Towards a theory of non-commutative optimization: geodesic first and second order methods for moment maps and polytopes

Peter Bürgisser* Cole Franks† Ankit Garg‡ Rafael Oliveira§
Michael Walter¶ Avi Wigderson‖

Abstract

This paper initiates a systematic development of a theory of *non-commutative* optimization, a setting which greatly extends ordinary (Euclidean) convex optimization. It aims to unify and generalize a growing body of work from the past few years which developed and analyzed algorithms for natural *geodesically convex* optimization problems on Riemannian manifolds that arise from the symmetries of non-commutative groups. More specifically, these are algorithms to minimize the *moment map* (a non-commutative notion of the usual *gradient*), and to test membership in *moment polytopes* (a vast class of polytopes, typically of exponential vertex and facet complexity, which quite magically arise from this a-priori non-convex, non-linear setting).

The importance of understanding this very general setting of geodesic optimization, as these works unveiled and powerfully demonstrate, is that it captures a diverse set of problems, many non-convex, in different areas of CS, math, and physics. Several of them were solved efficiently for the first time using non-commutative methods; the corresponding algorithms also lead to solutions of purely structural problems and to many new connections between disparate fields.

In the spirit of standard convex optimization, we develop two general methods in the geodesic setting, a first order and a second order method, which respectively receive first and second order information on the "derivatives" of the function to be optimized. These in particular subsume all past results. The main technical work, again unifying and extending much of the previous work, goes into identifying the key parameters of the underlying group actions which control convergence to the optimum in each of these methods. These non-commutative analogues of "smoothness" in the commutative case are far more complex, and require significant algebraic and analytic machinery (much existing and some newly developed here). Despite this complexity, the way in which these parameters control convergence in both methods is quite simple and elegant. We also bound these parameters in several general cases.

Our work points to intriguing open problems and suggests further research directions. We believe that extending this theory, namely understanding geodesic optimization better, is both mathematically and computationally fascinating; it provides a great meeting place for ideas and techniques from several very different research areas, and promises better algorithms for existing and yet unforeseen applications.

[*] Institut für Mathematik, Technische Universität Berlin, pbuerg@math.tu-berlin.de. Supported by DFG grant BU 1371 2-2 and by the ERC under the European's Horizon 2020 research and innovation programme (grant agreement no. 787840).

[†] Department of Mathematics, Rutgers University, wcf17@math.rutgers.edu. Supported in part by Simons Foundation award 332622.

[‡] Microsoft Research India, garga@microsoft.com.

[§] Department of Computer Science, University of Toronto, rafael@cs.toronto.edu. Part of this work was done while the author was visiting the Simons Institute for the Theory of Computing.

[¶] Korteweg-de Vries Institute for Mathematics, Institute for Theoretical Physics, Institute for Logic, Language, and Computation, QuSoft, University of Amsterdam, m.walter@uva.nl. Supported by an NWO Veni grant no. 680-47-459.

[‖] Institute for Advanced Study, Princeton, avi@ias.edu. Research partially supported by NSF grant CCF-1412958.

Contents

1 Introduction

1.1 High-level overview

Consider a group G that acts by *linear* transformations on the complex Euclidean space $V = \mathbb{C}^m$. This partitions V into *orbits*: For a vector $v \in V$, the orbit $\mathcal{O}_v$ is simply all vectors of the form $g \cdot v$ to which the action of a group element $g \in G$ can map v.

The most basic algorithmic question in this setting is as follows. Given a vector $v \in V$, compute (or approximate) the smallest ℓ_2-norm of any vector in the orbit of v, that is, $\inf\{\|w\|_2 : w \in \mathcal{O}_v\}$. Remarkably, this simple question, for different groups G, captures natural important problems in computational complexity, algebra, analysis, and quantum information. Even when restricted only to *commutative* groups, it already captures all linear programming problems!

Starting with [GGOW16], a series of recent works including [GGOW17, BGO+17, Fra18, KLLR18, AZGL+18, BFG+18] designed algorithms and analysis tools to handle this basic and other related optimization problems over *non-commutative* groups G. These provided efficient solutions for some applications, and *through algorithms*, the resolution of some purely structural mathematical open problems. We will mention some of these below.

A great deal of understanding gradually evolved in this sequence of works. These new algorithms are all essentially iterative methods, progressing from the input vector v to the desired optimum in small steps, as do convex optimization algorithms. This seems surprising, as the basic question above is patently *non-convex* for non-commutative groups (in the commutative case, a simple change of variables discussed below convexifies the problem). Indeed, neither the domain nor the function to be optimized are convex! However, in hindsight, a key to all of them are the notions of *geodesic convexity* (which generalizes the familiar Euclidean notion of convexity) and the *moment map* (which generalizes the familiar Euclidean gradient) in the curved space and new metrics induced by the group action. A rich duality theory of geometric invariant theory (greatly generalizing LP duality), together with tools from algebraic geometry, representation theory and differential equations are used in the convergence analysis of these algorithms.

