_i g_i^* g_i$$

for

$$
\begin{aligned}
g_1 &= 1 - \frac{1}{2}X^2Y^4 - \frac{1}{2}X^4Y^2 \\
g_2 &= XY^2 - \frac{1}{2}X^3Y^6 - \frac{1}{2}X^5Y^4 \\
g_3 &= X^2Y - \frac{1}{2}X^4Y^5 - \frac{1}{2}X^6Y^3 \\
g_4 &= X^2Y^4 - X^4Y^2 \\
g_5 &= X^3Y^6 - X^5Y^4 \\
g_6 &= X^4Y^5 - X^6Y^3
\end{aligned}
$$

and

$$\lambda_1 = \lambda_2 = \lambda_3 = 1, \quad \lambda_4 = \lambda_5 = \lambda_6 = \frac{3}{4}.$$

Remark 3.27. We point out that this yields a rational sum of squares certificate for $\check{f}(x^3, y^3)$ where $\check{f}(x,y) = 1 + x^4y^2 + x^2y^4 - 3x^2y^2$ is the commutative Motzkin polynomial.

References

[BMV75] Bessis, D., Moussa, P., Villani, M.: Monotonic converging variational approximations to the functional integrals in quantum statistical mechanics. J. Math. Phys. **16**(11), 2318–2325 (1975)

[Bur11] Burgdorf, S.: Sums of hermitian squares as an approach to the BMV conjecture. Linear Multilinear Algebra **59**(1), 1–9 (2011)

[BCKP13] Burgdorf, S., Cafuta, K., Klep, I., Povh, J.: The tracial moment problem and trace-optimization of polynomials. Math. Program. **137**(1–2), 557–578 (2013)

[CKP10] Cafuta, K., Klep, I., Povh, J.: A note on the nonexistence of sum of squares certificates for the Bessis-Moussa-Villani conjecture. J. Math. Phys. **51**(8), 083521, 10 (2010)

[CKP11] Cafuta, K., Klep, I., Povh, J.: NCSOStools: a computer algebra system for symbolic and numerical computation with noncommutative polynomials. Optim. Methods. Softw. **26**(3), 363–380 (2011). Available from http://ncsostools.fis.unm.si/

[CKP15] Cafuta, K., Klep, I., Povh, J.: Rational sums of hermitian squares of free noncommutative polynomials. Ars Math. Contemp. **9**(2), 253–269 (2015)

[CLR95] Choi, M.-D., Lam, T.Y., Reznick, B.: Sums of squares of real polynomials. In: *K*-Theory and Algebraic Geometry: Connections with Quadratic Forms and Division Algebras (Santa Barbara,CA, 1992). Proceedings of Symposia in Pure Mathematics, vol. 58, pp. 103–126. American Mathematical Society, Providence (1995)

[CDTA10] Collins, B., Dykema, K.J., Torres-Ayala, F.: Sum-of-squares results for polynomials related to the Bessis-Moussa-Villani conjecture. J. Stat. Phys. **139**(5), 779–799 (2010)

[Häg07] Hägele, D.: Proof of the cases $p \leq 7$ of the Lieb-Seiringer formulation of the Bessis-Moussa-Villani conjecture. J. Stat. Phys. **127**(6), 1167–1171 (2007)

[HdKR02] Halická, M., de Klerk, E., Roos, C.: On the convergence of the central path in semidefinite optimization. SIAM J. Optim. **12**(4), 1090–1099 (2002)

[Hil07] Hillar, C.J.: Advances on the Bessis-Moussa-Villani trace conjecture. Linear Algebra Appl. **426**(1), 130–142 (2007)

[KP10] Klep, I., Povh, J.: Semidefinite programming and sums of hermitian squares of noncommutative polynomials. J. Pure Appl. Algebra **214**, 740–749 (2010)

[KS08] Klep, I., Schweighofer, M.: Sums of hermitian squares and the BMV conjecture. J. Stat. Phys **133**(4), 739–760 (2008)

[LS09] Landweber, P.S., Speer, E.R.: On D. Hägele's approach to the Bessis-Moussa-Villani conjecture. Linear Algebra Appl. **431**(8), 1317–1324 (2009)

[LS04] Lieb, E.H., Seiringer, R.: Equivalent forms of the Bessis-Moussa-Villani conjecture. J. Stat. Phys. **115**(1–2), 185–190 (2004)

[MPRW09] Malick, J., Povh, J., Rendl, F., Wiegele, A.: Regularization methods for semidefinite programming. SIAM J. Optim. **20**(1), 336–356 (2009)

[Mit03] Mittelmann, H.D.: An independent benchmarking of SDP and SOCP solvers. Math. Program. B **95**, 407–430 (2003). http://plato.asu.edu/bench.html

[Par03] Parrilo, P.A.: Semidefinite programming relaxations for semialgebraic problems. Math. Program. **96**(2, Ser. B), 293–320 (2003)

[RFP10] Recht, B., Fazel, M., Parrilo, P.A.: Guaranteed minimum-rank solutions of linear matrix equations via nuclear norm minimization. SIAM Rev. **52**(3), 471–501 (2010)

[Rez78] Reznick, B.: Extremal PSD forms with few terms. Duke Math. J. **45**(2), 363–374 (1978)

[Sta13] Stahl, H.R.: Proof of the BMV conjecture. Acta Math. **211**(2), 255–290 (2013)

Chapter 4
Eigenvalue Optimization of Polynomials in Non-commuting Variables

4.1 Introduction

In Sect. 1.6 we introduced a natural notion of positivity that corresponds exactly to nc polynomials that are SOHS. Recall that an nc polynomial is *positive semidefinite* if it yields a positive semidefinite matrix when we replace the letters (variables) in the polynomial by symmetric matrices of the same order. Helton's Theorem 1.30 implies that positive semidefinite polynomials are exactly the SOHS polynomials, the set of which we denoted by Σ^2.

In this chapter we consider the following question: what is the smallest eigenvalue that a given nc polynomial can attain on a tuple of symmetric matrices from a given semialgebraic set? This is in general a difficult problem. Inspired by the commutative approach of Henrion and Lasserre [HL05] which has been implemented in GloptiPoly [HLL09] we propose a hierarchy of semidefinite programming problems yielding an increasing sequence of lower bounds for the optimum value that we are interested in.

Contrary to the commutative case we prove that for the unconstrained problems and the constrained problems over the nc ball and the nc polydisc the hierarchy of SDPs is finite, i.e., we can compute the global optimum by solving a single instance of a semidefinite programming problem. If the underlying semialgebraic set is more general but the related quadratic module is still archimedean, then the sequence is converging to the desired optimum. It is even finitely convergent if at some point we find a flat optimal solution of the corresponding SDP. We also prove that in all cases of finite convergence we are able to extract optimizers using the GNS construction, presented in Algorithm 1.1.

© The Author(s) 2016

S. Burgdorf et al., *Optimization of Polynomials in Non-Commuting Variables*,
SpringerBriefs in Mathematics, DOI 10.1007/978-3-319-33338-0_4

4.2 Unconstrained Optimization

4.2.1 Unconstrained Optimization as a Single SDP

We start with the question how close to (or how far from) positive semidefiniteness a given nc polynomial is. More precisely, given $f \in \operatorname{Sym}\mathbb{R}\langle \underline{X}\rangle$ of degree $2d$, what is its smallest eigenvalue:

$$\lambda_{\min}(f) \;:=\; \inf\bigl\{\langle f(\underline{A})\mathbf{v} \,|\, \mathbf{v}\rangle \mid \underline{A} \in \mathbb{S}^n,\ \mathbf{v} \text{ a unit vector}\bigr\}. \tag{Eig$_{\min}$}$$

Hence $\lambda_{\min}(f)$ is the greatest lower bound on the eigenvalues of $f(\underline{A})$ taken over all n-tuples $\underline{A}$ of real symmetric matrices of the same order (we denote such tuples by $\mathbb{S}^n$, where n is the number of nc variables). Therefore $(f - \lambda_{\min}(f))(\underline{A}) \succeq 0$ for all $\underline{A} \in \mathbb{S}^n$ and $\lambda_{\min}(f)$ is the largest real number with this property.

Problem (Eig$_{\min}$) is equivalent to (i.e., has equal optimum to):

$$\begin{aligned} \lambda_{\min}(f) \;=\; & \sup\ \lambda \\ & \text{s.t.}\ f(\underline{A}) - \lambda I \succeq 0,\ \forall \underline{A} \in \mathbb{S}^n. \end{aligned} \tag{Eig$'_{\min}$}$$

Helton's Theorem 1.30 implies that (Eig$'_{\min}$) is equivalent to:

$$\begin{aligned} \lambda_{\min}(f) \;=\; & \sup\ \lambda \\ & \text{s.t.}\ f - \lambda \in \Sigma^2_{2d}. \end{aligned} \tag{Eig$^{(d)}_{\mathrm{SDP}}$}$$

This is a semidefinite programming problem which can be explicitly stated as

$$\begin{aligned} & \sup f_1 - \langle E_{1,1} \,|\, F\rangle \\ & \text{s.t.} \qquad f - f_1 = \mathbf{W}_d^*(F - \langle E_{1,1} \,|\, F\rangle E_{1,1})\mathbf{W}_d \\ & \qquad\qquad F \succeq 0. \end{aligned} \tag{Eig$^{(d)}_{\mathrm{SDP}'}$}$$

By f_1 we denote the constant term of f and $E_{1,1}$ is the matrix with all entries 0 except for the $(1,1)$ entry which is 1.

Using standard Lagrange duality approach we obtain the dual of (Eig$^{(d)}_{\mathrm{SDP}}$):

$$\lambda_{\min}(f) = \sup_{f-\lambda\in\Sigma^2_{2d}} \lambda \;=\; \sup_{\lambda}\ \inf_{L\in(\Sigma^2_{2d})^\vee} (\lambda + L(f-\lambda)) \tag{4.1}$$

$$\leq \inf_{L\in(\Sigma^2_{2d})^\vee}\ \sup_{\lambda}\ (\lambda + L(f-\lambda)) \tag{4.2}$$

$$= \inf_{L\in(\Sigma^2_{2d})^\vee} (L(f) + \sup_{\lambda}\ \lambda(1-L(1))) \tag{4.3}$$

$$\begin{aligned} = \ \inf\ & L(f) \\ & L \in (\Sigma_{2d}^2)^\vee \\ & L(1) = 1 \end{aligned} \tag{4.4}$$

$$\begin{aligned} = \ \inf\ & \langle H_L \,|\, G_f \rangle \\ \text{s. t. } & (H_L)_{u,v} = (H_L)_{w,z} \text{ for all } u^*v = w^*z, \\ & (H_L)_{1,1} = 1, \\ & H_L \succeq 0. \end{aligned} \tag{$\mathrm{Eig}_{\mathrm{DSDP}}^{(d)}$}$$

$$=: \ L_{\mathrm{sohs}}$$

The resulting problem ($\mathrm{Eig}_{\mathrm{DSDP}}^{(d)}$) is obviously a semidefinite programming problem. The second equality in (4.1) is a standard transformation of the cone constraint: the inner minimization problem gives optimal value 0 if and only if $f - \lambda \in \Sigma_{2d}^2$. Inequality (4.2) is obvious. The inner problem in (4.3) is bounded (with optimum 0) if and only if $L(1) = 1$, i.e., L is unital. Formulation ($\mathrm{Eig}_{\mathrm{DSDP}}^{(d)}$) is based on the matrix formulation of $(\Sigma_{2d}^2)^\vee$ from Corollary 1.45. The matrix G_f is a Gram matrix for f.

Theorem 4.1. ($\mathrm{Eig}_{\mathrm{SDP}}^{(d)}$) *satisfies strong duality.*

Proof. Clearly, $\lambda_{\min}(f) \le L_{\mathrm{sohs}}$ (weak duality). The dual problem ($\mathrm{Eig}_{\mathrm{DSDP}}^{(d)}$) is always feasible (e.g., $H_L = E_{11}$ is feasible), hence $L_{\mathrm{sohs}} < \infty$.

Suppose first that ($\mathrm{Eig}_{\mathrm{SDP}}^{(d)}$) is feasible, hence $L_{\mathrm{sohs}} \ge \lambda_{\min}(f) > -\infty$. Note that $L(f - L_{\mathrm{sohs}}) \ge 0$ for all L in the dual cone $(\Sigma_{2d}^2)^\vee$. This means that $f - L_{\mathrm{sohs}}$ belongs to the closure of Σ_{2d}^2, so by Proposition 1.20, $f - L_{\mathrm{sohs}} \in \Sigma_{2d}^2$. Hence $L_{\mathrm{sohs}} \le \lambda_{\min}(f)$.

Let us consider the case when ($\mathrm{Eig}_{\mathrm{SDP}}^{(d)}$) is infeasible, i.e., $f \in \mathrm{Sym}\,\mathbb{R}\langle \underline{X} \rangle_{2d}$ is not bounded from below. Then for every $a \in \mathbb{R}$, $f - a$ is not an element of the closed convex cone Σ_{2d}^2. Thus by the Hahn–Banach separation theorem, there exists a linear functional $L : \mathrm{Sym}\,\mathbb{R}\langle \underline{X} \rangle_{2d} \to \mathbb{R}$ satisfying $L(\Sigma_{2d}^2) \subseteq [0, \infty)$, $L(1) = 1$ and $L(f) < a$. As a was arbitrary, this shows that the dual problem ($\mathrm{Eig}_{\mathrm{DSDP}}^{(d)}$) is unbounded, hence strong duality holds in this case as well. ■

We point out that unlike optimization of polynomials in commuting variables which requires a *sequence* of SDPs to compute the minimum, for nc polynomials a single SDP suffices to compute $\lambda_{\min}(f)$ (recall this is an unconstrained optimization).

Corollary 4.2. *Given $f \in \mathbb{R}\langle \underline{X} \rangle_{2d}$, we can compute $\lambda_{\min}(f)$ by solving a single SDP in the primal form* ($\mathrm{Eig}_{\mathrm{SDP}'}^{(d)}$) *or in the dual form* ($\mathrm{Eig}_{\mathrm{DSDP}}^{(d)}$).

Once we have computed the optimal value $\lambda_{\min}(f)$ we may also ask for the n-tuple $\underline{A}$ such that $\lambda_{\min}(f) = \lambda_{\min} f(\underline{A})$. We explain how to compute such $\underline{A}$ in the following section.

4.2.2 Extracting Optimizers for the Unconstrained Case

In this subsection we explain how to find a pair $(\underline{A}, \mathbf{v})$ such that $\lambda_{\min}(f) = \langle f(\underline{A})\mathbf{v} \,|\, \mathbf{v}\rangle$ provided such a pair exists. Flatness of the dual optimal solution is a key property that leads to such a pair.

Proposition 4.3. *Suppose $\lambda_{\min}(f) > -\infty$. If the infimum of* $(\mathrm{Eig}_{\mathrm{DSDP}}^{(d+1)})$ *is attained at a Hankel matrix H_L, then it is also attained at a Hankel matrix $H_{\hat{L}}$ which is 1-flat.*

Proof. Let

$$H_L = \begin{bmatrix} H_d & B \\ B^T & C \end{bmatrix}$$

for some B, C. Here H_d corresponds to rows and columns of H_L labeled by words of length $\leq d$. Since H_L and H_d are positive semidefinite, $B = H_d Z$ and $C \succeq Z^T H_d Z$ for some Z, as follows from Proposition 1.11. Now we form a “new” $H_{\hat{L}}$:

$$H_{\hat{L}} = \begin{bmatrix} H_d & B \\ B^T & Z^T H_d Z \end{bmatrix} = \begin{bmatrix} I & Z \end{bmatrix}^T H_d \begin{bmatrix} I & Z \end{bmatrix}. \tag{4.5}$$

This matrix is obviously 1-flat over H_d, positive semidefinite, and satisfies the nc Hankel condition (it is inherited from H_L since for all quadruples u, v, z, w of words of degree $d+1$ we have $u^*v = z^*w \iff u = z$ and $z = w$). Moreover, we have $\langle H_L \,|\, G_f \rangle = \langle H_{\hat{L}} \,|\, G_f \rangle$. ■

Proposition 4.4. *Let $f \in \mathbb{R}\langle \underline{X} \rangle_{2d}$. Then $\lambda_{\min}(f)$ is attained if and only if there is a feasible solution H_L for* $(\mathrm{Eig}_{\mathrm{DSDP}}^{(d+1)})$ *satisfying $\langle H_L \,|\, G_f \rangle = \lambda_{\min}(f)$.*

Proof. $(\Rightarrow)$ If $\lambda_{\min}(f) = \langle f(\underline{A})\mathbf{v} \,|\, \mathbf{v}\rangle$ holds for some $\underline{A} \in \mathbb{S}^n$ and unit vector $\mathbf{v} \in \mathbb{R}^n$, then $L(p) := \langle p(\underline{A})\mathbf{v} \,|\, \mathbf{v}\rangle$ is the linear functional with Hankel matrix H_L which is the desired feasible solution for $(\mathrm{Eig}_{\mathrm{DSDP}}^{(d+1)})$. $(\Leftarrow)$ By Proposition 4.3, we may assume that H_L is 1-flat over H_d (upper left hand block of H_L corresponding to degree $\leq d$). Now Theorem 1.69 and Algorithm 1.1 apply to the linear functional L corresponding to H_L and yield a tuple $\underline{A}$ of symmetric matrices and a vector $\mathbf{v}$ such that $L(f) = \langle f(\underline{A})\mathbf{v} \,|\, \mathbf{v}\rangle$. By construction, $\|\mathbf{v}\| = \sqrt{\langle \mathbf{v} \,|\, \mathbf{v}\rangle} = \sqrt{L(1)} = 1$. Hence $f(\underline{A})$ has (unit) eigenvector $\mathbf{v}$ with eigenvalue $\lambda_{\min}(f)$. ■

We can extract optimizers for ($\mathrm{Eig}_{\min}$) by the following algorithm:

Algorithm 4.1: Algorithm for finding optimal solutions for ($\mathrm{Eig}_{\min}$)

```
Input: $f \in \mathrm{Sym}\,\mathbb{R}\langle \underline{X}\rangle_{2d}$;
Solve ($\mathrm{Eig}_{\mathrm{DSDP}}^{(d+1)}$);
if the problem is unbounded or the optimum is not attained then
    Stop;
end
Let $H_L$ denote an optimizer. We modify $H_L$ into a 1-flat positive semidefinite matrix $H_{\hat{L}}$ as in (4.5). This matrix yields a positive linear map $\hat{L}$ on $\mathbb{R}\langle \underline{X}\rangle_{2d+2}$ which is 1-flat. In particular, $\hat{L}(f) = L(f) = \lambda_{\min}(f)$;
Use the finite dimensional GNS construction (Algorithm 1.1) on $\hat{L}$ to compute an $n$-tuple of symmetric matrices $\underline{A}$ and a vector $\mathbf{v}$ with $\hat{L}(f) = \lambda_{\min}(f) = \langle f(\underline{A})\mathbf{v} \,|\, \mathbf{v}\rangle$;
Output: $\hat{L}$, $\underline{A}$, $\mathbf{v}$;
```

In Step 6, to construct symmetric matrix representations $A_i \in \mathbb{R}^{r\times r}$ of the multiplication operators; we calculate their image according to a chosen basis $\mathscr{B}$ for $E = \mathrm{ran}\, H_{\hat{L}}$. To be more specific, the vector $A_i\mathbf{u_1}$, where $u_1 \in \langle \underline{X}\rangle_d$ is the first label in $\mathscr{B}$, can be written as a unique linear combination $\sum_{j=1}^{s} \lambda_j \mathbf{u_j}$ with words u_j labeling $\mathscr{B}$ such that $L\big((u_1X_i - \sum \lambda_j u_j)^*(u_1X_i - \sum \lambda_j u_j)\big) = 0$. Then $\big[\lambda_1 \ \ldots \ \lambda_s\big]^T$ will be the first column of A_i.

Example 4.5. Let $f = Y^2 + (XY-1)^*(XY-1)$. Clearly, $\lambda_{\min}(f) \geq 0$. However, $f\big(1/\varepsilon, \varepsilon\big) = \varepsilon^2$, so $\lambda_{\min}(f) = 0$ and hence $L_{\mathrm{sohs}} = 0$. On the other hand, $\lambda_{\min}(f)$ and the dual optimum L_{sohs} are not attained.

Let us first consider $\lambda_{\min}(f)$. Suppose (A,B) is a pair of matrices yielding a singular $f(A,B)$ and let $\mathbf{v}$ be a null vector. Then

$$B^2\mathbf{v} = 0 \quad \text{and} \quad (AB-I)^*(AB-I)\mathbf{v} = 0.$$

From the former we obtain $B\mathbf{v} = 0$, whence

$$\mathbf{v} = I\mathbf{v} = -(AB-I)\mathbf{v} = 0,$$

a contradiction.

We now turn to the nonexistence of a dual optimizer. Suppose otherwise and let H_L be the optimal solution of ($\mathrm{Eig}_{\mathrm{DSDP}}^{(2)}$) and L be the corresponding linear operator, i.e., $L : \mathrm{Sym}\,\mathbb{R}\langle \underline{X}\rangle_4 \to \mathbb{R}$ with $L(1) = 1$. We extend L to $\mathbb{R}\langle \underline{X}\rangle_4$ by symmetrization. That is,

$$L(p) := \frac{1}{2}L(p+p^*).$$

We note that L induces a semi-scalar product (i.e., a positive semidefinite bilinear form) $(p,q) \mapsto L(p^*q)$ on $\mathbb{R}\langle \underline{X}\rangle_2$ due to the positivity property. Since $L(f)=0$, we have

$$L(Y^2)=0 \quad \text{and} \quad L\big((XY-1)^*(XY-1)\big)=0.$$

Hence by the Cauchy–Schwarz inequality, $L(XY)=L(YX)=0$. Thus

$$0=L\big((XY-1)^*(XY-1)\big)=L\big((XY)^*(XY)\big)+L(1)\geq L(1)=1,$$

a contradiction.

Note that similar situation happens if we consider f as a commutative polynomial with global infimum 0 that is not attained.

4.3 Constrained Eigenvalue Optimization of Non-commutative Polynomials

4.3.1 Approximation Hierarchy

The main problem in constrained eigenvalue optimization of nc polynomials can be stated as follows. Given $f \in \operatorname{Sym}\mathbb{R}\langle \underline{X}\rangle_{2d}$ and a subset $S=\{g_1,g_2,\ldots,g_t\} \subseteq \operatorname{Sym}\mathbb{R}\langle \underline{X}\rangle$, compute

$$\lambda_{\min}(f,S) := \inf\big\{\langle f(\underline{A})\mathbf{v} \,|\, \mathbf{v}\rangle \mid \underline{A} \in \mathscr{D}_S^\infty, \mathbf{v} \text{ a unit vector}\big\}. \qquad (\text{Constr-Eig}_{\min})$$

Hence $\lambda_{\min}(f,S)$ is the greatest lower bound on the eigenvalues of $f(\underline{A})$ taken over all tuples $\underline{A}$ of bounded self-adjoint operators on a separable infinite dimensional Hilbert space which satisfy $g_i(\underline{A}) \succeq 0$ for all $g_i \in S$. That is, $(f-\lambda_{\min}(f,S))(\underline{A}) \succeq 0$ for all $\underline{A} \in \mathscr{D}_S^\infty$, and $\lambda_{\min}(f,S)$ is the largest real number with this property.

Similarly to the unconstrained case we can reformulate (Constr-Eig$_{\min}$) into

$$\begin{aligned} \lambda_{\min}(f,S) = \sup\ & \lambda \\ \text{s.t.}\ & f(\underline{A})-\lambda I \succeq 0,\ \forall \underline{A} \in \mathscr{D}_S^\infty. \end{aligned} \qquad (\text{Constr-Eig}'_{\min})$$

Following Pironio, Navascués, and Acín [PNA10] (see also [CKP12]) we get the hierarchy of primal lower bounds for $\lambda_{\min}(f,S)$, which is essentially based on Proposition 1.25:

$$\begin{aligned} \lambda_{\min}(f,S) \ \geq\ f_{\text{sohs}}^{(s)} := \sup\ & \lambda \\ \text{s.t.}\ & f-\lambda \in M_{S,2s}, \end{aligned} \qquad (\text{Constr-Eig}_{\text{SDP}}^{(s)})$$

for $s \geq d$ (for $s < d$ the problem is infeasible). Recall $M_{S,2s}$ is the truncated quadratic module generated by S—see (1.7). Problem (Constr-Eig$_{\mathrm{SDP}}^{(s)}$) is a semidefinite programming problem, as seen from the following proposition:

Proposition 4.6. *Let $f = \sum_w f_w w \in \operatorname{Sym}\mathbb{R}\langle \underline{X}\rangle_{2d}$ and $S = \{g_1, \dots, g_t\} \subseteq \operatorname{Sym}\mathbb{R}\langle \underline{X}\rangle$ with $g_i = \sum_{w\in\langle \underline{X}\rangle_{\deg g_i}} g^i_w w$. Then $f \in M_{S,2d}$ if and only if there exists a positive semidefinite matrix A of order $\sigma(d)$ and positive semidefinite matrices B^i of order $\sigma(d_i)$ ($d_i = \lfloor d - \deg(g_i)/2 \rfloor$) such that for all $w \in \langle \underline{X}\rangle_{2d}$,*

$$f_w = \sum_{\substack{u,v\in\langle \underline{X}\rangle_d\\ u^*v=w}} A_{u,v} + \sum_i \sum_{\substack{u,v\in\langle \underline{X}\rangle_{d_i}, z\in\langle \underline{X}\rangle_{\deg g_i}\\ u^*zv=w}} g^i_z B^i_{u,v}. \tag{4.6}$$

Proof. We start with the "only if" part. Suppose $f \in M_{S,2d}$, hence there exist nc polynomials $a_j = \sum_{w\in\langle \underline{X}\rangle_d} a^j_w w$ and $b_{i,j} = \sum_{w\in\langle \underline{X}\rangle_{d_i}} b^{i,j}_w w$ such that

$$f = \sum_j a_j^* a_j + \sum_{i,j} b_{i,j}^* g_i b_{i,j}.$$

In particular this means that for every $w \in \langle \underline{X}\rangle_{2d}$ the following must hold:

$$\begin{aligned} f_w &= \sum_j \sum_{\substack{u,v\in\langle \underline{X}\rangle_d\\ u^*v=w}} a^j_u a^j_v u^*v + \sum_{i,j} \sum_{\substack{u,v\in\langle \underline{X}\rangle_{d_i}, z\in\langle \underline{X}\rangle_{\deg g_i}\\ u^*zv=w}} b^{i,j}_u b^{i,j}_v g^i_z u^*zv \\ &= \sum_{\substack{u,v\in\langle \underline{X}\rangle_d\\ u^*v=w}} u^*v \sum_i a^i_u a^i_v + \sum_i \sum_{\substack{u,v\in\langle \underline{X}\rangle_{d_i}, z\in\langle \underline{X}\rangle_{\deg g_i}\\ u^*zv=w}} g^i_z u^*zv \sum_j b^{i,j}_u b^{i,j}_v. \end{aligned}$$

If we define the matrix A of order $\sigma(d)$ and matrices B^i of order $\sigma(d_i)$ by $A_{u,v} = \sum_i a^i_u a^i_v$ and $B^i_{u,v} = \sum_j b^{i,j}_u b^{i,j}_v$, then these matrices are positive semidefinite and satisfy (4.6).

To prove the "if" part we use that A and B^i are positive semidefinite, therefore we can find (column) vectors A_i and $B_{i,j}$ such that $A = \sum_i A_i A_i^T$ and $B^i = \sum_j B_{i,j} B_{i,j}^T$. These vectors yield nc polynomials $a_i = A_i^T \mathbf{W}_{\sigma(d)}$ and $b_{i,j} = B_{i,j}^T \mathbf{W}_{\sigma(d_i)}$, which give a certificate for $f \in M_{S,2d}$. ■

Remark 4.7. The last part of the proof of Proposition 4.6 explains how to construct the certificate for $f \in M_{S,2d}$. First we solve the semidefinite feasibility problem in the variables $A \in \mathbb{S}^+_{\sigma(d)}$, $B^i \in \mathbb{S}^+_{\sigma(d_i)}$ subject to constraints (4.6). Then we compute by Cholesky or eigenvalue decomposition column vectors $A_i \in \mathbb{R}^{\sigma(d)}$ and $B_{i,j} \in \mathbb{R}^{\sigma(d_i)}$ which yield desired polynomial certificates $a_i \in \mathbb{R}\langle \underline{X}\rangle_d$ and $b_{i,j} \in \mathbb{R}\langle \underline{X}\rangle_{d_i}$.

Proposition 4.6 implies that (Constr-Eig$_{\mathrm{SDP}}^{(s)}$) is an SDP. It can be explicitly presented as

$$\begin{aligned} f^{(s)}_{\mathrm{sohs}} &= \sup\ f_1 - A_{1,1} - \textstyle\sum_i g^i_1 B^i_{1,1} \\ \text{s. t.}\quad f_w &= \sum_{\substack{u,v\in\langle \underline{X}\rangle_s\\ u^*v=w}} A_{u,v} + \sum_i \sum_{\substack{u,v\in\langle \underline{X}\rangle_{d_i}, z\in\langle \underline{X}\rangle_{\deg g_i}\\ u^*zv=w}} g^i_z B^i_{u,v} \\ &\quad \text{for all } 1\neq w\in\langle \underline{X}\rangle_{2d}, \\ A &\in \mathbb{S}^+_{\sigma(d)},\ B^i \in \mathbb{S}^+_{\sigma(d_i)}, \end{aligned} \tag{Constr-Eig$^{(s)}_{\mathrm{SDP}'}$}$$

where we use $d_i = \lfloor s - \deg(g_i)/2 \rfloor$ (note that here d_i depends on s).

To construct the dual to (Constr-Eig$^{(s)}_{\mathrm{SDP}'}$) we first consider the dual cone to $M_{S,2s}$.

Lemma 4.8. *Let the constant polynomial* 1 *belong to S. Then*

$$\begin{aligned} M^{\vee}_{S,2s} =& \{L : \mathbb{R}\langle \underline{X}\rangle_{2s} \to \mathbb{R} \mid L \textit{ linear},\ L(p^*gp)\geq 0\ \forall p\in\mathbb{R}\langle \underline{X}\rangle_{d_i},\ g\in S\} \\ \cong& \{H_L \mid H_L\in\mathbb{S}^+_{\sigma(s)}, H^{\Uparrow}_{L,g_i}\in\mathbb{S}^+_{\sigma(d_i)} \forall g_i\in S\}, \end{aligned}$$

where H_L is a Hankel matrix and $H^{\Uparrow}_{L,g_i}$ are localizing matrices corresponding to H_L and S.

By repeating the line of reasoning (4.1)–(4.4) we obtain the dual for (Constr-Eig$^{(s)}_{\mathrm{SDP}}$):

$$\begin{aligned} f^{(s)}_{\mathrm{sohs}} \ \leq\ L^{(s)}_{\mathrm{sohs}} =&\ \inf\langle H_L \,|\, G_f\rangle \\ &\text{s. t. } H_L \text{ satisfies nc Hankel condition} \\ &\quad H_L \in \mathbb{S}^+_{\sigma(s)}, \\ &\quad H^{\Uparrow}_{L,g_i} \in \mathbb{S}^+_{\sigma(d_i)}\ \forall g_i\in S \\ &\quad (H_L)_{1,1} = 1. \end{aligned} \tag{Constr-Eig$^{(s)}_{\mathrm{DSDP}}$}$$

We can prove that the dual problems have Slater points under mild conditions, i.e., $L^{(s)}_{\mathrm{sohs}} = f^{(s)}_{\mathrm{sohs}}$, for all $s\geq d$, cf. [CKP12, Proposition 4.4].

Proposition 4.9. *Suppose $\mathscr{D}_S$ contains an ε-neighborhood of 0. Then the SDP* (Constr-Eig$^{(s)}_{\mathrm{DSDP}}$) *admits Slater points.*

Proof. For this it suffices to find a linear map $L : \mathrm{Sym}\,\mathbb{R}\langle \underline{X}\rangle_{2s} \to \mathbb{R}$ satisfying $L(p^*p) > 0$ for all nonzero $p\in\mathbb{R}\langle \underline{X}\rangle_s$, and $L(b^*g_ib) > 0$ for all nonzero $b\in\mathbb{R}\langle \underline{X}\rangle_{d_i}$. We again exploit the fact that there are no nonzero polynomial identities that hold for all orders of matrices, which was used already in Sect. 1.7.

Let us choose $N > s$ and enumerate a dense subset $\mathscr{U}$ of $N\times N$ matrices from $\mathscr{D}_S$ (for instance, take all $N\times N$ matrices from $\mathscr{D}_S$ with entries in $\mathbb{Q}$), that is,

$$\mathscr{U} = \{\underline{A}^{(k)} := (A^{(k)}_1,\ldots,A^{(k)}_n) \mid k\in\mathbb{N}, \underline{A}^{(k)}\in\mathscr{D}_S(N)\}.$$

To each $\underline{A} \in \mathscr{U}$ we associate the linear map

$$L_{\underline{A}} : \operatorname{Sym}\mathbb{R}\langle \underline{X} \rangle_{2s} \to \mathbb{R}, \qquad f \mapsto \operatorname{tr} f(\underline{A}).$$

Form

$$L := \sum_{k=1}^{\infty} 2^{-k} \frac{L_{\underline{A}^{(k)}}}{\|L_{\underline{A}^{(k)}}\|}.$$

We claim that L is the desired linear functional.