The main objective of this paper is to unify and generalize these works, in a way which naturally extends the familiar first and second order methods of standard convex optimization. We design geodesic analogs of these methods, which, respectively, have oracle access to first and second order "derivatives" of the function being optimized. Our first order method (which is a non-commutative version of gradient descent) replaces and extends the use of "alternate minimization" in most past works, and thus can accommodate more general group actions. Our second order method greatly generalizes the one used for the particular group action corresponding to operator scaling in [AZGL+18]. It may be thought of as a geodesic analog of the "trust region method" [CGT00] or the "box-constrained Newton method" [CMTV17, AZLOW17] applied to a regularized function. For both methods, in this non-commutative setting, we recover the familiar convergence behavior of the classical commutative case: to achieve "proximity" ε to the optimum, our first order method converges in $O(1/\varepsilon)$ iterations and our second order method in $O(\operatorname{poly}\log(1/\varepsilon))$ iterations.

As in the commutative case, the fundamental challenge is to understand the "constants" hidden in the big-O notation of each method. These depend on "smoothness" properties of the function optimized, which in turn are determined by the action of the group G on the space V that defines the optimization problem. The main technical contributions of the theory we develop are to identify the key parameters which control this dependence, and to bound them for various actions to obtain concrete running time bounds. These parameters depend on a combination of algebraic and geometric properties of the group action, in particular the irreducible representations occurring in it. As mentioned, despite the technical complexity of defining (and bounding) these parameters, the way they control convergence of the algorithms is surprisingly elegant.

We also develop important technical tools which naturally extend ones common in the commutative theory, including regularizers, diameter bounds, numerical stability, and initial starting points, which are key to the design and analysis of the presented (and hopefully future) algorithms in the geodesic setting.

As in previous works, we also address other optimization problems beyond the basic "norm minimization" question above, in particular the minimization of the moment map (which turns out to be a dual problem), and the membership problem for *moment polytopes*; a very rich class of polytopes (typically with exponentially many vertices and facets) which arises magically from any such group action.

The paradigm of optimization described above resulted in efficient algorithms for problems from various diverse areas of CS and mathematics. We mention some of these applications in the following Section 1.2. Section 1.3 describes the basic setting of non-commutative optimization. Next, in Section 1.4, we formally define the problems we are studying and survey what is known about them in the commutative and non-commutative case. Finally, in Sections 1.5 and 1.6, we describe our contributions and results.

1.2 Some unexpected applications and connections

We mention here some of the diverse applications of the paradigm of optimization over non-commutative groups:

1. **Algebraic identities:** Given an arithmetic formula (with inversion gates) in non-commuting variables, is it identically zero?

2. **Quantum information:** Given density matrices describing local quantum states of various parties, is there a global pure state consistent with the local states?

3. **Eigenvalues of sums of Hermitian matrices:** Given three real n-vectors, do there exist three Hermitian $n \times n$ matrices A, B, C with these *prescribed* spectra, such that $A + B = C$?

4. **Analytic inequalities:** Given m linear maps $A_i : \mathbb{R}^n \to \mathbb{R}^{n_i}$ and $p_1, \ldots, p_m \geqslant 0$, does there exist a finite constant C such that for all integrable functions $f_i : \mathbb{R}^{n_i} \to \mathbb{R}_+$ we have

$$\int_{x\in\mathbb{R}^n} \prod_{i=1}^m f_i(A_i x) dx \ \leqslant C \prod_{i=1}^m \|f_i\|_{1/p_i}?$$

   These inequalities are the celebrated Brascamp-Lieb inequalities, which capture the Cauchy-Schwarz, Hölder, Loomis-Whitney, and many further inequalities.

At first glance, it is far from obvious that solving any of these problems has any relation to *either* optimization or groups. We will clarify this mystery below, showing not only how symmetries naturally exist in all of them, but also how these help both in formulating them as optimization problems over groups, suggesting natural algorithms (or at least heuristics) for them, and finally in providing tools for analyzing these algorithms. It perhaps should be stressed that symmetries exist in many examples in which the relevant groups are commutative (e.g., perfect matching in bipartite graphs, matrix scaling, and more generally in linear, geometric, and hyperbolic programming);

however in these cases, standard convex optimization or combinatorial algorithms can be designed and analyzed *without* any reference to these existing symmetries. Making this connection explicit is an important part of our exposition.

Polynomial time algorithms were first given in [GGOW16] for Problem 1 (the works [IQS17b, DM17, IQS17a] later discovered completely different algebraic algorithms), in [BFG$^+$18] for Problem 2 (cf. [VDDM03] and the structural results [Kly04, DH05, CM06, CHM07, WDGC13, Wal14, CDKW14]), in [KT99, DLM06, MNS12, BI13, Fra18] for Problem 3 (the celebrated structural result in [KT99] and the algorithmic results of [DLM06, MNS12, BI13] solved the decision problem, while [Fra18] solved the search problem), and in [GGOW17] for Problem 4. However the algorithms in [GGOW17, Fra18, BFG$^+$18] remain exponential time in various input parameters, exemplifying only one aspect of many in which the current theory and understanding is lacking.

The unexpected connections revealed in this study are far richer than the mere relevance of optimization and symmetries to such problems. One type are connections between problems in disparate fields. For example, the analytic Problem 4 turns out to be a special case of the algebraic Problem 1. Moreover, Problem 1 has (well-studied) differently looking but equivalent formulations in quantum information theory and in invariant theory, which are automatically solved by the same algorithm. Another type of connections are of purely structural open problems solved through such geodesic algorithms, reasserting the importance of the computational lens in mathematics. One was the first dimension-independent bound on the Paulsen problem in operator theory, obtained ingeniously through such an algorithm in [KLLR18] (this work was followed by [HM18], who gave a strikingly simpler proof and stronger bounds). Another was a quantitative bound on the continuity of the best constant C in Problem 4 (in terms of the input data), important for non-linear variants of such inequalities. This bound was obtained through the algorithm in [GGOW17] and relies on its efficiency; previous methods used compactness arguments that provided no bounds.

We have no doubt that more unexpected applications and connections will follow. The most extreme and speculative perhaps among such potential applications is to develop a deterministic polynomial-time algorithm for the polynomial identity testing (PIT) problem. Such an algorithm will imply major algebraic or Boolean lower bounds, namely either separating VP from VNP, or proving that NEXP has no small Boolean circuits [KI04]. We note that this goal was a central motivation of the initial work in this sequence [GGOW16], which provided such a deterministic algorithm for Problem 1 above, the non-commutative analog of PIT. The "real" PIT problem (in which variables commute) also has a natural group of symmetries acting on it, which does not quite fall into the frameworks developed so far (including the one of this paper). Yet, the hope of proving lower bounds via optimization methods is a fascinating (and possibly achievable) one. This agenda of hoping to shed light on the PIT problem by the study of invariant theoretic questions was formulated in the fifth paper of the Geometric Complexity Theory (GCT) series [Mul12, Mul17].

1.3 Non-commutative optimization: a primer

We now give an introduction to non-commutative optimization and contrast its geometric structure and convexity properties with the familiar commutative setting. The basic setting is that of a continuous group G acting (linearly) on an m-dimensional complex vector space $V \cong \mathbb{C}^m$. For this section, and the rest of the introduction, think of G as either the group of $n \times n$ complex invertible matrices, denoted $\mathrm{GL}(n)$, or the group of *diagonal* such matrices, denoted $\mathrm{T}(n)$. The latter corresponds to the commutative case and the former is a paradigmatic example of the non-commutative case. An *action* (also called a *representation*) of a group G on a complex vector space V

is a group homomorphism $\pi : G \to GL(V)$, that is, an association of an invertible linear map $\pi(g)$ from $V \to V$ for every group element $g \in G$ satisfying $\pi(g_1 g_2) = \pi(g_1)\pi(g_2)$ for all $g_1, g_2 \in G$. Further suppose that V is also equipped with an inner product $\langle \cdot, \cdot \rangle$ and hence a norm $\|v\| := \langle v, v \rangle$.[1]

Given a vector $v \in V$ one can consider the optimization problem of taking the infimum of the norm in the *orbit* of the vector v under the action of G. More formally, define the *capacity* of v by[2]

$$\operatorname{cap}(v) := \inf_{g \in G} \|\pi(g)v\|.$$

This notion generalizes the matrix and operator capacities developed in [GY98, Gur04a] (to see this, carry out the optimization over one of the two group variables) as well as the polynomial capacity of [Gur06]. It turns out that this simple-looking optimization problem is already very general in the commutative case and, in the non-commutative case, captures *all* examples discussed in Section 1.2.

Let us first consider the commutative case, $G = T(n)$ acting on V. In this simple case, *all* actions π have a very simple form. We give two equivalent descriptions, first of how any representation π splits into one-dimensional irreducible representations, and another describing π as a natural scaling action on n-variate polynomials with m monomials.

The irreducible representations are given by an orthonormal basis $v_1, \ldots, v_m$ of V such that the v_j are simultaneous eigenvectors of all the matrices $\pi(g)$. That is, for all $g = \operatorname{diag}(g_1, \ldots, g_n) \in T(n)$ and $j \in [m]$,

$$\pi(g)v_j = \lambda_j(g)v_j, \quad \text{where} \quad \lambda_j(g) = \textstyle\prod_{i=1}^n g_i^{\omega_{j,i}} \tag{1.1}$$

for fixed integer vectors $\omega_1, \ldots, \omega_m \in \mathbb{Z}^n$, which are called *weights* and encode the simultaneous eigenvalues, and completely determine the action. Below we also refer to the weights of representation π of $GL(n)$, defined as the weights of π restricted to $T(n)$.