Obviously, $L(p^*p) \geq 0$ for all $p \in \mathbb{R}\langle \underline{X} \rangle_s$. Suppose $L(p^*p) = 0$ for some $p \in \mathbb{R}\langle \underline{X} \rangle_s$. Then $L_{\underline{A}^{(k)}}(p^*p) = 0$ for all $k \in \mathbb{N}$, i.e., for all k we have

$$\operatorname{tr} p^*(\underline{A}^{(k)}) p(\underline{A}^{(k)}) = 0,$$

hence $p^*(\underline{A}^{(k)}) p(\underline{A}^{(k)}) = 0$, therefore $p(\underline{A}^{(k)}) = 0$. Since $\mathscr{U}$ was dense in $\mathscr{D}_S(N)$, by continuity it follows that p vanishes on all n-tuples from $\mathscr{D}_S(N)$. Since N was arbitrary p vanishes on all n-tuples from $\mathscr{D}_S$, therefore it vanishes also on an ε-neighborhood of 0 hence $p = 0$ by Lemma 1.36.

Similarly, since free algebra has no zero divisors, $L(b^* g_i b) = 0$ implies $b = 0$ for all $b \in \mathbb{R}\langle \underline{X} \rangle_{d_i}$. ■

Remark 4.10. Having Slater points for (Constr-Eig$^{(s)}_{\mathrm{DSDP}}$) is important for the clean duality theory of SDP to kick in [VB96, dK02]. In particular, there is no duality gap, so $L^{(s)}_{\mathrm{sohs}} = f^{(s)}_{\mathrm{sohs}}$.

Corollary 4.11. *Suppose $\mathscr{D}_S$ contains an ε-neighborhood of 0 and the quadratic module M_S is archimedean. Then the following is true for every $f \in \operatorname{Sym}\mathbb{R}\langle \underline{X} \rangle$:*

$$\lim_{s\to\infty} f^{(s)}_{\mathrm{sohs}} = \lim_{s\to\infty} L^{(s)}_{\mathrm{sohs}} = \lambda_{\min}(f, S) \tag{4.7}$$

Proof. For every $\lambda < \lambda_{\min}(f,S)$ we have $f - \lambda$ is positive definite on $\mathscr{D}_S^{\infty}$, therefore Theorem 1.32 implies that $f - \lambda \in M_S$ which means $f - \lambda \in M_{S,2s_\lambda}$ for appropriate s_λ. Hence $\lambda_{\min}(f,S) \geq f^{(s_\lambda)}_{\mathrm{sohs}} = L^{(s_\lambda)}_{\mathrm{sohs}} \geq \lambda$. Since $\lambda < \lambda_{\min}(f,S)$ was arbitrary, the statement follows. ■

Another very important question is if $L^{(s)}_{\mathrm{sohs}}$ is attained. This is closely related to extraction of optimizers for (Constr-Eig$_{\min}$), as we shall explain in the next section.

4.3.2 Extracting Optimizers

In this section we study two questions: (1) is the convergence in (4.7) finite and (2) can we extract optimizers for (Constr-Eig$_{\min}$), i.e., can we construct $\underline{A} \in \mathscr{D}_S^{\infty}$ and unit vector $\mathbf{v}$ such that $\lambda_{\min}(f,S) = \langle f(\underline{A})\mathbf{v} \mid \mathbf{v} \rangle$? We recall Theorem 1.69 adapted to our situation: a sufficient condition (close to being necessary) for positive answers to both questions is flatness of the optimal solution for (Constr-Eig$_{\mathrm{DSDP}}^{(s)}$) for some $s \geq d$.

Theorem 4.12. *Suppose $\mathscr{D}_S$ contains an ε-neighborhood of* 0. *Let H_L be an optimal solution for* (Constr-Eig$_{\mathrm{DSDP}}^{(s)}$) *for $s \geq d+\delta$, which is δ-flat* ($\delta = \lceil \max_i \deg(g_i)/2 \rceil$). *Then there exist $\underline{A} \in \mathscr{D}_S(r)$ for some r and a unit vector* $\mathbf{v}$ *such that*

$$\lambda_{\min}(f,S) = \langle H_L \,|\, G_f \rangle = \langle f(\underline{A})\mathbf{v} \,|\, \mathbf{v} \rangle.$$

Theorem 4.12 implies that for solving (Constr-Eig$_{\mathrm{SDP}}^{(s)}$) it is crucial to compute a δ-flat optimal solution for (Constr-Eig$_{\mathrm{DSDP}}^{(s)}$) for some $s \geq d+\delta$. Recently Nie [Nie14] presented a hierarchy of semidefinite programming problems, similar to (Constr-Eig$_{\mathrm{DSDP}}^{(s)}$), with a random objective function that under mild conditions converges to a flat solution. Motivated by his ideas we present the following algorithm:

If Algorithm 4.2 returns a flat solution, then we can enter this solution to Algorithm 1.1 to obtain $\underline{A} \in \mathscr{D}_S$ and vector $\mathbf{v}$ such that $\lambda_{\min}(f,S) = \langle f(\underline{A})\mathbf{v}, \mathbf{v}\rangle$.

In Step 7 of Algorithm 4.2 we are solving the following semidefinite program:

$$\begin{array}{ll} \inf & \langle H_L \,|\, R \rangle \\ \text{s.t.} & (H_L)_{u,v} = L(u^*v), \ \text{ for all } u,v \in \langle \underline{X} \rangle_s \\ & (H_L)_{u,v} = L^{(s)}(u^*v), \ \text{ for all } u,v \in \langle \underline{X} \rangle_{s-\delta} \\ & H_L \in \mathbb{S}^+_{\sigma(s)}, \ \ H^{\Uparrow}_{L,g_i} \in \mathbb{S}^+_{\sigma(d_i)}, \ \ \forall i \\ & (H^{\Uparrow}_{L,g_i})_{u,v} = L(u^* g_i v), \ \text{ for all } u,v \in \langle \underline{X} \rangle_{d_i} \\ & L \ \text{ linear functional on } \mathbb{R}\langle \underline{X} \rangle_{2s}. \end{array} \qquad (\text{Constr-Eig}_{\mathrm{RAND}}^{(s)})$$

Algorithm 4.2: Randomized algorithm for finding flat optimal solutions for (Constr-Eig$_{\mathrm{DSDP}}^{(s)}$)

```
   Input: $f \in \operatorname{Sym}\mathbb{R}\langle \underline{X} \rangle_{2d}$, $S = \{g_1,\dots,g_t\}$, $\delta = \lceil \max_i \deg(g_i)/2 \rceil$, $\delta_{\max}$;
1  $H_{\mathrm{flat}} = 0$;
2  for $s = d+\delta, d+\delta+1, \dots, d+\delta+\delta_{\max}$ do
3      Compute $H_L^{(s)}$ – the optimal solution for (Constr-Eig$_{\mathrm{DSDP}}^{(s)}$);
4      if $H_L^{(s)}$ is $\delta$-flat then
5          $H_{\mathrm{flat}} = H_L^{(s)}$. Stop;
6      end
7      Compute $H_{\mathrm{rand}}^{(s)}$ – the optimal solution for (Constr-Eig$_{\mathrm{RAND}}^{(s)}$);
8      if $H_{\mathrm{rand}}^{(s)}$ is $\delta$-flat then
9          $H_{\mathrm{flat}} = H_{\mathrm{rand}}^{(s)}$. Stop;
10     end
11 end
   Output: $H_{\mathrm{flat}}$;
```

The objective function is *random*: we use R which is a random positive definite Gram matrix (corresponding to a random sum of hermitian squares polynomial). In practice (also in the `NCSOStools` [CKP11] implementation) we repeat Step 7 several times since it is cheaper to compute (Constr-Eig$^{(s)}_{\mathrm{RAND}}$) multiple times than going to the next value of s.

The second constraint in (Constr-Eig$^{(s)}_{\mathrm{RAND}}$) implies that the solution L of this problem must coincide with $L^{(s)}$ on $\langle \underline{X}\rangle_{2(s-\delta)}$, i.e., the nc Hankel matrices solving (Constr-Eig$^{(s)}_{\mathrm{DSDP}}$) and (Constr-Eig$^{(s)}_{\mathrm{RAND}}$) have the same upper left-hand corner, indexed by words in $\langle \underline{X}\rangle_{s-\delta}$.

Random polynomials were generated using a sparse random symmetric matrix (with elements coming from a standard normal distribution) of order $\sigma(d)$ with proportion of nonzero elements 0.2, for $n = 2,3$ and $2d = 2,4,6$. We called in Matlab

```
>>R=sprandn(length(W),length(W),0.2);
>>R=R+R';
>>poly = W'*R*W;
```

Here `W` is the vector with all monomials of degree $\leq d$.

Following Nie we expect that Algorithm 4.2 will often find a δ-flat extension. As we reported in [KP16] Algorithm 4.2 almost always returns a flat optimal solution when we optimize a random polynomial over $S = \{1 - \sum_i X_i^4\}$. The flat solution was found in almost all cases in Step 7. We tested δ-flatness of $H^{(s)}_{\mathrm{rand}}$ by comparing the rank of $H^{(s)}_{\mathrm{rand}}$ with the rank of its top left-hand part and by computing $\mathtt{err_{flat}}$ from (1.15).

Once we have a δ-flat solution for (Constr-Eig$^{(s)}_{\mathrm{DSDP}}$) we extract an optimizer, i.e., a pair $\underline{A} \in \mathscr{D}_S(r)$ and $\mathbf{v} \in \mathbb{R}^r$ such that $\lambda_{\min}(f,S) = \langle f(\underline{A})\mathbf{v} \,|\, \mathbf{v}\rangle$, by running the GNS construction, i.e., by executing Step 6 from Algorithm 4.1.

4.4 Constrained Optimization over the Nc Ball and the Nc Polydisc

4.4.1 Approximation Hierarchies Contain Only One Member

In this section we consider constrained eigenvalue optimization over nc semialgebraic sets defined by $\mathbb{B} = \{1 - \sum_i X_i^2\}$ (nc ball) and $\mathbb{D} = \{1 - X_1^2, 1 - X_2^2, \ldots, 1 - X_n^2\}$ (nc polydisc). We denote these semialgebraic sets by $\mathscr{D}_{\mathbb{B}}$ and $\mathscr{D}_{\mathbb{D}}$, respectively (see Definition 1.23).

Remark 4.13. Note that for obvious reasons $\mathscr{D}_{\mathbb{B}}$ and $\mathscr{D}_{\mathbb{D}}$ contain an nc ε-neighborhood of 0, since $\mathscr{D}_{\mathbb{B}}$ itself is an $\mathscr{N}_\varepsilon$ for $\varepsilon = 1$ while $\mathscr{D}_{\mathbb{D}}$ contains $\mathscr{N}_1$. This in particular implies that $M_{\mathbb{B},2d}$ and $M_{\mathbb{D},2d}$ are closed convex cones in the finite dimensional real vector space $\operatorname{Sym}\mathbb{R}\langle \underline{X}\rangle_{2d}$, see Proposition 1.38.

We first prove a stronger version of Proposition 1.25.

Proposition 4.14.

(i) *Suppose* $f = \sum_i g_i^* g_i + \sum_i h_i^*(1 - \sum_j X_j^2)h_i \in M_{\mathbb{B},2d}$. *Then*

$$f|_{\mathscr{D}_{\mathbb{B}}} = 0 \quad \Leftrightarrow \quad g_i = h_i = 0 \textit{ for all } i.$$

(ii) *Suppose* $f = \sum_i g_i^* g_i + \sum_{i,j} h_{i,j}^*(1 - X_j^2)h_{i,j} \in M_{\mathbb{D},2d}$. *Then*

$$f|_{\mathscr{D}_{\mathbb{D}}} = 0 \quad \Leftrightarrow \quad g_i = h_{i,j} = 0 \textit{ for all } i,j.$$

Proof. We only need to prove the ($\Rightarrow$) implication, since ($\Leftarrow$) is obvious. We give the proof of (4.14); the proof of (4.14) is a verbatim copy.

Consider $f = \sum_i g_i^* g_i + \sum_i h_i^*(1 - \sum_j X_j^2)h_i \in M_{\mathbb{B},2d}$ satisfying $f(\underline{A}) = 0$ for all $\underline{A} \in \mathscr{D}_{\mathbb{B}}$. Let us choose $N > d$ and $\underline{A} \in \mathscr{D}_{\mathbb{B}}(N)$. Obviously we have

$$g_i(\underline{A})^T g_i(\underline{A}) \succeq 0 \quad \text{and} \quad h_i(\underline{A})^T(1 - \sum_j A_j^2)h_i(\underline{A}) \succeq 0.$$

Since $f(\underline{A}) = 0$ this yields

$$g_i(\underline{A}) = 0 \quad \text{and} \quad h_i(\underline{A})^T(1 - \sum_j A_j^2)h_i(\underline{A}) = 0 \text{ for all } i.$$

Since N was chosen arbitrary g_i vanishes on $\mathscr{D}_{\mathbb{B}}$. By Remark 4.13 and Lemma 1.35, $g_i = 0$ for all i. Likewise, $h_i^*(1 - \sum_j X_j^2)h_i = 0$ for all i. As there are no zero divisors in the free algebra $\mathbb{R}\langle \underline{X} \rangle$, the latter implies $h_i = 0$. ■

The following theorem drastically simplifies constrained optimization over the nc ball or the nc polydisc.

Theorem 4.15 (Nichtnegativstellensatz). *Let* $f \in \operatorname{Sym} \mathbb{R}\langle \underline{X} \rangle_{2d+1}$.

(i) $f|_{\mathscr{D}_{\mathbb{B}}} \succeq 0$ *if and only if* $f \in M_{\mathbb{B},2d+2}$.
(ii) $f|_{\mathscr{D}_{\mathbb{D}}} \succeq 0$ *if and only if* $f \in M_{\mathbb{D},2d+2}$.

Proof. We prove (i) and leave (ii) as an exercise for the reader. The implication ($\Leftarrow$) is trivial (cf. Proposition 1.25), so we only consider the converse.

Assume $f \notin M_{\mathbb{B},2d+2}$. By Remark 4.13, the truncated quadratic module $M_{\mathbb{B},2d+2}$ is closed. So by the Hahn–Banach theorem there exists a linear functional

$$L : \mathbb{R}\langle \underline{X} \rangle_{2d+2} \to \mathbb{R}$$

satisfying

$$L\big(M_{\mathbb{B},2d+2}\big) \subseteq [0, \infty), \quad L(f) < 0.$$

We modify L by adding to it a small multiple of a linear functional $L^+ : \mathbb{R}\langle \underline{X}\rangle_{2d+2} \to \mathbb{R}$ that is nonnegative on $M_{\mathbb{B},2d+2}$ and strictly positive on Σ^2_{2d}; such an L^+ was constructed in the proof of Proposition 4.9. This new L satisfies

$$L : \mathbb{R}\langle \underline{X}\rangle_{2d+2} \to \mathbb{R}$$

and

$$L\big(M_{\mathbb{B},2d+2}\big) \subseteq [0,\infty), \quad L\big(\Sigma^2_{2d} \setminus \{0\}\big) \subseteq (0,\infty), \quad L(f) < 0. \tag{4.8}$$

Let $\check{L} := L_{2d+1} = L|_{\mathbb{R}\langle \underline{X}\rangle_{2d+1}}$, which is L, restricted to $\mathbb{R}\langle \underline{X}\rangle_{2d+1}$.

There is a positive 1-flat linear functional $\hat{L} : \mathbb{R}\langle \underline{X}\rangle_{2d+2} \to \mathbb{R}$ extending $\check{L}$. To prove this let consider the Hankel matrix H_L presented in block form

$$H_L = \begin{bmatrix} H_{\check{L}} & B \\ B^T & C \end{bmatrix}.$$

The top left block $H_{\check{L}}$ is indexed by words of degree $\leq d$, and the bottom right block C is indexed by words of degree $d+1$. By (4.8), H_L is positive definite.

We shall modify C to make the new matrix flat over $H_{\check{L}}$. Since H_L is positive definite, Schur complement arguments from Proposition 1.11 imply that there exists Z with $B = H_{\check{L}}Z$ and $C \succeq Z^T H_{\check{L}} Z$. Let us form

$$H_{\hat{L}} = \begin{bmatrix} H_{\check{L}} & B \\ B^T & Z^T H_{\check{L}} Z \end{bmatrix}.$$

Then $H_{\hat{L}} \succeq 0$ and $H_{\hat{L}}$ is a 1-flat over $H_{\check{L}}$ by construction. It also satisfies the Hankel condition (1.13), since there are no constraints related to the bottom right block. (Note: this uses the non-commutativity and the fact that we are considering only extensions by 1°.) Thus $H_{\hat{L}}$ is a positive semidefinite Hankel matrix and yields a positive linear functional $\hat{L} : \mathbb{R}\langle \underline{X}\rangle_{2d+2} \to \mathbb{R}$ which is 1-flat (see Remark 1.43).