A natural way to view all these actions is as follows. The natural action of $T(n)$ on $\mathbb{C}^n$ by matrix-vector multiplication, induces an action of $T(n)$ on n-variate polynomials $V = \mathbb{C}[x_1, x_2, \ldots, x_n]$: simply, any group element $g = \operatorname{diag}(g_1, \ldots, g_n)$ “scales” each x_i to $g_i x_i$. Note that any monomial $x^\omega = \prod_{i=1}^n x_i^{\omega(i)}$ (where ω is the integer vector of exponents) is an eigenvector of this action, with an eigenvalue $\lambda(g) = \prod_{i=1}^n g_i^{\omega(i)}$.

Now fix m integer vectors ω_j as above. Consider the linear space of n-variate Laurent polynomials (i.e., polynomials where the variables can have negative exponents, too) with the following m monomials: $v_j = x^{\omega_j} = \prod_{i=1}^n x_i^{\omega_{j,i}}$. The action on any polynomial $v = \sum_{j=1}^m c_j v_j$ is precisely the one described above, scaling each coefficient by the eigenvalue of its monomial. The norm $\|v\|$ of a polynomial is the sum of the square moduli of its coefficients. Now let us calculate the capacity of this action. For any $v = \sum_{j=1}^m c_j v_j$,

$$\operatorname{cap}(v)^2 = \inf_{g_1,\ldots,g_n \in \mathbb{C}^*} \textstyle\sum_{j=1}^m |c_j|^2 \prod_{i=1}^n |g_i|^{2\omega_{j,i}} = \inf_{x \in \mathbb{R}^n} \sum_{j=1}^m |c_j|^2 e^{x \cdot \omega_j}, \tag{1.2}$$

where we used the change of variables $x_i = \log |g_i|^2$, which makes the problem convex (in fact, log-convex)! This class of optimization problems (of optimizing norm in the orbit of a commutative group) is known as *geometric programming* and is well-studied in the optimization literature (see, e.g.,

[1] In general, the theory works whenever the group is connected, algebraic and reductive, and our results hold in this generality. However, for purposes of exposition we only discuss very simple groups in this introduction. We also suppress some technical details which are spelled out later, e.g., that the representations are rational and map unitary matrices to unitary matrices (both are essentially without loss of generality).

[2] For notational convenience, we suppress the dependence of $\operatorname{cap}(v)$ on the group G and representation π (likewise for the null cone and the moment polytopes defined below).

Chapter 4.5 in [BV04]). Hence for non-commutative groups, one can view $\mathrm{cap}(v)$ as *non-commutative geometric programming*. Is there a similar change of variables that makes the problem convex in the non-commutative case? It does not seem so. However, the non-commutative case also satisfies a notion of convexity, known as geodesic convexity, which we will study next.

1.3.1 Geodesic convexity

Geodesic convexity generalizes the notion of convexity in the Euclidean space to arbitrary Riemannian manifolds. We will not go into the notion of geodesic convexity in this generality but just mention what it amounts to in our concrete setting of norm optimization for $G = \mathrm{GL}(n)$.

It turns out the appropriate way to define geodesic convexity in this case is as follows. Fix an action π of $\mathrm{GL}(n)$ and a vector v. Then $\log\|\pi(e^{tH}g)v\|$ is convex in the real parameter t for every Hermitian matrix H and $g \in \mathrm{GL}(n)$. This notion of convexity is quite similar to the notion of Euclidean convexity, where a function is convex iff it is convex along all lines. However, it is far from obvious how to import optimization techniques from the Euclidean setting to work in this non-commutative geodesic setting. An essential ingredient we describe next is the geodesic notion of a gradient, called the *moment map*.

1.3.2 Moment map

The moment map is by definition the gradient of the function $\log\|\pi(g)v\|$ (understood as a function of v), at the identity element of the group, $g = I$. It captures how the norm of the vector v changes when we act on it by infinitesimal perturbations of the identity.

Again, we start with the commutative case $G = T(n)$ acting on the multivariate Laurent polynomials. For a ("direction") vector $h \in \mathbb{R}^n$ and a real ("perturbation") parameter t, let $e^{th} = \mathrm{diag}\left(e^{th_1}, \ldots, e^{th_n}\right)$. Then, for $G = T(n)$, the moment map is the function $\mu\colon V\setminus\{0\} \to \mathbb{R}^n$, defined by the following property:

$$\mu(v) \cdot h = \partial_{t=0}\left[\log\left\|\pi(\mathrm{diag}(e^{th})v\right\|\right],$$

for all $h \in \mathbb{R}^n$. That is, the directional derivative in direction h is given by the dot product $\mu(v) \cdot h$. Here one can see that the moment map matches the notion of Euclidean gradient. For the action of $T(n)$ in Eq. (1.1),

$$\mu(v) = \nabla_{x=0} \log\left(\sum_{j=1}^{m} |c_j|^2 e^{x\cdot\omega_j}\right) = \frac{\sum_{j=1}^{m} |c_j|^2 \omega_j}{\sum_{j=1}^{m} |c_j|^2}. \tag{1.3}$$

Note that the gradient $\mu(v)$ at any point v is a convex combination of the weights! Viewing v as a polynomial, the gradient thus belongs to the so-called *Newton polytope* of v, namely the convex hull of the exponent vectors of its monomials! Conversely, every point in that polytope is a gradient of some polynomial v with these monomials. We will soon return to this curious fact!