The linear functional $\hat{L}$ satisfies the assumptions of Theorem 1.27 and Remark 1.28. Hence there is an n-tuple $\underline{A}$ of symmetric matrices of order $s \leq \sigma(d)$ (the order follows from the construction, see the text below) and a vector $\mathbf{v} \in \mathbb{R}^s$ such that

$$\hat{L}(p^* q) = \langle p(\underline{A})\mathbf{v} \,|\, q(\underline{A})\mathbf{v} \rangle$$

for all $p, q \in \mathbb{R}\langle \underline{X}\rangle$ with $\deg p + \deg q \leq 2d$. By linearity,

$$\langle f(\underline{A})\mathbf{v} \,|\, \mathbf{v} \rangle = \hat{L}(f) = L(f) < 0. \tag{4.9}$$

It remains to be seen that $\underline{A}$ is a row contraction, i.e., $1 - \sum_j A_j^2 \succeq 0$. For this we need to recall the construction of the A_j from the proof of Theorem 1.27.

Let $E = \operatorname{ran} H_{\hat{L}}$. There exist s linearly independent columns of $H_{\check{L}}$ labeled by words $w \in \langle \underline{X}\rangle$ with $\deg w \leq d$ which form a basis $\mathscr{B}$ of E. The scalar product on E is induced by $\hat{L}$, and A_i is the left multiplication with X_i on E, i.e., $A_i\colon \mathbf{u} \mapsto A_i\mathbf{u}$ for $u \in \langle \underline{X}\rangle_d$, where $A_i\mathbf{u}$ is the column of $H_{\check{L}}$ or B corresponding to $X_i u \in \langle \underline{X}\rangle_{d+1}$. Let $\mathbf{u} \in E$ be arbitrary. Then there are $\alpha_v \in \mathbb{R}$ for $v \in \langle \underline{X}\rangle_d$ with

$$\mathbf{u} = \sum_{v \in \langle \underline{X} \rangle_d} \alpha_v \mathbf{v}.$$

Write $u = \sum_v \alpha_v v \in \mathbb{R}\langle \underline{X} \rangle_d$. Now compute

$$\begin{aligned}\langle (1 - \sum_j A_j^2)\mathbf{u} \,|\, \mathbf{u} \rangle &= \sum_{v,v' \in \langle \underline{X} \rangle_d} \alpha_v \alpha_{v'} \langle (1 - \sum_j A_j^2)\mathbf{v} \,|\, \mathbf{v}' \rangle \\ &= \sum_{v,v'} \alpha_v \alpha_{v'} \langle \mathbf{v} \,|\, \mathbf{v}' \rangle - \sum_{v,v'} \alpha_v \alpha_{v'} \sum_j \langle A_j \mathbf{v} \,|\, A_j \mathbf{v}' \rangle \\ &= \sum_{v,v'} \alpha_v \alpha_{v'} \hat{L}(v'^* v) - \sum_{v,v'} \alpha_v \alpha_{v'} \sum_j \hat{L}(v'^* X_j^2 v) \\ &= \hat{L}(u^* u) - \sum_j \hat{L}(u^* X_j^2 u) = L(u^* u) - \sum_j \hat{L}(u^* X_j^2 u).\end{aligned} \tag{4.10}$$

Here, the last equality follows from the fact that $\hat{L}_{2d+1} = \check{L} = L_{2d+1} = L|_{\mathbb{R}\langle \underline{X} \rangle_{2d+1}}$. We now estimate the summands $\hat{L}(u^* X_j^2 u)$. By construction the bottom right corner of H_L is greater (w.r.t. to positive semidefiniteness) than the bottom right corner of $H_{\hat{L}}$, therefore

$$\hat{L}(u^* X_j^2 u) = H_{\hat{L}}(X_j u, X_j u) \leq H_L(X_j u, X_j u) = L(u^* X_j^2 u). \tag{4.11}$$

Using (4.11) in (4.10) yields

$$\begin{aligned}\langle (1 - \sum_j A_j^2)\mathbf{u} \,|\, \mathbf{u} \rangle &= L(u^* u) - \sum_j \hat{L}(u^* X_j^2 u) \\ &\geq L(u^* u) - \sum_j L(u^* X_j^2 u) = L\big(u^* (1 - \sum_j X_j^2) u\big) \geq 0,\end{aligned}$$

where the last inequality is a consequence of (4.8).

All this shows that $\underline{A}$ is a row contraction, that is, $\underline{A} \in \mathscr{D}_{\mathbb{B}}$. As in (4.9),

$$\langle f(\underline{A})\mathbf{v} \,|\, \mathbf{v} \rangle = L(f) < 0,$$

contradicting our assumption $f|_{\mathscr{D}_{\mathbb{B}}} \succeq 0$ and finishing the proof of Theorem 4.15. ■

Proposition 4.16. *Let $f \in \operatorname{Sym} \mathbb{R}\langle \underline{X} \rangle_{2d}$. There exists an n-tuple $\underline{A} \in \mathscr{D}_{\mathbb{B}}(\sigma(d))$, and a unit vector $\mathbf{v} \in \mathbb{R}^{\sigma(d)}$ such that*

$$\lambda_{\min}(f, \mathbb{B}) = \langle f(\underline{A})\mathbf{v} \,|\, \mathbf{v} \rangle.$$

In other words, the infimum in (Constr-Eig$_{\min}$) *is really a minimum. An analogous statement holds for $\lambda_{\min}(f, \mathbb{D})$.*

Proof. Note that $f \succeq 0$ on $\mathscr{D}_{\mathbb{B}}$ if and only if $f \succeq 0$ on $\mathscr{D}_{\mathbb{B}}(\sigma(d))$; cf. the proof of Lemma 1.35. Thus in (Constr-Eig$_{\min}$) we are optimizing

$$(\underline{A}, \mathbf{v}) \mapsto \langle f(\underline{A})\mathbf{v} \,|\, \mathbf{v} \rangle \tag{4.12}$$

over $(\underline{A},\mathbf{v}) \in \mathscr{D}_{\mathbb{B}}(\sigma(d)) \times \{\mathbf{v} \in \mathbb{R}^{\sigma(d)} \mid \|\mathbf{v}\| = 1\}$, which is evidently a compact set. Hence by continuity of (4.12) the infimum is attained. The proof for the corresponding statement for $\lambda_{\min}(f,\mathbb{D})$ is the same. ■

Proposition 4.17. *Let $f \in \operatorname{Sym}\mathbb{R}\langle\underline{X}\rangle_{2d+1}$. Then there exist linear functionals*

$$L^{\mathbb{B}}, L^{\mathbb{D}} : \operatorname{Sym}\mathbb{R}\langle\underline{X}\rangle_{2d+2} \to \mathbb{R}$$

which are feasible for (Constr-Eig$^{(s)}_{\mathrm{DSDP}}$) *with $s = d+1$ and $S = \mathbb{B},\mathbb{D}$, respectively, and the following is true:*

$$L^{\mathbb{B}}(f) = \lambda_{\min}(f,\mathbb{B}) \quad \text{and} \quad L^{\mathbb{D}}(f) = \lambda_{\min}(f,\mathbb{D}).$$

Proof. We prove the statement for $L^{\mathbb{B}}$. Proposition 4.16 implies that there exist $\underline{A} \in \mathscr{D}_{\mathbb{B}}(\sigma(d))$ and $\mathbf{v}$ such that $\lambda_{\min}(f,\mathbb{B}) = \langle f(\underline{A})\mathbf{v} \mid \mathbf{v}\rangle$. Let us define $L^{\mathbb{B}}(g) := \langle g(\underline{A})\mathbf{v} \mid \mathbf{v}\rangle$ for $g \in \operatorname{Sym}\mathbb{R}\langle\underline{X}\rangle_{2d+2}$. Then $L^{\mathbb{B}}$ (actually its Hankel matrix) is feasible for (Constr-Eig$^{(s)}_{\mathrm{DSDP}}$) and $L^{\mathbb{B}}(f) = \lambda_{\min}(f,\mathbb{B})$. The same proof works for $\mathbb{D}$. ■

Corollary 4.18. *The hierarchy* (Constr-Eig$^{(s)}_{\mathrm{DSDP}}$) *of lower bounds for $\lambda_{\min}(f,S)$ is finite, when $S = \mathbb{B}$ or $S = \mathbb{D}$. We need to solve only the member of the hierarchy corresponding to $s = d+1$.*

4.4.2 Extracting Optimizers

In this subsection we explain how an optimizer $(\underline{A},\mathbf{v})$ can be extracted from the solutions of the SDPs we constructed in the previous subsection. The explanation is done for $\mathbb{B}$ but the same line of reasoning works for $\mathbb{D}$.

The following theorem demonstrates that Algorithm 4.2 works much better, i.e., is comparable with Algorithm 4.3, when we are optimizing over the nc ball or the nc polydisc.

Theorem 4.19. *If S is the nc ball $\mathbb{B} = \{1 - \sum_j X_j^2\}$ or the nc polydisc $\mathbb{D} = \{1 - X_1^2, \ldots, 1 - X_n^2\}$, then Algorithm 4.2 always finds a 1-flat solution in the first iteration of the for loop.*

Proof. In this case we have $\delta = 1$. Proposition 4.17 implies that for $s = d+1$ the optimal solution $L^{(d+1)}$, computed in Step 3 of Algorithm 4.2, gives value equal to $\lambda_{\min}(f,S)$. Although we can transform it into a 1-flat solution, as explained in Algorithm 4.3, $L^{(d+1)}$ itself is not necessarily 1-flat. In such a case Algorithm 4.2 comes to Step 7 and computes $L^{(d+1)}_{\mathrm{rand}}$. We claim that it is always 1-flat. Let

$$H_{\mathrm{rand}} = \begin{bmatrix} \check{H} & B \\ B^T & C \end{bmatrix}$$

Algorithm 4.3: Extracting optimal solutions for (Constr-Eig$_{\min}$) over $\mathbb{B}$

```
Input: $f \in \mathrm{Sym}\,\mathbb{R}\langle \underline{X}\rangle_{2d+1}$, $S = \mathbb{B}$;
Solve (Constr-Eig$^{(s)}_{\mathrm{DSDP}}$) for $s = d+1$. Let $L$ denote an optimizer, i.e., $L(f) = \lambda_{\min}(f,\mathbb{B})$
    with Hankel matrix $H_L = \begin{bmatrix} H_{\check{L}} & B \\ B^T & C \end{bmatrix}$;
Modify $H_L$: $H_{\hat{L}} = \begin{bmatrix} H_{\check{L}} & B \\ B^T & Z^T H_{\check{L}} Z \end{bmatrix}$, where $Z$ satisfies $H_{\check{L}} Z = B$;
$H_{\hat{L}}$ yields a flat positive linear map $\hat{L}$ on $\mathbb{R}\langle \underline{X}\rangle_{2d+2}$ satisfying $\hat{L}_{2d+1} = L_{2d+1}$. In particular,
    $\hat{L}(f) = L(f) = \lambda_{\min}(f,\mathbb{B})$;
Use the GNS construction (Algorithm 1.1) on $\hat{L}$ to compute $\underline{A} \in \mathscr{D}_{\mathbb{B}}$ and a unit vector $\mathbf{v}$
    with $\hat{L}(f) = \langle f(\underline{A})\mathbf{v} \,|\, \mathbf{v}\rangle = \lambda_{\min}(f,\mathbb{B})$ ;
Output: $\hat{L}$, $\underline{A}$, $\mathbf{v}$;
```

be the (Hankel) matrix, corresponding to $L^{(d+1)}_{\mathrm{rand}}$. Rows of $\check{H}$ and B are labeled by words of length $\le d$ and the rows of B^T and C by words of length $d+1$. Since $H_{\mathrm{rand}} \succeq 0$, we have $B = \check{H}Z$ for some matrix Z and $C \succeq Z^T\check{H}Z$, see (1.14) and the related comments.

Write

$$H_{\mathrm{rand}} = \begin{bmatrix} \check{H} & \check{H}Z \\ Z^T\check{H}^T & Z^T\check{H}Z \end{bmatrix} + \begin{bmatrix} 0 & 0 \\ 0 & C - Z^T\check{H}Z \end{bmatrix}. \tag{4.13}$$

The first matrix is obviously feasible for almost all constraints in (Constr-Eig$^{(s)}_{\mathrm{RAND}}$) and the second is positive semidefinite. The only constraint that is not obvious is $H^{\Uparrow}_{L,g_i} \in \mathbb{S}^{+}_{\sigma(d_i)}$, which is equivalent to nonnegativity of the linear functional on polynomials $p^*(1-\sum_i X_i^2)p$ and $p^*(1-X_j^2)p$ for the nc ball and nc polydisc, respectively.

Let $\tilde{L}$ be the linear functional corresponding to the first matrix on the right-hand side of (4.13). Let H_Δ denote the second matrix in (4.13). Observe that by construction, $\tilde{L}$, $L^{(d+1)}$, and $L^{(d+1)}_{\mathrm{rand}}$ coincide on words of length at most $2d$ (they have the same left upper corner). Then for $p \in \mathbb{R}\langle \underline{X}\rangle_d$,

$$\begin{aligned} \tilde{L}\big(p^*(1-X_i^2)p\big) &= \tilde{L}(p^*p) - \tilde{L}(p^*X_i^2p) \\ &= L^{(d+1)}_{\mathrm{rand}}(p^*p) - \big(L^{(d+1)}_{\mathrm{rand}}(p^*X_i^2p) - (H_\Delta)_{pX_i,pX_i}\big) \\ &= L^{(d+1)}_{\mathrm{rand}}\big(p^*(1-X_i^2)p\big) + (H_\Delta)_{pX_i,pX_i} \ge 0, \end{aligned}$$

whence $\tilde{L}$ is feasible for (Constr-Eig$^{(s)}_{\mathrm{RAND}}$). (We used that $H_\Delta \succeq 0$, a consequence of $C \succeq Z^T\check{H}Z$.) Similar reasoning works for $1 - \sum_i X_i^2$.

In (Constr-Eig$^{(s)}_{\text{RAND}}$) we minimize $\langle H_L \,|\, R\rangle$ for R random positive definite matrix. Therefore

$$\begin{aligned}\langle H_{\text{rand}} \,|\, R\rangle &= \langle H_{\tilde{L}} \,|\, R\rangle + \langle C - Z^T \check{H} Z \,|\, \hat{R}\rangle \\ &\geq \langle H_{\tilde{L}} \,|\, R\rangle.\end{aligned} \tag{4.14}$$

Here $\hat{R}$ is the diagonal block of R corresponding to words of length $d+1$—the bottom right part.

Since $\tilde{L}$ is feasible for (Constr-Eig$^{(s)}_{\text{RAND}}$), the minimum of (4.14) is attained where the second summand is zero. Since R is positive definite this happens if and only if $C = Z^T \check{H} Z$, i.e., $L^{(d+1)}_{\text{rand}} = \tilde{L}$, hence $L^{(d+1)}_{\text{rand}}$ is 1-flat. ■

Remark 4.20. We implemented Algorithm 4.2 in our open source Matlab package `NCSOStools` [CKP11] and numerical evidence corroborates Theorem 4.19. If flatness is checked by computing ranks with accuracy up to 10^{-6}, then we get flat solutions in all the examples we tested. Furthermore, Algorithm 4.2 works very well in practice. It often returns flat solutions when S is archimedean even if it is not the nc ball or nc polydisc. However, see also Example 4.21 below; the question which archimedean S admits flat extensions is difficult.

Example 4.21. Consider S as in Remark 1.24. Then $\mathscr{D}_S$ is empty, M_S is archimedean, and $\mathscr{D}_S^\infty \neq \varnothing$. None of the dual solutions can be flat, as each flat linear functional would yield a point in $\mathscr{D}_S$.

4.5 Implementation

We implemented an algorithm to compute $\lambda_{\min}(f)$ and $\lambda_{\min}(f,S)$ in `NCSOStools` within the function `NCeigMin`. When we are further interested in extracting optimizers we run `NCeigOpt`. The randomized Algorithm 4.2 is coded in `NCeigOptRand`. However, due to special properties of the unconstrained case and of the eigenvalue optimization over the nc ball and nc polydisc, we coded these algorithms (Algorithms 4.1 and 4.3) separately into `NCopt`, `NCoptBall`, and `NCoptCube`, while computing only the optimum values can be also done by `NCmin, NCminBall`, and `NCminCube`.

Example 4.22. Let us consider $f = 2 - X^2 + XY^2X - Y^2$. We first compute $\lambda_{\min}(f)$.

```
>> NCvars x y
>> f = 2 - x^2 + x*y^2*x - y^2;
>> opt = NCeigMin(f);
```

We get a message that f is unbounded from below, i.e., $\lambda_{\min}(f) = -\infty$.

If we want to compute $\lambda_{\min}(f,S)$ over the nc ball, we run

```
>> opt = NCminBall(f);
```

and obtain $\lambda_{\min}(f,\mathbb{B}) = -1$. Finding an optimizer, i.e., a pair $(\underline{A},\mathbf{v})$ with $\underline{A} \in \mathscr{D}_S(r)$ for some r and a unit vector $\mathbf{v}$ such that

$$\langle f(\underline{A})\mathbf{v} \,|\, \mathbf{v}\rangle \;=\; \lambda_{\min}(f,\mathbb{B}) = -1$$

can be done by calling `NCoptBall`:

```
>> [X,fX,eig_val,eig_vec,A,fA]=NCoptBall(f);
```

This gives a matrix X of size 2×25 each of whose rows represents one symmetric 5×5 matrix,

$$\mathtt{A} = \mathrm{reshape}(X(1,:),5,5) = \begin{bmatrix} -0.0000 & 0.7107 & -0.0000 & 0.0000 & 0.0000 \\ 0.7107 & 0.0000 & -0.0000 & 0.3536 & -0.0000 \\ -0.0000 & -0.0000 & -0.0000 & 0.0000 & 0.4946 \\ 0.0000 & 0.3536 & 0.0000 & 0.0000 & 0.0000 \\ 0.0000 & -0.0000 & 0.4946 & 0.0000 & 0.0000 \end{bmatrix}$$

$$\mathtt{B} = \mathrm{reshape}(X(2,:),5,5) = \begin{bmatrix} -0.0000 & 0.0000 & 0.7035 & 0.0000 & 0.0000 \\ 0.0000 & -0.0000 & 0.0000 & -0.0000 & 0.0000 \\ 0.7035 & 0.0000 & 0.0000 & -0.3588 & 0.0000 \\ 0.0000 & -0.0000 & -0.3588 & 0.0000 & -0.0000 \\ 0.0000 & 0.0000 & 0.0000 & -0.0000 & 0.0000 \end{bmatrix}$$

such that

$$f(A,B) = \mathtt{fX} = \begin{bmatrix} 1.0000 & -0.0000 & -0.0000 & 0.0011 & -0.0000 \\ -0.0000 & 1.5091 & -0.0000 & -0.0000 & -0.0000 \\ -0.0000 & -0.0000 & 1.1317 & -0.0000 & -0.0000 \\ 0.0011 & -0.0000 & -0.0000 & 1.7462 & 0.0000 \\ -0.0000 & -0.0000 & -0.0000 & 0.0000 & 1.9080 \end{bmatrix}$$

with eigenvalues $[1.0000, 1.1317, 1.5091, 1.7462, 1.9080]$. So the minimal eigenvalue of $f(A,B)$ is 1 and the corresponding eigenvector is (rounded to four digit accuracy) $\mathbf{v} = [-1.0000\;\; -0.0000\;\; -0.0000\;\; 0.0015\;\; -0.0000]^T$.

Example 4.23. Let us consider $f = XYX$ and $S = \mathbb{D} = \{1-X^2, 1-Y^2\}$. We can write it as

$$f = -2.5 + Y^2 + (1-X^2) + (1-Y^2) + \frac{1}{2}X(1+Y)^2X + \frac{1}{2}X(1-X^2)X + \\ +\frac{1}{2}X(1-Y^2)X + \frac{1}{2}X^4 + \frac{1}{2}(1-X^2),$$

hence $\lambda_{\min}(f,\mathbb{D}) \geq -2.5$. We use NCSOStools

```
>> NCvars x y
>> f = x*y*x;
>> [opt,S,D1,D2,b,SDP,Zd,H,H1] = NCminCube(f)
```

to obtain numerical evidence that $\lambda_{\min}(f,\mathbb{D}) = -1$. By some manual rounding of entries from Zd, which is the optimal solution of (Constr-Eig$^{(s)}_{\text{SDP}'}$), we see

$$f = -1+\alpha(1-X^2)^2+\frac{1}{2}X(1+y)^2X+\beta(1-X^2)+$$
$$+(1-\beta)X(1-X^2)X+\frac{1}{2}X(1-Y^2)X,$$

where $\alpha = 1-\frac{\sqrt{3}}{3}$ and $\beta = \frac{\sqrt{3}}{3}$. Therefore $\lambda_{\min}(f,\mathbb{D}) \geq -1$. We can also extract optimizers by

```
>> [X,fX,eig_val,eig_vec]=NCoptCube(f);
```

which gives us

$$\mathtt{A} = \text{reshape}(X(1,:),5,5) = \begin{bmatrix} 0.0000 & 0.9979 & 0.0000 & -0.0000 & -0.0000 \\ 0.9979 & -0.0000 & 0.0000 & -0.0000 & -0.0646 \\ 0.0000 & 0.0000 & 0.0000 & 0.7470 & -0.0000 \\ -0.0000 & -0.0000 & 0.7470 & -0.0000 & -0.0000 \\ -0.0000 & -0.0646 & -0.0000 & -0.0000 & -0.0000 \end{bmatrix}$$

$$\mathtt{B} = \text{reshape}(X(2,:),5,5) = \begin{bmatrix} -0.0000 & -0.0000 & 0.7880 & -0.0000 & -0.0000 \\ -0.0000 & -1.0000 & -0.0000 & -0.0000 & -0.0000 \\ 0.7880 & -0.0000 & 0.0000 & 0.0000 & 0.4460 \\ -0.0000 & -0.0000 & 0.0000 & -0.0000 & -0.0000 \\ -0.0000 & -0.0000 & 0.4460 & -0.0000 & -0.0000 \end{bmatrix}$$

such that

$$f(A,B) = \mathtt{fX} = \begin{bmatrix} -0.9958 & 0.0000 & -0.0000 & 0.0000 & 0.0645 \\ 0.0000 & -0.0000 & 0.0000 & 0.5659 & -0.0000 \\ -0.0000 & 0.0000 & -0.0000 & 0.0000 & 0.0000 \\ 0.0000 & 0.5659 & 0.0000 & 0.0000 & -0.0000 \\ 0.0645 & -0.0000 & 0.0000 & -0.0000 & -0.0042 \end{bmatrix}$$

which has eigenvalues $[-1.0000 \;\; -0.0000 \;\; 0.5659 \;\; -0.5659 \;\; -0.0000]^T$ and the eigenvector corresponding to smallest eigenvalue -1 is

$$\mathbf{v} = [-0.9979 \;\; 0.0000 \;\; -0.0000 \;\; 0.0000 \;\; 0.0646]^T.$$

This is a numerical certificate that $\lambda_{\min}(f,\mathbb{D}) = -1$.

We point out that the optimum of f (considered as polynomial in commutative variables) over the $[-1,1]^2$ square in $\mathbb{R}^2$ is also -1.

Example 4.24. For $f = 2 - X^2 + XY^2X - Y^2$ and $S = \{4 - X^2 - Y^2, XY + YX - 2\}$ we immediately obtain by calling `NCeigOptRand` a dual optimal solution that is flat ($\mathtt{err}_{\mathtt{flat}} \approx 10^{-7}$):

```
>> NCvars x y
>> f = 2 - x^2 + x*y^2*x - y^2;
>> S = {4-x^2-y^2,x*y+y*x-2};
>> [X,fX,eig_min,flat,err_flat] = NCeigOptRand(f,S,4);
```

This means that we have a numerical certificate that $\lambda_{\min}(f,S) = f_{\mathrm{sohs}}^{(s)} = L_{\mathrm{sohs}}^{(s)} = -1$ for $s = 2,3\ldots$.

Example 4.25. Let us consider the non-commutative version of Motzkin polynomial $f_{\mathrm{Mot}} = XY^4X + YX^4Y - 3XY^2X + 1$, introduced in Example 3.26 and the constraint set $S = \{1 - X^4 - Y^4\}$. Using `NCeigOptRand` we obtain for $s = 3$ a flat dual optimum solution, hence f_{Mot} has minimum eigenvalue over S equal to $f_{\mathrm{sohs}}^{(s)} = L_{\mathrm{sohs}}^{(s)} = -0.160813$ for $s = 3,4,\ldots$.

4.5.1 Application to Quantum Mechanics

We demonstrate how `NCSOStools` can be used to derive upper bounds for Bell inequalities in quantum mechanics. Bell inequalities provide a method to investigate *entanglement*, one of the most peculiar features of quantum mechanics. Entanglement allows two or more parties to be correlated in a non-classical way, and is often studied through the set of bipartite quantum correlations, which consist of the conditional probabilities that two physically separated parties can generate by performing measurements on a shared entangled state.

If we fix a finite number of measurements and outcomes, the set of correlations achievable using classically correlated instructions is a polytope. Hence it can be characterized by its finite number of facets, which correspond to the Bell inequalities. A typical Bell inequality is given as $f = \sum_{i,j} c_{i,j} p(i,j) \leq C$, with coefficients $c_{i,j} \in \mathbb{R}$, conditional probabilities $p(i,j)$, and a constant C depending on the given linear relation. To further understand the possibilities of quantum correlations one is additionally interested in how far one can get beyond the classical Bell inequality, i.e., can one find quantum correlations such that $f > C$ holds, and what is the possible maximum. This subject is known as maximal Bell violation in the literature.

In other words, the goal is to maximize f under the condition that the $p(i,j)$ are generated by the so-called non-local quantum measurements, i.e., for all i,j we have an expression $p(i,j) = \mathbf{v}^T X_i Y_j \mathbf{v}$ with a unit vector $\mathbf{v}$ and self-adjoint operators X_i, Y_j, with the additional constraint that all the $X_i's$ commute with all the $Y_j's$. One often also assumes that the operators are projections. Replacing $p(i,j)$ with this expression we can consider f as an nc polynomial in the variables $\underline{X},\underline{Y}$. Maximizing f is then nothing else than finding the biggest eigenvalue f can attain when running over the $X_i's$ and $Y_i's$. Hence, by considering $-f$, we obtain an eigenvalue minimization problem and can apply the approximation hierarchy using SOHS to derive upper bounds for the maximal violation.

Example 4.26. The most famous inequality is given by the Clauser, Horne, Shimony, and Holt (CHSH) inequality [CHSH69]. Let us assume we have a quantum system consisting of two measurement for each party, each with the two outcomes ± 1. In quantum mechanics the measurements can be modeled by four unitary operators X_1, X_2, Y_1, Y_2, i.e., they satisfy the following conditions: $X_1^2 = 1$, $X_2^2 = 1$, $Y_1^2 = 1$, and $Y_2^2 = 1$. Since we are further interested in the non-local behavior of our quantum system we get as additional constraint that the operators X_i commute with the operators Y_j.

The CHSH inequality is stated in terms of expectation values instead of probabilities, but the concept remains the same. The expectation value $\langle X\rangle$ of an operator X in relation the quantum system $\mathbf{v}$ (given as a unit vector) is defined as $\langle X\rangle = \langle \mathbf{v}^T X \,|\, \mathbf{v}\rangle$. So, the only difference is that the operators are now unitaries instead of projections. Consider the linear relation

$$\langle X_1 Y_1\rangle + \langle X_1 Y_2\rangle + \langle X_2 Y_1\rangle - \langle X_2 Y_2\rangle.$$

This relation is bounded classically by 2, whereas the maximum Bell violation is $2\sqrt{2}$. We now demonstrate the latter bound using `NCSOStools` [CKP11]. First set up the linear relation we are interested in.

```
>> NCvars x1 x2 y1 y2;
>> g = x1*y1+x1*y2+x2*y1-x2*y2;
>> f = (g + g')/2;
```

Then we add the constraints we get from the quantum model of a non-local measurement

```
>> S = {x1^2-1, 1-x1^2, x2^2-1, 1-x2^2, y1^2-1,...
1-y1^2, 1-y2^2, y2^2-1,x1*y1-y1*x1, y1*x1-x1*y1,...
x1*y2-y2*x1, y2*x1-x1*y2, x2*y1-y1*x2,...
y1*x2-x2*y1, x2*y2-y2*x2, y2*x2-x2*y2};
```

Calling

```
>> opt = NCeigMin(-f,S,2);
```

gives then already the desired bound `-opt` $= 2\sqrt{2}$.

We finish by another example where the first level of the hierarchy does not yet give the exact value. To the best of our knowledge the exact value for this example is still unknown.

Example 4.27. Consider the I_{3322}-inequality [CG04], where each party can perform one of three possible measurements (X_1, X_2, X_3 and Y_1, Y_2, Y_3) each with two outcomes. We are interested in finding upper bounds for the following relation of joint probabilities

$$\begin{aligned} &p(X_1,Y_1)+p(X_1,Y_2)+p(X_1,Y_3)+p(X_2,Y_1)+p(X_2,Y_2)-p(X_2,Y_3)\\ &\quad +p(X_3,Y_1)-p(X_3,Y_2)-p(X_1)-2p(Y_1)-p(Y_2), \end{aligned}$$

where we put the operator itself in the argument instead of just the index. Using again that $p(X_i,Y_j)=\mathbf{v}^T X_i Y_j \mathbf{v}$ for some unit vector $\mathbf{v}$, this problem can be written more compactly as maximizing the eigenvalue of f, where f is given by the nc polynomial $X_1(Y_1+Y_2+Y_3)+X_2(Y_1+Y_2-Y_3)+X_3(Y_1-Y_2)-X_1-2Y_1-Y_2$. The semialgebraic set S we are maximizing over is given by the conditions on non-local quantum measurement, i.e., the operators are (positive semidefinite) projections and all $X_i{}'s$ commute with all $Y_j{}'s$.

To get upper bounds using the SOHS hierarchy set up the system

```
>> NCvars x1 x2 x3 y1 y2 y3;
>> g= x1*(y1+y2+y3)+x2*(y1+y2-y3)+x3*(y1-y2) ...
-x1-2*y1-y2;
>> f = (g + g')/2;
>> S = {x1, x2, x3, y1, y2, y3, x1^2-x1, x1-x1^2,...
 x2^2-x2, x2-x2^2,x3^2-x3,x3-x3^2, y1^2-y1,...
 y1-y1^2, y2-y2^2, y2^2-y2,y3^2-y3, y3-y3^2,...
 x1*y1-y1*x1, y1*x1-x1*y1, x1*y2-y2*x1,...
 y2*x1-x1*y2, x1*y3-y3*x1, y3*x1-x1*y3,...
 x2*y1-y1*x2, y1*x2-x2*y1, x2*y2-y2*x2,...
 y2*x2-x2*y2, x2*y3-y3*x2, y3*x2-x2*y3,...
 x3*y1-y1*x3, y1*x3-x3*y1, x3*y2-y2*x3,...
 y2*x3-x3*y2, x3*y3-y3*x3, y3*x3-x3*y3 };
```

Calling

```
>> opt = NCeigMin(-f,S,2);
```

gives the upper bound `0.375`. Calling the next level

```
>> opt = NCeigMin(-f,S,4);
```

gives the bound `0.2509400561`.

These bounds have already been computed by Doherty et al. [DLTW08] where they calculated by hand the dual problem of the original maximization problem and fed this into an SDP solver. `NCSOStools` [CKP11] now provides a direct way to perform these kind of computations. A nice list of derived upper bounds for Bell inequalities using the approximation hierarchy based on SOHS can be found in [PV09].

References

[CKP11] Cafuta, K., Klep, I., Povh, J.: NCSOStools: a computer algebra system for symbolic and numerical computation with noncommutative polynomials. Optim. Methods. Softw. **26**(3), 363–380 (2011). Available from http://ncsostools.fis.unm.si/

[CKP12] Cafuta, K., Klep, I., Povh, J.: Constrained polynomial optimization problems with noncommuting variables. SIAM J. Optim. **22**(2), 363–383 (2012)

[CHSH69] Clauser, J.F., Horne, M.A., Shimony, A., Holt, R.A.: Proposed experiment to test local hidden-variable theories. Phys. Rev. Lett. **23**, 880–884 (1969)

[CG04] Collins, D., Gisin, N.: A relevant two qubit Bell inequality inequivalent to the CHSH inequality. J. Phys. A **37**(5), 1775–1787 (2004)

[dK02] de Klerk, E.: Aspects of Semidefinite Programming. Applied Optimization, vol. 65. Kluwer Academic, Dordrecht (2002)

[DLTW08] Doherty, A.C., Liang, Y.-C., Toner, B., Wehner, S.: The quantum moment problem and bounds on entangled multi-prover games. In: 23rd Annual IEEE Conference on Computational Complexity, 2008. CCC'08, pp. 199–210. IEEE, LOS ALAMOS (2008)

[HL05] Henrion, D., Lasserre, J.-B.: Detecting global optimality and extracting solutions in GloptiPoly. In: Positive Polynomials in Control. Lecture Notes in Control and Information Sciences, vol. 312, pp. 293–310. Springer, Berlin (2005)

[HLL09] Henrion, D., Lasserre, J.-B., Löfberg, J.: GloptiPoly 3: moments, optimization and semidefinite programming. Optim. Methods Softw. **24**(4–5), 761–779 (2009)

[KP16] Klep, I., Povh, J.: Constrained trace-optimization of polynomials in freely noncommuting variables. J. Glob. Optim. **64**, 325–348 (2016)

[Nie14] Nie, J.: The $\mathscr{A}$-truncated $\mathscr{K}$-moment problem. Found. Comput. Math. **14**(6), 1243–1276 (2014)

[PV09] Pál, K.F., Vértesi, T.: Quantum bounds on Bell inequalities. Phys. Rev. A **79**, 022120 (2009)

[PNA10] Pironio, S., Navascués, M., Acín, A.: Convergent relaxations of polynomial optimization problems with noncommuting variables. SIAM J. Optim. **20**(5), 2157–2180 (2010)

[VB96] Vandenberghe, L., Boyd, S.: Semidefinite programming. SIAM Rev. **38**(1), 49–95 (1996)

Chapter 5
Trace Optimization of Polynomials in Non-commuting Variables

5.1 Introduction

In Chap. 3 trace-positivity together with the question how to detect it was explored in details. Due to hardness of the decision problem “Is a given nc polynomial f trace-positive?” we proposed a relaxation of the problem, i.e., we are asking if f is cyclically equivalent to SOHS. The tracial Gram matrix method based on the tracial Newton polytope was proposed (see Sects. 3.3 and 3.4) to efficiently detect such polynomials.