We now proceed to the non-commutative case, focusing on $G = \mathrm{GL}(n)$. Denote by $\mathrm{Herm}(n)$ the set of $n \times n$ complex Hermitian matrices. Here "directions" will be parametrized by $H \in \mathrm{Herm}(n)$. For the case of $G = \mathrm{GL}(n)$, the moment map is the function $\mu\colon V \setminus \{0\} \to \mathrm{Herm}(n)$ defined (in complete analogy to the commutative case above) by the following property that

$$\mathrm{tr}[\mu(v)H] = \partial_{t=0}\left[\log\left\|\pi(e^{tH})v\right\|\right]$$

for all $H \in \mathrm{Herm}(n)$. That is, the directional derivative in direction H is given by $\mathrm{tr}[\mu(v)H]$.

Remark 1.1. *The reason we are restricting to directions in $\mathbb{R}^n$ in the $T(n)$ case and to directions in Herm_n in the $GL(n)$ case is that imaginary and skew-Hermitian directions, respectively, do not change the norm.*

In the commutative case, Eq. (1.3) is a convex combination of the weights ω_j. Thus, the image of μ is the convex hull of the weights – a convex polytope. This brings us to moment polytopes.

1.3.3 Moment polytopes

One can ask whether the above fact is true for actions of $GL(n)$ i.e., is the set $\{\mu(v) : v \in V \setminus \{0\}\}$ convex? This turns out to be blatantly false. Consider the action of $GL(n)$ on $\mathbb{C}^n$ by matrix-vector multiplication. The moment map in this setting is $\mu(v) = vv^\dagger/\|v\|^2$, and its image is clearly not convex. However, there is still something deep and non-trivial that can be said. Given a Hermitian matrix $H \in \mathrm{Herm}(n)$, define its *spectrum* to be the vector of its eigenvalues arranged in non-increasing order. That is, $\mathrm{spec}(H) := (\lambda_1, \dots, \lambda_n)$, where $\lambda_1 \geqslant \cdots \geqslant \lambda_n$ are the eigenvalues of H. Amazingly, the set of spectra of moment map images, that is,

$$\Delta := \left\{\mathrm{spec}\big(\mu(v)\big) : 0 \neq v \in V\right\}, \tag{1.4}$$

is a convex polytope for every representation π [NM84, Kos73, Ati82, GS82, Kir84a]! These polytopes are called *moment polytopes*.

Let us mention two important examples of moment polytopes. The examples are for actions of products of $GL(n)$'s but the above definitions generalize almost immediately.

Example 1.2 (Horn's problem). *Let* $G = GL(n) \times GL(n) \times GL(n)$ *act on* $V = \mathrm{Mat}(n) \oplus \mathrm{Mat}(n)$ *as follows:* $\pi(g_1, g_2, g_3)(X, Y) := (g_1 X g_3^{-1}, g_2 Y g_3^{-1})$. *The moment map in this case is*

$$\mu(X, Y) = \frac{(XX^\dagger, YY^\dagger, -(X^\dagger X + Y^\dagger Y))}{\|X\|_F^2 + \|Y\|_F^2}.$$

Using that $XX^\dagger$ *and* $X^\dagger X$ *are PSD and isospectral, we obtain the following moment polytopes, which characterize the eigenvalues of sums of Hermitian matrices, i.e., Horn's problem (see, e.g., [Ful00, BVW16]):*

$$\Delta = \left\{(\mathrm{spec}(A), \mathrm{spec}(B), \mathrm{spec}(-A-B)) \ : \ A, B \in \mathrm{Mat}(n),\ A \geqslant 0,\ B \geqslant 0,\ \mathrm{tr}\, A + \mathrm{tr}\, B = 1\right\}$$

These polytopes are known as the Horn polytopes *and correspond to Problem 3 in Section 1.2. They have been characterized mathematically in [Kly98, KT99, BK06, Res10] and algorithmically in [DLM06, MNS12, BI13].*

The preceding is one of the simplest example of a moment polytope associated with a quiver (in this case, the so-called *star quiver* with two edges, see Figure 1.1, (a)). We refer [DW17] to an introduction to quivers and briefly give the relevant definitions.