In this chapter we turn our attention to trace optimization of nc polynomials. We are interested in computing the smallest number the trace of a given nc polynomial can attain or approaches over a given nc semialgebraic set of symmetric matrices. This is in general a very difficult question, so we employ approximation tools again and present a *tracial* Lasserre relaxation scheme [Las01, Las09]. It yields again a hierarchy of semidefinite programming problems resulting in an increasing sequence of lower bounds for the optimum value. Finally we also shortly discuss the extraction of optimizers.

5.2 Unconstrained Trace Optimization

The purpose of this section is twofold. First we formulate the unconstrained trace optimization problem and second we present a Lasserre type of approximation hierarchy consisting of semidefinite programming problems. We also explore the duality properties.

Let $f \in \mathbb{R}\langle \underline{X} \rangle$ be given. We are interested in the *trace-minimum* of f, that is,

$$\operatorname{tr}_{\min}(f) := \inf\{\operatorname{tr} f(\underline{A}) \mid \underline{A} \in \mathbb{S}^n\}. \tag{Tr$_{\min}$}$$

© The Author(s) 2016

S. Burgdorf et al., *Optimization of Polynomials in Non-Commuting Variables*,
SpringerBriefs in Mathematics, DOI 10.1007/978-3-319-33338-0_5

This is a hard problem. For instance, a good understanding of trace-positive polynomials is likely to lead to a solution of the Connes' embedding conjecture [Con76], an outstanding open problem from operator algebras; see [KS08]. Another way to see the hardness is due to a result of Ji [Ji13] who proved that deciding whether the quantum chromatic number of a graph is at most three is NP-hard. This problem in turn is a conic optimization problem which is dual to an optimization problem over certain trace-positive polynomials, see [LP15] for details.

We can rewrite ($\mathrm{Tr}_{\min}$) as

$$\operatorname{tr}_{\min}(f) = \sup\{a \mid \operatorname{tr}(f-a)(\underline{A}) \geq 0,\ \forall \underline{A} \in \mathbb{S}^n\}. \tag{$\mathrm{Tr}_{\min'}$}$$

We again assume $\sup \varnothing = -\infty$. Nc polynomials from Θ^2 are trace-positive therefore it is natural to consider the following relaxation of ($\mathrm{Tr}_{\min'}$):

$$\operatorname{tr}_{\Theta^2}(f) := \sup\{a \mid f - a \in \Theta^2_{2d}\}, \tag{$\mathrm{Tr}_{\mathrm{sohs}}$}$$

where $2d = \operatorname{cdeg} f$ (if $\operatorname{cdeg} f$ is an odd number, then $\operatorname{tr}_{\min}(f) = \operatorname{tr}_{\Theta^2}(f) = -\infty$, hence we do not need to consider this case).

Remark 5.1. Since we are only interested in the trace of nc polynomials $f \in \mathbb{R}\langle \underline{X} \rangle$, when evaluated on elements from $\mathbb{S}^n$, $\mathscr{D}_S$, or $\mathscr{D}_S^{\mathrm{II}_1}$ we use that $\operatorname{tr} f(\underline{A}) = \operatorname{tr} f^*(\underline{A})$ for all $\underline{A}$; hence there is no harm in replacing f by its symmetrization $\frac{1}{2}(f + f^*)$. Thus we will focus in this chapter on *symmetric* nc polynomials.

Lemma 5.2. *Let $f \in \operatorname{Sym}\mathbb{R}\langle \underline{X} \rangle$. Then $\operatorname{tr}_{\Theta^2}(f) \leq \operatorname{tr}_{\min}(f)$.*

Proof. Indeed, if $a \in \mathbb{R}$ is such that $f - a \in \Theta^2$, then $0 \leq \operatorname{tr}(f-a) = \operatorname{tr} f - \operatorname{tr} a = \operatorname{tr} f - a$, hence $\operatorname{tr} f \geq a$. ■

In general we do not have equality in Lemma 5.2. For instance, the Motzkin polynomial f satisfies $\operatorname{tr}_{\min}(f) = 0$ and $\operatorname{tr}_{\Theta^2}(f) = \sup \varnothing := -\infty$, see [KS08] and Example 5.14. Nevertheless, $\operatorname{tr}_{\Theta^2}(f)$ gives a solid approximation of $\operatorname{tr}_{\min}(f)$ for most of the examples and is easier to compute. It is obtained by solving an instance of SDP.

Suppose $f \in \operatorname{Sym}\mathbb{R}\langle \underline{X} \rangle$ is of degree $\leq 2d$ (with constant term f_1). Let $\mathbf{W}_d$ be a vector of all words up to degree d with first entry equal to 1. Then ($\mathrm{Tr}_{\mathrm{sohs}}$) rewrites into

$$\begin{array}{ll} \sup f_1 - \langle E_{1,1} \mid F \rangle & \\ \text{s.t.} & f - f_1 \overset{\text{cyc}}{\sim} \mathbf{W}_d^*(G - \langle E_{1,1} \mid F \rangle E_{1,1})\mathbf{W}_d \\ & F \succeq 0. \end{array} \tag{$\mathrm{Tr}_{\mathrm{SDP}}$}$$

Here $E_{1,1}$ is again the matrix with all entries 0 except for the $(1,1)$-entry which is 1. The cyclic equivalence translates into a set of linear constraints, cf. Proposition 1.51.

In general ($\mathrm{Tr}_{\mathrm{SDP}}$) does not satisfy the Slater condition. Nevertheless:

Theorem 5.3. ($\mathrm{Tr}_{\mathrm{SDP}}$) *satisfies strong duality.*

Proof. The proof is essentially the same as that of Theorem 4.1 so is omitted. We only mention an important ingredient is the closedness of the cone Θ^2 which is a trivial corollary of Proposition 1.58. ■

Repeating the Lagrangian procedure from (4.1)–(4.4) we obtain the dual to ($\mathrm{Tr}_{\mathrm{SDP}}$):

$$\begin{aligned} L_{\Theta^2}(f) &= \inf L(f) \\ &\quad \text{s.t. } L(1) = 1 \\ &\qquad\quad L \in (\Theta^2_{2d})^\vee \end{aligned}$$

Following Remark 1.64 we rewrite this problem into an explicit semidefinite programming problem:

$$\begin{aligned} L_{\Theta^2}(f) &= \inf \, \langle H_L \,|\, G_f \rangle \\ &\quad \text{s.t. } (H_L)_{u,v} = (H_L)_{w,z} \text{ for all } u^*v \overset{\mathrm{cyc}}{\sim} w^*z, \\ &\qquad (H_L)_{1,1} = 1, \\ &\qquad\quad H_L \succeq 0. \end{aligned} \tag{$\mathrm{Tr}_{\mathrm{DSDP}}$}$$

Recall that H_L from the SDP above is a tracial Hankel matrix. It is of order $\sigma(d)$. By Theorem 5.3, we have $\mathrm{tr}_{\Theta^2}(f) = L_{\Theta^2}(f)$. The question is, does $\mathrm{tr}_{\Theta^2}(f) = L_{\Theta^2}(f) = \mathrm{tr}_{\min}(f)$ hold? This is true for the case of unconstrained eigenvalue optimization (see Theorem 5.3), while in the unconstrained trace optimization it only holds under additional assumptions. We show that if the optimum solution of ($\mathrm{Tr}_{\mathrm{DSDP}}$) satisfies a flatness condition (see Definitions 1.47 and 1.49), then the answer to the question is affirmative. In particular, the proposed Θ^2-relaxation is then exact. Furthermore, in this case we can even extract global trace-minimizers of f.

Theorem 5.4. *If the optimizer H_L^{opt} of* ($\mathrm{Tr}_{\mathrm{DSDP}}$) *satisfies the* flatness condition*, i.e., the linear functional underlying H_L^{opt} is 1-flat, then the Θ^2-relaxation is exact:*

$$\mathrm{tr}_{\Theta^2}(f) = L_{\Theta^2}(f) = \mathrm{tr}_{\min}(f).$$

Proof. The first equality is strong duality shown in Theorem 5.3. For the second equality, if the linear functional L^{opt} corresponding to H_L^{opt} satisfies the flatness condition, then by Theorem 1.71 there exist finitely many n-tuples $\underline{A}^{(j)}$ of symmetric matrices and positive scalars $\lambda_j > 0$ with $\sum_j \lambda_j = 1$ such that

$$L^{\mathrm{opt}}(f) = \sum_j \lambda_j \mathrm{tr} f(\underline{A}^{(j)}).$$

Hence $L_{\Theta^2}(f) = L^{\mathrm{opt}}(f) \leq \mathrm{tr}_{\min}(f)$ and equality follows from weak duality. ■

5.3 Constrained Trace Optimization

In this section we present the tracial version of Lasserre's relaxation scheme to minimize the trace of an nc polynomial.

Let $S \subseteq \operatorname{Sym}\mathbb{R}\langle \underline{X}\rangle$ be finite and let $f \in \operatorname{Sym}\mathbb{R}\langle \underline{X}\rangle$. We are interested in the smallest trace the polynomial f attains on $\mathscr{D}_S$, i.e.,

$$\operatorname{tr}_{\min}(f,S) := \inf\big\{\operatorname{tr} f(\underline{A}) \mid \underline{A} \in \mathscr{D}_S\big\}. \tag{Constr-Tr$_{\min}$}$$

Hence $\operatorname{tr}_{\min}(f,S)$ is the greatest lower bound on the trace of $f(\underline{A})$ for tuples of symmetric matrices $\underline{A} \in \mathscr{D}_S$, i.e., $\operatorname{tr}(f(\underline{A}) - \operatorname{tr}_{\min}(f,S)\underline{A}) \geq 0$ for all $\underline{A} \in \mathscr{D}_S$, and $\operatorname{tr}_{\min}(f,S)$ is the largest real number with this property.

We introduce $\operatorname{tr}^{\mathrm{II}_1}_{\min}(f,S) \in \mathbb{R}$ as the trace-minimum of f on $\mathscr{D}_S^{\mathrm{II}_1}$. Since $\mathscr{D}_S^{\mathrm{II}_1} \supseteq \mathscr{D}_S$, we have $\operatorname{tr}^{\mathrm{II}_1}_{\min}(f,S) \leq \operatorname{tr}_{\min}(f,S)$. As mentioned in Remark 1.61 (see also Proposition 1.63), $\operatorname{tr}^{\mathrm{II}_1}_{\min}(f,S)$ is more approachable than $\operatorname{tr}_{\min}(f,S)$. In fact, in this section we shall present Lasserre's relaxation scheme producing a sequence of computable lower bounds $\operatorname{tr}^{(s)}_{\Theta^2}(f,S)$ monotonically converging to $\operatorname{tr}^{\mathrm{II}_1}_{\min}(f,S)$. Here, as always, the constraint set S is assumed to produce an archimedean quadratic module M_S. From Proposition 1.62 we can bound $\operatorname{tr}^{\mathrm{II}_1}_{\min}(f,S)$ from below by

$$\begin{aligned}\operatorname{tr}^{(s)}_{\Theta^2}(f,S) := {} & \sup \ \ \lambda \\ & \ \text{s.t.} \ \ f - \lambda \in \Theta^2_{S,2s},\end{aligned} \tag{Constr-Tr$^{(s)}_{\mathrm{SDP}}$}$$

for $2s \geq \operatorname{cdeg} f$. For $2s < \operatorname{cdeg} f$, (Constr-Tr$^{(s)}_{\mathrm{SDP}}$) is infeasible.

For each *fixed* s, (Constr-Tr$^{(s)}_{\mathrm{SDP}}$) is an SDP (see Proposition 5.7 below) and leads to the tracial version of the Lasserre relaxation scheme.

Corollary 5.5. *Let $S \subseteq \operatorname{Sym}\mathbb{R}\langle \underline{X}\rangle$, and let $f \in \operatorname{Sym}\mathbb{R}\langle \underline{X}\rangle$. If M_S is archimedean, then*

$$\operatorname{tr}^{(s)}_{\Theta^2}(f,S) \underset{s\to\infty}{\longrightarrow} \operatorname{tr}^{\mathrm{II}_1}_{\min}(f,S). \tag{5.1}$$

The sequence $\operatorname{tr}^{(s)}_{\Theta^2}(f,S)$ is monotonically increasing and bounded from above, but the convergence in (5.1) *is not finite in general.*

Proof. This follows from Proposition 1.63. For each $m \in \mathbb{N}$, there is $s(m) \in \mathbb{N}$ with

$$f - \operatorname{tr}^{\mathrm{II}_1}_{\min}(f,S) + \frac{1}{m} \in \Theta^2_{S,2s(m)}.$$

In particular,

$$\operatorname{tr}^{(s(m))}_{\Theta^2}(f) \geq \operatorname{tr}^{\mathrm{II}_1}_{\min}(f,S) - \frac{1}{m}.$$

Since also

$$\operatorname{tr}_{\Theta^2}^{(s(m))}(f) \leq \operatorname{tr}_{\min}^{\mathrm{II}_1}(f,S),$$

we obtain

$$\lim_{s\to\infty} \operatorname{tr}_{\Theta^2}^{(s)}(f,S) = \lim_{m\to\infty} \operatorname{tr}_{\Theta^2}^{(s(m))}(f) = \operatorname{tr}_{\min}^{\mathrm{II}_1}(f,S).$$

■

Example 5.6. For a simple example with non-finite convergence, consider

$$p = (1-X^2)(1-Y^2) + (1-Y^2)(1-X^2),$$

and

$$S = \{1-X^2, 1-Y^2\}.$$

Then $\operatorname{tr}_{\min}^{\mathrm{II}_1}(p,S) = 0$, but $p \notin \Theta_S^2$ [KS08, Example 4.3]. The first few lower bounds for $\operatorname{tr}_{\min}^{\mathrm{II}_1}(p,S)$ are in the second column of Table 5.1.

Generally we are interested in $\operatorname{tr}_{\min}^{\mathrm{II}_1}(f,S)$, but there is no good procedure or algorithm for computing it. Therefore we stick to $\operatorname{tr}_{\Theta^2}^{(s)}(f,S)$ since its computational feasibility comes from the fact that verifying whether $f \in \Theta_{S,2s}^2$ is a semidefinite programming feasibility problem when S is finite.

Proposition 5.7. *Let $f = \sum_w f_w w \in \operatorname{Sym}\mathbb{R}\langle \underline{X}\rangle$ and $S = \{g_1,\ldots,g_t\} \subseteq \operatorname{Sym}\mathbb{R}\langle \underline{X}\rangle$ with $g_i = \sum_{w\in\langle \underline{X}\rangle_{\deg g_i}} g_w^i w$. Then $f \in \Theta_{S,2s}^2$ if and only if there exists a positive semidefinite matrix A of order $\sigma(s)$ and positive semidefinite matrices B^i of order $\sigma(s_i)$ (recall that $s_i = \lfloor s - \deg(g_i)/2\rfloor$) such that for all $w \in \langle \underline{X}\rangle_{2s}$,*

$$f_w = \sum_{\substack{u,v\in\langle \underline{X}\rangle_s \\ u^*v \overset{\mathrm{cyc}}{\sim} w}} A_{u,v} + \sum_i \sum_{\substack{u,v\in\langle \underline{X}\rangle_{s_i}, z\in\langle \underline{X}\rangle_{\deg g_i} \\ u^*zv \overset{\mathrm{cyc}}{\sim} w}} g_z^i B_{u,v}^i. \tag{5.2}$$

Proof. We start with the "only if" part. Suppose $f \in \Theta_{S,2s}^2$, hence there exist nc polynomials $a_i = \sum_{w\in\langle \underline{X}\rangle_s} a_w^i w$ and $b_{i,j} = \sum_{w\in\langle \underline{X}\rangle_{s_i}} b_w^{i,j} w$ such that $f \overset{\mathrm{cyc}}{\sim} \sum_i a_i^* a_i + \sum_{i,j} b_{i,j}^* g_i b_{i,j}$. In particular this means that for every $w \in \langle \underline{X}\rangle_{2s}$ the following must hold:

$$\begin{aligned} f_w &= \sum_i \sum_{\substack{u,v\in\langle \underline{X}\rangle_s \\ u^*v \overset{\mathrm{cyc}}{\sim} w}} a_u^i a_v^i + \sum_{i,j} \sum_{\substack{u,v\in\langle \underline{X}\rangle_{s_i}, z\in\langle \underline{X}\rangle_{\deg g_i} \\ u^*zv \overset{\mathrm{cyc}}{\sim} w}} b_u^{i,j} b_v^{i,j} g_z^i \\ &= \sum_{\substack{u,v\in\langle \underline{X}\rangle_s \\ u^*v \overset{\mathrm{cyc}}{\sim} w}} \sum_i a_u^i a_v^i + \sum_i \sum_{\substack{u,v\in\langle \underline{X}\rangle_{s_i}, z\in\langle \underline{X}\rangle_{\deg g_i} \\ u^*zv \overset{\mathrm{cyc}}{\sim} w}} g_z^i \sum_j b_u^{i,j} b_v^{i,j}. \end{aligned}$$

If we define a matrix A of order $\sigma(s)$ and matrices B^i of order $\sigma(s_i)$ by $A_{u,v} = \sum_i a^i_u a^i_v$ and $B^i_{u,v} = \sum_j b^{i,j}_u b^{i,j}_v$, then these matrices are positive semidefinite and satisfy (5.2).

To prove the "if" part we use that A and B^i are positive semidefinite, therefore we can find (column) vectors A_i and $B_{i,j}$ such that $A = \sum_i A_i A_i^T$ and $B^i = \sum_j B_{i,j} B_{i,j}^T$. These vectors yield nc polynomials $a_i = A_i^T \mathbf{W}_{\sigma(s)}$ and $b_{i,j} = B_{i,j}^T \mathbf{W}_{\sigma(s_i)}$, which give a certificate for $f \in \Theta^2_{S,2s}$. ■

Remark 5.8. The last part of the proof of Proposition 5.7 explains how to construct the certificate for $f \in \Theta^2_{S,2s}$. First we solve the semidefinite feasibility problem in the variables $A \in \mathbb{S}^+_{\sigma(s)}$, $B^i \in \mathbb{S}^+_{\sigma(s_i)}$ subject to constraints (5.2). Then we use the Cholesky or eigenvalue decomposition to compute column vectors $A_i \in \mathbb{R}^{\sigma(s)}$ and $B_{i,j} \in \mathbb{R}^{\sigma(s_i)}$ which yield desired polynomial certificates $a_i \in \mathbb{R}\langle \underline{X}\rangle_s$ and $b_{i,j} \in \mathbb{R}\langle \underline{X}\rangle_{s_i}$.

By Proposition 5.7, (Constr-Tr$^{(s)}_{\mathrm{SDP}}$) is an SDP. It can be explicitly presented as

$$\begin{aligned}
\operatorname{tr}^{(s)}_{\Theta^2}(f,S) = \sup\ & f_1 - A_{1,1} - \textstyle\sum_i g^i_1 B^i_{1,1} \\
\text{s. t. } & f_w = \sum_{\substack{u,v \in \langle \underline{X}\rangle_s \\ u^*v \overset{\text{cyc}}{\sim} w}} A_{u,v} + \sum_i \sum_{\substack{u,v \in \langle \underline{X}\rangle_{s_i}, z \in \langle \underline{X}\rangle_{\deg g_i} \\ u^*zv \overset{\text{cyc}}{\sim} w}} g^i_z B^i_{u,v} \\
& \qquad \text{for all } 1 \neq w \in \langle \underline{X}\rangle_{2s}, \\
& A \in \mathbb{S}^+_{\sigma(s)}, \quad B^i \in \mathbb{S}^+_{\sigma(s_i)},
\end{aligned} \tag{Constr-Tr$^{(s)}_{\mathrm{SDP}'}$}$$

where we use $s_i = \lfloor s - \deg(g_i)/2 \rfloor$.

Lemma 5.9. *The dual semidefinite program to* (Constr-Tr$^{(s)}_{\mathrm{SDP}}$) *and* (Constr-Tr$^{(s)}_{\mathrm{SDP}'}$) *is*

$$\begin{aligned}
L^{(s)}_{\Theta^2}(f,S) = \inf\ & L(f) \\
\text{s.t. } & L : \mathbb{R}\langle \underline{X}\rangle_{2s} \to \mathbb{R} \ \text{ is linear and symmetric}, \\
& L(1) = 1, \\
& L(pq - qp) = 0, \ \text{ for all } p,q \in \mathbb{R}\langle \underline{X}\rangle_s, \\
& L(q^*q) \geq 0, \ \text{ for all } q \in \mathbb{R}\langle \underline{X}\rangle_s, \\
& L(h^* g_i h) \geq 0, \ \text{ for all } i \text{ and all } h \in \mathbb{R}\langle \underline{X}\rangle_{s_i}, \\
& \qquad \text{where } s_i = \lfloor s - \deg(g_i)/2 \rfloor.
\end{aligned} \tag{Constr-Tr$^{(s)}_{\mathrm{DSDP}}$}$$

Proof. For this proof it is beneficial to adopt a functional analytic viewpoint of (Constr-Tr$^{(s)}_{\mathrm{SDP}}$) and (Constr-Tr$^{(s)}_{\mathrm{SDP}'}$).

We have the following chain of reasoning, similar to (4.1)–(4.4) (recall $2s \geq \lceil \operatorname{cdeg} f \rceil$):

$$\begin{aligned}
\sup\{\lambda \mid f - \lambda \in \Theta^2_{S,2s}\} &= \sup\ \{\lambda \mid f - \lambda \in \overline{\Theta^2_{S,2s}}\} \\
&= \sup\ \{\lambda \mid \forall L \in \left(\Theta^2_{S,2s}\right)^{\vee} : L(f - \lambda) \geq 0\}
\end{aligned} \tag{5.3}$$

$$= \sup \left\{\lambda \mid \forall L \in \left(\Theta^2_{S,2s}\right)^\vee \text{ with } L(1) = 1 : L(f) \geq \lambda\right\} \tag{5.4}$$

$$= \inf \left\{L(f) \mid L \in \left(\Theta^2_{S,2s}\right)^\vee \text{ with } L(1) = 1\right\}. \tag{5.5}$$

(Recall that $\left(\Theta^2_{S,2s}\right)^\vee$ is the set of all linear functionals $\mathbb{R}\langle \underline{X}\rangle_{2s} \to \mathbb{R}$ nonnegative on $\Theta^2_{S,2s}$.) The last equality is trivial. We next give the reasoning behind the third equality. Clearly, "$\leq$" holds since every λ feasible for the right-hand side of (5.3) is also feasible for the right-hand side of (5.4). To see the reverse inequality we consider an arbitrary λ feasible for (5.4). Note that $\lambda \leq f_1 = \tilde{L}(f)$, where $\tilde{L} \in \left(\Theta^2_{S,2s}\right)^\vee$ maps every polynomial into its constant term. We shall prove that $L(f-\lambda) \geq 0$ for every $L \in \left(\Theta^2_{S,2s}\right)^\vee$. Consider an arbitrary $L \in \left(\Theta^2_{S,2s}\right)^\vee$ and define $\hat{L} = \frac{L+\varepsilon}{L(1)+\varepsilon}$ for some $\varepsilon > 0$. Then $\hat{L}(1) = 1$ and $\hat{L} \in \left(\Theta^2_{S,2s}\right)^\vee$, therefore $\hat{L}(f-\lambda) \geq 0$, whence $L(f-\lambda) \geq \varepsilon(\lambda - 1)$. Since ε was arbitrary we get $L(f-\lambda) \geq 0$.

The problem $\inf\{L(f) \mid L \in \left(\Theta^2_{S,2s}\right)^\vee \text{ with } L(1) = 1\}$ is an SDP, and this is easily seen to be equivalent to the problem (Constr-$\mathrm{Tr}^{(s)}_{\mathrm{DSDP}}$) given above. Indeed, if $L \in \left(\Theta^2_{S,2s}\right)^\vee$, $L(1) = 1$, then L must be nonnegative on the terms (1.18) and on every commutator, therefore L is feasible for the constraints in (Constr-$\mathrm{Tr}^{(s)}_{\mathrm{DSDP}}$). ■

Proposition 5.10. *Suppose $\mathscr{D}_S$ contains an ε-neighborhood of 0. Then the SDP (Constr-$\mathrm{Tr}^{(s)}_{\mathrm{DSDP}}$) admits Slater points.*

Proof. Since the constructed linear functional in the proof of Proposition 4.9 is tracial, the same proof can be applied here and is thus omitted. ■

Remark 5.11. As in the eigenvalue case, having Slater points for (Constr-$\mathrm{Tr}^{(s)}_{\mathrm{DSDP}}$) is important for the duality theory. In particular, there is no duality gap, so for every $s \geq 1$

$$L^{(s)}_{\Theta^2}(f,S) = \mathrm{tr}^{(s)}_{\Theta^2}(f,S)$$

and

$$L_{\Theta^2}(f,S) := \lim_{s\to\infty} L^{(s)}_{\Theta^2}(f,S) = \mathrm{tr}^{\mathrm{II}_1}_{\min}(f,S).$$

Algorithms to compute the lower bounds $\mathrm{tr}^{(s)}_{\Theta^2}(f,S) = L^{(s)}_{\Theta^2}(f,S)$ for $\mathrm{tr}^{\mathrm{II}_1}_{\min}(f,S)$ and $\mathrm{tr}_{\min}(f,S)$ are implemented in `NCSOStools` [CKP11] . We demonstrate it on a few examples at the end of the chapter.

5.4 Flatness and Extracting Optimizers

In this section we assume $S \subseteq \operatorname{Sym}\mathbb{R}\langle\underline{X}\rangle$ is finite, and $f \in \operatorname{Sym}\mathbb{R}\langle\underline{X}\rangle_{2d}$. Let M_S be archimedean. In this case $\mathscr{D}_S^{\mathrm{II}_1}$ is bounded and hence $\operatorname{tr}_{\min}^{\mathrm{II}_1}(f,S) > -\infty$. Since M_S is archimedean, for s big enough (Constr-$\mathrm{Tr}_{\mathrm{SDP}}^{(s)}$) will be feasible.

Like in constrained eigenvalue optimization, flatness is a sufficient condition for finite convergence of the bounds $\operatorname{tr}_{\Theta^2}^{(s)}(f,S) = L_{\Theta^2}^{(s)}(f,S)$ and exactness of the relaxed solution; it also enables the extraction of optimizers.

We first recall a variant of Theorem 1.71 adapted to this setting.

Theorem 5.12. *Suppose L^{opt} is an optimal solution of* (Constr-$\mathrm{Tr}_{\mathrm{DSDP}}^{(s)}$) *for some $s \geq d+\delta$ that is δ-flat. Then there are finitely many n-tuples $\underline{A}^{(j)}$ of symmetric matrices in $\mathscr{D}_S$ and positive scalars $\lambda_j > 0$ with $\sum_j \lambda_j = 1$ such that*

$$L^{\mathrm{opt}}(f) = \sum_j \lambda_j \operatorname{tr} f(\underline{A}^{(j)}). \tag{5.6}$$

In particular, $\operatorname{tr}_{\min}(f,S) = \operatorname{tr}_{\min}^{\mathrm{II}_1}(f,S) = L_{\Theta^2}^{(s)}(f,S) = \operatorname{tr}_{\Theta^2}^{(s)}(f,S)$.

We propose Algorithm 5.1 to find solutions of (Constr-$\mathrm{Tr}_{\mathrm{DSDP}}^{(s)}$) for $s \geq d+\delta$ which are δ-flat enabling us to extract a minimizer of (Constr-$\mathrm{Tr}_{\mathrm{SDP}}^{(s)}$). It is a variant of Algorithm 4.2 and performs surprisingly well; e.g., it finds flat solutions in all tested situations where finite convergence was numerically detected (i.e., at least two consequent bounds were equal).

Algorithm 5.1: Randomized algorithm to find flat solutions for problem (Constr-$\mathrm{Tr}_{\mathrm{DSDP}}^{(s)}$)

```
Input: $f \in \operatorname{Sym}\mathbb{R}\langle\underline{X}\rangle$ with $\deg f = 2d$, $S = \{g_1,\dots,g_t\}$,
       $\delta = \lceil \max_i \deg(g_i)/2 \rceil$, $\delta_{\max}$;
$L_{\mathrm{flat}} = 0$;
for $s = d+\delta, d+\delta+1, \dots, d+\delta+d_{\max}$ do
    Compute $L^{(s)}$ – the optimal solution for (Constr-$\mathrm{Tr}_{\mathrm{DSDP}}^{(s)}$);
    if $L^{(s)}$ is $\delta$-flat then
        $L_{\mathrm{flat}} = L^{(s)}$. Stop;
    end
    Compute $L_{\mathrm{rand}}^{(s)}$;
    if $L_{\mathrm{rand}}^{(s)}$ is $\delta$-flat then
        $L_{\mathrm{flat}} = L_{\mathrm{rand}}^{(s)}$. Stop;
    end
end
Output: $L_{\mathrm{flat}}$;
```

In Step 7 we solve the SDP which is obtained from (Constr-Tr$^{(s)}_{\mathrm{DSDP}}$) by fixing the upper left-hand corner of the Hankel matrix to be equal to the upper left-hand corner of the Hankel matrix of $L^{(s)}$ and by taking a full random objective function—like in (Constr-Eig$^{(s)}_{\mathrm{RAND}}$). We repeat this step several (e.g. 10) times. In our experiments, this algorithm very often returns flat solutions if the module $\Theta^2_{S,2d}$ is archimedean. On the other hand, there is little theoretical evidence supporting this performance.

We repeat Steps 1–3 at most $\delta_{\max}+1$ times, where $\delta_{\max}$ is for computational complexity reasons chosen so that $d+\delta+\delta_{\max}$ is at most 10, when we have two nc variables, and is at most 8 if we have three nc variables. Otherwise the complexity of the underlying SDP exceeds the capability of our current hardware. We implemented Steps 1–3 from 5.1 in the `NCSOStools` function `NCtraceOptRand`.

In [KP16] we report numerical results obtained by running Algorithm 5.1 on random polynomials. We generated random polynomials as in Sect. 4.3.2 and we check for δ-flatness by computing ranks much like in Sect. 4.3.2. In all cases we took the tolerance to be $\min\{30\cdot \mathtt{err}_{\mathtt{flat}}, 10^{-3}\}$.

With this tolerance we can observe (as in Sect. 4.3.2) that in almost all tested (random) cases Algorithm 5.1 returned a flat optimal solution already after the first step, i.e., for $s=d+\delta$; see [KP16, Table 4] for concrete results.

Once we have a flat optimum solution for (Tr$_{\mathrm{DSDP}}$) or (Constr-Tr$^{(s)}_{\mathrm{DSDP}}$) we can extract optimizers, i.e., compute an n-tuple of symmetric matrices $\underline{A}$, which is in $\mathscr{D}_S$ when we consider the constrained case, such that $\operatorname{tr}(\underline{A})$ is equal to $\operatorname{tr}_{\min}(f)$ and $\operatorname{tr}_{\min}(f,S)$, respectively, by running Algorithm 1.2.

5.5 Implementation

We can compute the unconstrained and constrained trace optimum exactly only for very simple and nice examples. For all other cases we shall use numerical algorithms. The software package `NCSOStools` contains `NCcycMin` to compute the unconstrained trace optimum (i.e., $\operatorname{tr}_{\Theta^2}(f) = L_{\Theta^2}(f)$) and `NCcycOpt` to extract the related optimizers if the dual optimal solution is 1-flat. Likewise we have `NCtraceOpt` to compute $\operatorname{tr}^{(s)}_{\Theta^2}(f,S)$ and `NCtraceOptRand` to compute flat solutions together with $\operatorname{tr}^{\mathrm{II}_1}_{\min}(f,S)$ when a flat solution is found. In this case we also extract optimizers by running Algorithm 1.2.