Example 1.3 (Quivers). *A* quiver *is a directed graph* Q *with loops and parallel edges allowed (see Figure 1.1 for two examples). A quiver representation of* Q *(not to be confused with a group representation!) assigns a vector space* $V_x = \mathbb{C}^{n_x}$ *to each vertex* x *and a linear map* $V_x \to V_y$ *to each edge* $x \to y$ *in* Q*. Thus, if we fix the dimensions* n_x*, the space of all quiver representations forms a vector space* $V = \bigoplus_{x \to y} \mathrm{Mat}(n_y, n_x)$*. This space carries a natural action of the group* $G = \prod_x GL(n_x)$*. We will call this action of* G *on* V *the* group representation *associated with the quiver* Q *and dimension vector* $\mathbf{n} = (n_x)$*. Regarding their moment polytopes, an important slice corresponding to semi-invariants has been characterized in [Kin94, DW00, SVdB01, Res12] and the polytopes have been described completely in [BVW18b, BVW19].*

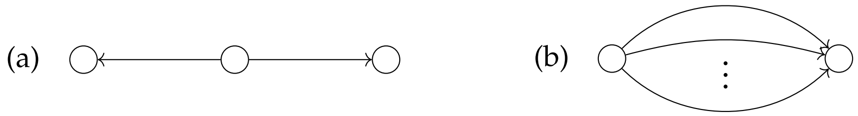

Figure 1.1: Two examples of quivers. (a) The star quiver with two edges. The associated moment polytopes model Horn's problem (Example 1.2). (b) The generalized Kronecker quiver with k parallel edges, which corresponds to a variant of the left-right action (cf. Example 1.6).

Example 1.4 (Tensor action)**.** $G = GL(n) \times GL(n) \times GL(n)$ *acts on* $V = \mathbb{C}^n \otimes \mathbb{C}^n \otimes \mathbb{C}^n$*, as follows:* $\pi(g_1, g_2, g_3)v := (g_1 \otimes g_2 \otimes g_3)v$*. We can think of vectors* $v \in V$ *as tripartite quantum states with local dimension* n*. Then the moment map for this group action captures precisely the notion of* quantum marginals*. That is,* $\mu(v) = (\rho_1, \rho_2, \rho_3)$*, where* $\rho_k = \mathrm{tr}_{k^c}(vv^\dagger)$ *denotes the reduced density matrix describing the state of the* k^{th} *particle. This corresponds to Problem 2 in Section 1.2.*

The moment polytopes in this case are known as Kronecker polytopes*, since they can be equivalently described in terms of the Kronecker coefficients of the symmetric group. These polytopes have been studied in [Kly04, DH05, CM06, CHM07, WDGC13, Wal14, CDKW14, VW17, BFG⁺18].*

There is a more refined notion of a moment polytope. One can look at the collection of spectra of moment maps of vectors in the orbit of a particular vector $v \in V$. Its closure,

$$\Delta(v) := \overline{\{\mathrm{spec}(\mu(w)) : w \in \mathcal{O}_v\}}$$

is a convex polytope as well, called the *moment polytope of* v [NM84, Bri87]! It can equivalently be defined as the spectra of moment map images of the orbit's closure in projective space.

1.3.4 Null cone

Fix a representation π of a group G on a vector space V (recall G is $T(n)$ or $GL(n)$ for the introduction). The *null cone* for this group action is defined as the set of vectors v such that $\mathrm{cap}(v) = 0$:

$$\mathcal{N} := \{v \in V : \mathrm{cap}(v) = 0\}$$

In other words, v is in the null cone if and only if 0 lies in the orbit-closure of v. It is of importance in invariant theory due to the results of Hilbert and Mumford [Hil93, Mum65] which state that the null cone is the algebraic variety defined by non-constant homogeneous invariant polynomials of the group action (see, e.g., the excellent textbooks [DK15, Stu08]).

Let us see what the null cone for the action of $T(n)$ in Eq. (1.1) is. Recall from Eq. (1.2), the formulation for $\mathrm{cap}(v)$. It is easy to see that $\mathrm{cap}(v) = 0$ iff there exists $x \in \mathbb{R}^n$ such that $x \cdot \omega_j < 0$ for all $j \in \mathrm{supp}(v)$, where $\mathrm{supp}(v) = \{j \in [m] : c_j \neq 0\}$ for $v = \sum_{j=1}^m c_j v_j$. Thus the property of v being in the null cone is captured by a simple linear program defined by $\mathrm{supp}(v)$ and the weights ω_j's. Hence the null cone membership problem for non-commutative group actions can be thought of as *non-commutative linear programming*.

We know by Farkas' lemma that there exists $x \in \mathbb{R}^n$ such that $x \cdot \omega_j < 0$ for all $j \in \mathrm{supp}(v)$ iff 0 does not lie in $\mathrm{conv}\{\omega_j : j \in \mathrm{supp}(v)\}$. In other words, $\mathrm{cap}(v) = 0$ iff $0 \notin \Delta(v)$. Is this true in the non-commutative world? It is! This is the Kempf-Ness theorem [KN79] and it is a consequence of the geodesic convexity of the function $g \to \log\|\pi(g)v\|$. The Kempf-Ness theorem can be thought of as a *non-commutative duality theory* paralleling the linear programming duality given by Farkas' lemma (which corresponds to the commutative world). Let us now mention an example of an interesting null cone in the non-commutative case.

Example 1.5 (Operator scaling, or left-right action)**.** *$G = \mathrm{SL}(n) \times \mathrm{SL}(n)$ (where $\mathrm{SL}(n)$ denotes the group of $n \times n$ matrices with determinant 1) acts on $\mathrm{Mat}(n)^k$ as follows: $\pi(g,h)(X_1,\dots,X_k) := (gX_1h^T,\dots,gX_kh^T)$. This family of actions is called the* left-right *action.*

The null cone for this action captures non-commutative singularity *(see, e.g., [IQS17b, GGOW16, DM17, IQS17a]) and Problem 1 in Section 1.2. The left-right action has been crucial in getting deterministic polynomial time algorithms for the non-commutative rational identity testing problem* [IQS17b, GGOW16, DM17, IQS17a]*. The commutative analogue is the famous polynomial identity testing (PIT) problem, for which designing a deterministic polynomial time algorithm remains a major open question in derandomization and complexity theory.*

Example 1.6 (Generalized Kronecker quivers)**.** *Also sometimes referred to as the left-right action, the action $\pi(g,h)(X_1,\dots,X_k) := (gX_1h^{-1},\dots,gX_kh^{-1})$ of $G = \mathrm{GL}(n) \times \mathrm{GL}(n)$ on k-tuples of matrices $(X_1,\dots,X_k)$ can be obtained from action of Example 1.5 via the isomorphism $h \mapsto (h^{-1})^T$ of $\mathrm{GL}(n)$. This action is also associated to a quiver, namely the* generalized Kronecker quiver *(see Example 1.3 and Figure 1.1, (b)).*

Example 1.7 (Simultaneous conjugation)**.** *Similarly, $G = \mathrm{GL}(n)$ acts on k-tuples of matrices in $V = (\mathrm{Mat}(n))^d$ by $\pi(g)(X_1,\dots,X_k) := (gX_1g^{-1},\dots,gX_kg^{-1})$. This example is associated to the quiver with a single vertex and k self-loops (cf. Example 1.3).*

1.4 Computational problems and state of the art

In this section, we describe the main computational questions that are of interest for the optimization problems discussed in the previous section and then discuss what is known about them in the commutative and non-commutative worlds.

Problem 1.8 (Null cone membership)**.** *Given (π, v), determine if v is in the null cone, i.e., if $\mathrm{cap}(v) = 0$. Equivalently, test if $0 \notin \Delta(v)$.*

The null cone membership problem for $\mathrm{GL}(n)$ is interesting only when the action $\pi(g)$ is given by rational functions in the $g_{i,j}$ rather than polynomials. This is completely analogous to the commutative case (e.g., the convex hull of weights ω_j with positive entries never contains the origin). In the important case that π is homogeneous, the null cone membership problem is interesting precisely when the total degree is zero, so that scalar multiples of the identity matrix act trivially. Thus, in this case the null cone membership problem for $G = \mathrm{GL}(n)$ is equivalent to the one for $G = \mathrm{SL}(n)$. We will come back to this perspective in Section 1.6.

Problem 1.9 (Scaling)**.** *Given (π, v, ε) such that $0 \in \Delta(v)$, output a group element $g \in G$ such that $\|\mathrm{spec}(\mu(g)v)\|_2 = \|\mu(\pi(g)v)\|_F \leqslant \varepsilon$.*

In particular, the following promise problem can be reduced to Problem 1.9: Given (π, v, ε), decide whether $0 \notin \Delta(v)$ under the promise that either $0 \in \Delta(v)$ or 0 is ε-far from $\Delta(v)$. In fact, there always exists $\varepsilon > 0$, depending only on the group action, such that this promise is satisfied! Thus, the null cone membership problem can always be reduced to the scaling problem (see Corollary 1.18 below).

We develop theory in Section 3.4 showing that an efficient agorithm to minimize the norm on an orbit closure of a vector v (i.e., approximate the capacity of v) under the promise that $0 \in \Delta(v)$ results in an efficient algorithm for the scaling problem and hence for the null cone membership problem. This motivates our next computational problem.

Problem 1.10 (Norm minimization)**.** *Given* (π, v, ε) *such that* $\operatorname{cap}(v) > 0$*, output a group element* $g \in G$ *such that* $\log\|\pi(g)v\| - \log\operatorname{cap}(v) \leqslant \varepsilon$*.*

We also consider the moment polytope membership problem for an arbitrary point $p \in \mathbb{Q}^n$.

Problem 1.11 (Moment polytope membership)**.** *Given* (π, v, p)*, determine if* $p \in \Delta(v)$*.*

The moment polytope membership problem is more general than the null cone membership problem, but there is a reduction from the former to the latter via the "shifting trick" in the next subsection. This forms the basis of our algorithms for the moment polytope membership problem. As in the case of the null cone, we consider a scaling version of the moment polytope membership problem.