Example 5.13. Let

$$\begin{aligned} f = {} & 3 + X_1^2 + 2X_1^3 + 2X_1^4 + X_1^6 - 4X_1^4X_2 + X_1^4X_2^2 + 4X_1^3X_2 + 2X_1^3X_2^2 - 2X_1^3X_2^3 \\ & + 2X_1^2X_2 - X_1^2X_2^2 + 8X_1X_2X_1X_2 + 2X_1^2X_2^3 - 4X_1X_2 + 4X_1X_2^2 + 6X_1X_2^4 - 2X_2 \\ & + X_2^2 - 4X_2^3 + 2X_2^4 + 2X_2^6. \end{aligned}$$

The minimum of f on $\mathbb{R}^2$ is 1.0797. Using NCcycMin we obtain the floating point trace-minimum $\operatorname{tr}_{\Theta^2}(f) = 0.2842$ for f which is different from the commutative minimum. In particular, the minimizers will not be scalar matrices. The dual optimal solution for ($\mathrm{Tr}_{\mathrm{DSDP}}$) is of rank 4 and 1-flat. Thus the matrix representation of the multiplication operators A_i is given by 4×4 matrices (see the proof of Theorem 1.69 and Algorithm 1.1):

$$A_1 = \begin{bmatrix} -1.0761 & 0.5319 & 0.1015 & 0.2590 \\ 0.5319 & 0.4333 & -0.3092 & 0.2008 \\ 0.1015 & -0.3092 & -0.2633 & 0.9231 \\ 0.2590 & 0.2008 & 0.9231 & -0.3020 \end{bmatrix},$$

$$A_2 = \begin{bmatrix} 0.7107 & 0.2130 & 0.7090 & 0.4415 \\ 0.2130 & 0.2087 & 0.3878 & -0.9321 \\ 0.7090 & 0.3878 & -0.5016 & -0.0757 \\ 0.4415 & -0.9321 & -0.0757 & 0.1393 \end{bmatrix}.$$

The Artin–Wedderburn decomposition for the matrix $*$-algebra $\mathscr{A}$ generated by A_1, A_2 gives in this case only one block. Using NCcycOpt, which essentially implements Algorithm 1.2 leads to the trace-minimizer

$$\hat{A}_1 = \begin{bmatrix} -1.0397 & -0.0000 & 0.1024 & 0.6363 \\ -0.0000 & -1.0397 & -0.6363 & 0.1024 \\ 0.1024 & -0.6363 & 0.4356 & -0.0000 \\ 0.6363 & 0.1024 & -0.0000 & 0.4356 \end{bmatrix},$$

$$\hat{A}_2 = \begin{bmatrix} -0.4246 & 0.0000 & -0.1377 & -0.8559 \\ 0.0000 & -0.4246 & 0.8559 & -0.1377 \\ -0.1377 & 0.8559 & 0.7031 & 0.0000 \\ -0.8559 & -0.1377 & 0.0000 & 0.7031 \end{bmatrix}.$$

The reader can easily verify that $\operatorname{tr}(f(\hat{A}_1, \hat{A}_2)) = 0.2842$.

Note that $\mathscr{A}$ is (as a real $*$-algebra) isomorphic to $M_2(\mathbb{C})$. For instance,

$$A \approx \tilde{A} = \begin{bmatrix} -1.0397 & 0.6363 + 0.1024\mathrm{i} \\ 0.6363 - 0.1024\mathrm{i} & 0.4356 \end{bmatrix},$$

$$B \approx \tilde{B} = \begin{bmatrix} -0.4246 & -0.8559 - 0.1377\mathrm{i} \\ -0.8559 + 0.1377\mathrm{i} & 0.7031 \end{bmatrix}.$$

In this case it is possible to find a unitary matrix $U \in \mathbb{C}^{2\times 2}$ with $A' = U^* \tilde{A} U \in \mathbb{R}^{2\times 2}$ and $B' = U^* \tilde{B} U \in \mathbb{R}^{2\times 2}$, e.g.,

$$U = \begin{bmatrix} 0.9803+0.1576\mathrm{i} & 0.1176+0.0189\mathrm{i} \\ 0.1191 & -0.9929 \end{bmatrix},$$

$$A' = \begin{bmatrix} -0.8663 & -0.8007 \\ -0.8007 & 0.2622 \end{bmatrix}, \quad B' = \begin{bmatrix} -0.6136 & 0.7089 \\ 0.7089 & 0.8921 \end{bmatrix}.$$

Then $(A', B') \in \left(\mathbb{S}^{2\times 2}\right)^2$ is also a trace-minimizer for f.

Example 5.14. We demonstrate our software for constrained trace optimization for the set $S = \{1 - X^2, 1 - Y^2\}$ with the polynomial

$$p = (1-X^2)(1-Y^2) + (1-Y^2)(1-X^2)$$

from Example 5.6, and a non-commutative version of the Motzkin polynomial from Example 4.25,

$$q = XY^4X + YX^4Y - 3XY^2X + 1.$$

It is obvious (see Example 5.6 and [KS08, Example 4.3]) that $\operatorname{tr}^{\mathrm{II}_1}_{\min}(p,S) = 0$. Similarly, $\operatorname{tr}^{\mathrm{II}_1}_{\min}(q,S) = 0$ (see [KS08, Example 4.4]). We use NCSOStools as follows:

```
>> NCvars x y
>> S = {1 - x^2, 1 - y^2};
>> p = (1-x^2)*(1-y^2)+(1-y^2)*(1-x^2);
>> q = x*y^4*x+y*x^4*y-3*x*y^2*x+1;
```

To compute the sequence of lower bounds $\operatorname{tr}^{(s)}_{\Theta^2}(p,S)$ for $\operatorname{tr}^{\mathrm{II}_1}_{\min}(p,S)$ we call

```
>> [opt,decom_sohs,decom_S,base] = NCtraceOpt(p,S,2*s);
```

with $s = 2,3,4,5$. Similarly we obtain bounds for q. Results are reported in Table 5.1.

We can see that the sequence of bounds $\operatorname{tr}^{(s)}_{\Theta^2}(p,S)$ of p increases and does not reach the limit for $s \le 5$. Actually, it never reaches $\operatorname{tr}^{\mathrm{II}_1}_{\min}(p,S)$; see Example 5.6. On the other hand, the sequence of bounds for q is finite and reaches the optimal value already for $s = 3$ ($\operatorname{tr}^{(2)}_{\Theta^2}(q,S)$ is not defined).

Table 5.1 Lower bounds $\operatorname{tr}^{(s)}_{\Theta^2}(f,S)$ for p and q over $S = \{1-X^2, 1-Y^2\}$

s	$\operatorname{tr}^{(s)}_{\Theta^2}(p,S)$	$\operatorname{tr}^{(s)}_{\Theta^2}(q,S)$
2	−0.2500	n.d.
3	−0.0178	0
4	−0.0031	0
5	−0.0010	0

Table 5.2 Lower bounds $\mathrm{tr}_{\Theta^2}^{(s)}(p,S)$, $\mathrm{tr}_{\Theta^2}^{(s)}(q,S)$, and $\mathrm{tr}_{\Theta^2}^{(s)}(r,S)$ over $S = \{1-X, 1-Y, 1+X, 1+Y\}$

s	$\mathrm{tr}_{\Theta^2}^{(s)}(p,S)$	$\mathrm{tr}_{\Theta^2}^{(s)}(q,S)$	$\mathrm{tr}_{\Theta^2}^{(s)}(r,S)$
2	−2.0000	n.d.	−1.0000
3	−0.2500	−0.0261	−1.0000
4	−0.0178	0.0000	−1.0000
5	−0.0031	0.0000	−1.0000

Example 5.15. Let p,q be as in Example 5.14 and let $r = XYX$. Let us define $S = \{1-X, 1-Y, 1+X, 1+Y\}$. The resulting sequences from the relaxation are in Table 5.2 and show that there is again no convergence in the first four steps for p, while for q we get convergence at $s=4$ and for r we get the optimal value immediately (at $s=2$).

To compute, e.g., $\mathrm{tr}_{\Theta^2}^{(5)}(p,S)$ we need to solve (Constr-$\mathrm{Tr}_{\mathrm{SDP}'}^{(s)}$) which has 3739 linear constraints and five positive semidefinite constraints with matrix variables of order 63, 31, 31, 31, 31.

Example 5.16. Let us consider $p = XY$, $q = 1 + X(Y-2) + Y(X-2)$, $f = p^*q + q^*p$ and $S = \{4-X^2, 4-Y^2\}$. If we use NCSOStools and call

```
>> NCvars x y
>> p = x*y;q = 1+x*(y-2)+y*(x-2);f = p'*q+q'*p;
>> S = {4-x^2,4-y^2};
>> [opt_2,decom_1,dec_S1,base1] = NCtraceOpt(f,S,4);
>> [opt_3,decom_2,dec_S2,base2] = NCtraceOpt(f,S,6);
>> [opt_4,decom_3,dec_S3,base3] = NCtraceOpt(f,S,8);
```

we obtain `opt_2` $= \mathrm{tr}_{\Theta^2}^{(2)}(f,S) = -8$ and `opt_3` $= \mathrm{tr}_{\Theta^2}^{(3)}(f,S) = \mathrm{tr}_{\Theta^2}^{(4)}(f,S) = -5.2165$. This was checked numerically but running `NCtraceOptRand` did not finish with a numerical proof of 1-flat solutions, so we cannot claim that $\mathrm{tr}_{\min}^{\mathrm{II}_1}(f,S)$ is equal to -5.2165.

It is easy to see that the (commutative) minimum of f on $\mathscr{D}_S \cap \mathbb{R}^2 = [-2,2]^2$ is -4.5.

Example 5.17. Let us compute the trace-minimum of $f = 2 - X^2 + XY^2X - Y^2$ over the semialgebraic set defined by $S = \{4 - X^2 - Y^2, XY + YX - 2\}$.

```
>> NCvars x y
>> f = 2 - x^2 + x*y^2*x - y^2;
>> S={4-x^2-y^2,x*y+y*x-2};
>> [X,fX,tr_val,flat,err_flat]=NCtraceOptRand(f,S,4);
```

Firstly we see that `flat` $=1$ which means that the method has found a flat optimal solution with `err_flat`$\approx 4\cdot 10^{-8}$. This gives a matrix X of size 2×16; each row represents one symmetric 4×4 matrix,

$$A = \mathrm{reshape}(X(1,:),4,4) = \begin{bmatrix} -0.0000 & 1.4044 & -0.1666 & -0.0000 \\ 1.4044 & 0.0000 & 0.0000 & 1.1329 \\ -0.1666 & 0.0000 & -0.0000 & -0.8465 \\ -0.0000 & 1.1329 & -0.8465 & 0.0000 \end{bmatrix}$$

$$B = \mathrm{reshape}(X(2,:),5,5) = \begin{bmatrix} -0.0000 & 0.8465 & 1.1329 & 0.0000 \\ 0.8465 & 0.0000 & 0.0000 & -0.1666 \\ 1.1329 & 0.0000 & 0.0000 & -1.4044 \\ 0.0000 & -0.1666 & -1.4044 & 0.0000 \end{bmatrix}$$

such that A and B are from $\mathscr{D}_S(4)$ and

$$fX = f(A,B) = \begin{bmatrix} -1.0000 & 0.0000 & 0.0000 & -0.0000 \\ 0.0000 & -1.0000 & 0.0000 & 0.0000 \\ 0.0000 & 0.0000 & -1.0000 & 0.0000 \\ -0.0000 & 0.0000 & 0.0000 & -1.0000 \end{bmatrix}$$

with (normalized) trace equal to `trace_val` $= -1$.

References

[CKP11] Cafuta, K., Klep, I., Povh, J.: NCSOStools: a computer algebra system for symbolic and numerical computation with noncommutative polynomials. Optim. Methods. Softw. **26**(3), 363–380 (2011). Available from http://ncsostools.fis.unm.si/

[Con76] Connes, A.: Classification of injective factors. Cases II_1, II_∞, III_λ, $\lambda \neq 1$. Ann. Math. (2) **104**(1), 73–115 (1976)

[Ji13] Ji, Z.: Binary Constraint System Games and Locally Commutative Reductions. arXiv preprint. arXiv:1310.3794 (2013)

[KP16] Klep, I., Povh, J.: Constrained trace-optimization of polynomials in freely noncommuting variables. J. Glob Optim. **64**, 325–348 (2016)

[KS08] Klep, I., Schweighofer, M.: Connes' embedding conjecture and sums of hermitian squares. Adv. Math. **217**(4), 1816–1837 (2008)

[Las01] Lasserre, J.B.: Global optimization with polynomials and the problem of moments. SIAM J. Optim. **11**(3), 796–817 (2000/01)

[Las09] Lasserre, J.B.: Moments, Positive Polynomials and Their Application. Imperial College Press, London (2009)

[LP15] Laurent, M., Piovesan, T.: Conic approach to quantum graph parameters using linear optimization over the completely positive semidefinite cone. SIAM J. Optim. **25**(4), 2461–2493 (2015)

List of Symbols

Symbol	**Description**
$\operatorname{conv}(A)$	Convex hull of the set A
$\operatorname{diag}(A)$	Vector with diagonal entries of the matrix A
$\operatorname{Diag}(A_1,\dots,A_n)$	Block diagonal matrix with matrices A_i on the main diagonal
$\mathbf{e}_i$	ith standard unit vector
I_n	The $n \times n$ identity matrix
$\lambda_i(X)$	ith smallest eigenvalue of a symmetric matrix X
$\lambda_{\min}(X)$ $(\lambda_{\max}(X))$	The smallest (largest) eigenvalue of a symmetric matrix X
$\mathscr{M}_n$	The vector space of $n \times n$ real matrices
$\mathscr{M}_{m,n}$	The vector space of $m \times n$ real matrices
$\mathbb{S}_n$	The vector space of $n \times n$ symmetric matrices
$\mathbb{S}_n^+$	The cone of positive semidefinite $n \times n$ matrices
$\mathbb{S}_n^{++}$	The cone of positive definite $n \times n$ matrices
$\operatorname{tr}(A)$	Normalized trace of a square matrix A
$X \succeq Y$	$X - Y \in \mathbb{S}_n^+$ $(X - Y \succeq 0)$
$\langle x \mid y \rangle$	$x^T y$
$\langle X \mid Y \rangle$	$\operatorname{tr}(X^T Y)$
$\langle \underline{X} \rangle$	The monoid, freely generated by the letters $X_1,\dots,X_n$
$\mathbb{R}\langle \underline{X} \rangle$	The free algebra with generating set $\{X_1,\dots,X_n\}$
$\operatorname{Sym}\mathbb{R}\langle \underline{X} \rangle$	The set of all symmetric elements form $\mathbb{R}\langle \underline{X} \rangle$
Σ^2	Sums of hermitian squares
Σ^2_{2d}	Sums of hermitian squares of degree $\leq 2d$
$(\Sigma^2_{2d})^\vee$	The dual cone to Σ^2_{2d}

© The Author(s) 2016

S. Burgdorf et al., *Optimization of Polynomials in Non-Commuting Variables*,
SpringerBriefs in Mathematics, DOI 10.1007/978-3-319-33338-0

$\mathbf{W}_d$	Vector with all words of degree $\leq d$
M_S	Quadratic module generated by S
$M_{S,2d}$	Truncated quadratic module generated by S with elements of degree $\leq 2d$
$\mathscr{D}_S$	Semialgebraic set associated to $S \subseteq \operatorname{Sym}\mathbb{R}\langle\underline{X}\rangle$
$\mathscr{D}_S(k)$	$\mathscr{D}_S \cap \mathbb{S}_k^n$
$\mathscr{D}_S^\infty$	Operator semialgebraic set
$\mathscr{D}_S^{\mathscr{F}}$	$\mathscr{F}$-semialgebraic set
$\mathscr{N}_\varepsilon$	An nc ε-neighborhood of 0
$\mathscr{B}(\underline{A},\varepsilon)$	An nc ε-neighborhood of $\underline{A}$
$\operatorname{cc}(f)$	Commutative collapse of f
$\deg_{\underline{\alpha}} w$	α-degree of f
$\overset{\text{cyc}}{\sim}$	Cyclic equivalence on $\mathbb{R}\langle\underline{X}\rangle$
$[f]$	Canonical representative of $f \in \mathbb{R}\langle\underline{X}\rangle$
$\lambda_{\min}(f)$	Smallest eigenvalue $f \in \operatorname{Sym}\mathbb{R}\langle\underline{X}\rangle$ attains
$\lambda_{\min}(f,S)$	Smallest eigenvalue of $f \in \operatorname{Sym}\mathbb{R}\langle\underline{X}\rangle$ on $\mathscr{D}_S^\infty$
$f_{\text{sohs}}^{(s)}, L_{\text{sohs}}^{(s)}$	sth approximation for $\lambda_{\min}(f,S)$
$\mathbb{B}$, $\mathbb{D}$	Nc ball, nc polydisc (respectively)
$\operatorname{tr}_{\min}(f)$	Smallest trace of f on $\mathbb{S}^n$
$\operatorname{tr}_{\Theta^2}(f)$	SDP approximation for $\operatorname{tr}_{\min}(f)$
$\operatorname{tr}_{\min}(f,S)$	Smallest trace of f on $\mathscr{D}_S$
$\operatorname{tr}_{\min}^{\mathrm{II}_1}(f,S)$	Smallest trace of f on $\mathscr{D}_S^{\mathrm{II}_1}$
$\operatorname{tr}_{\Theta^2}^{(s)}(f,S)$, $L_{\Theta^2}^{(s)}(f,S)$	SDP approximations for $\operatorname{tr}_{\min}^{\mathrm{II}_1}(f,S)$ of order s

Index

© The Author(s) 2016

S. Burgdorf et al., *Optimization of Polynomials in Non-Commuting Variables*,
SpringerBriefs in Mathematics, DOI 10.1007/978-3-319-33338-0