Problem 1.12 (p-scaling)**.** *Given* (π, v, p, ε) *such that* $p \in \Delta(v)$*, output an element* $g \in G$ *such that* $\|\operatorname{spec}(\mu(\pi(g)v)) - p\|_2 \leqslant \varepsilon$*.*

The above problem has been referred to as *nonuniform scaling* [BFG$^+$18] or, for operators, matrices and tensors, as *scaling with specified or prescribed marginals* [Fra18]. The following problem can be reduced to Problem 1.12: Given (π, v, p, ε), decide whether $p \in \Delta(v)$ under the promise that either $p \in \Delta(v)$ or p is ε-far from $\Delta(v)$. We later combine the shifting trick with our duality theory to show that there is a value $\varepsilon > 0$ with bitsize polynomial in the input size such that this is promise is always satisfied. Thus, the moment polytope membership problem can be reduced to p-scaling (see Corollary 3.31 in Section 3.6 and Lemma 7.18 in Section 1.6).

There are multiple interesting input models for these problems. One could explicitly describe the weights $\omega_1, \ldots, \omega_m$ for an action of $T(n)$ (Eq. (1.1)) and then describe v as $\sum_{j=1}^m c_j v_j$ by describing the c_j's. The analogous description in the non-commutative world would be to describe the irreducible representations occuring in V. Alternately, one could give black box access to the function $\|\pi(g)v\|$, or to the moment map $\mu(\pi(g)v)$, etc. Sometimes π can be a non-uniform input as well, such as a fixed family of representations like the simultaneous left-right action Example 1.5 as done in [GGOW16]. The inputs p and ε will be given in their binary descriptions but we will see that some of the algorithms run in time polynomial in their unary descriptions.

Remark 1.13 (Running time in terms of ε)**.** *By standard considerations about the bit complexity of the facets of the moment polytope, it can be shown that polynomial time algorithms for the scaling problems (Problems 1.9 and 1.12) result in polynomial time algorithms for the exact versions (Problems 1.8 and 1.11, respectively). Polynomial time requires, in particular,* $\operatorname{poly}(\log(1/\varepsilon))$ *dependence on* ε*; a* $\operatorname{poly}(1/\varepsilon)$ *dependence is only known to suffice in special cases.*

1.4.1 Commutative groups and geometric programming

In the commutative case, the preceding problems are reformulations of well-studied optimization problems and much is known about them computationally. To see this, consider the action of $T(n)$ as in Eq. (1.1), and a vector $v = \sum_{j=1}^m c_j v_j$. It follows from Section 1.3.4 that v is in the null cone iff $0 \notin \Delta(v) = \operatorname{conv}\{\omega_j : c_j \neq 0\}$. Recall from Eq. (1.2), the formulation for $\operatorname{cap}(v)$. Since this formulation is convex, it follows that, given $\omega_1, \ldots, \omega_m \in \mathbb{Z}^n$ (recall this is the description of π) and $c_1, \ldots, c_m \in \mathbb{Q}[i]$ (each entry described in binary), there is a polynomial-time algorithm for the null cone membership problem via linear programming [Kha79, Kar84]. The same is true for the moment polytope membership problem. The capacity optimization problem is an instance of *(unconstrained) geometric programming* and one can design polynomial time algorithms in the same input model. It is hard to find an exact reference for this, but this can be done, for example, using the ellipsoid algorithm as done for the same problem in slightly different settings in the papers [Gur04b, SV14, SV19]. There has been work in the oracle setting as well, in which one has oracle access to the function $\|\pi(g)v\|$. The advantage of the oracle setting is that one can handle exponentially large representations of $T(n)$ when it is not possible to describe all the weights explicitly. A very general result of this form is proved in [SV19]. While not explicitly mentioned in [SV19], their techniques can also be used to design polynomial time algorithms for *commutative* null cone and moment polytope membership in the oracle setting. Thus, in the commutative case, Problems 1.8, 1.10 and 1.11 are well-understood.

1.4.2 Non-commutative actions

Comparatively very little is known in the non-commutative case. The two non-trivial group actions for which there are known polynomial-time algorithms for null cone membership (Problem 1.8) are *simultaneous conjugation* [RS05, FS13] and the *left-right* action [IQS17b, GGOW16, DM17, IQS17a]. Approximate algorithms for null cone membership have been designed for the *tensor action* of products of $\mathrm{SL}(n)$'s [BGO+17]. However the running time is exponential in the binary description of ε (i.e., polynomial in $1/\varepsilon$). This is the reason the algorithm does not lead to a polynomial time algorithm for the exact null cone membership problem for the tensor action.

Moment polytope membership is already interesting for the polytope Δ in (1.4), the moment polytope of the entire representation V (not restricted to any orbit closure). Even here, efficient algorithms are only known in very special cases, such as for the Horn polytope (Example 1.2) [DLM06, MNS12, BI13]. The structural results in [BS00, Res10, VW17] characterize Δ in terms of linear inequalities (it is known that in general there are exponentially many). Mathematically, this is related to the asymptotic vanishing of certain representation-theoretic multiplicities [Bri87, CDW12, BVW18a] whose non-vanishing is in general NP-hard to decide [IMW17]. [BCMW17] proved that the membership problem for Δ is in NP $\cap$ coNP. As Δ and $\Delta(v)$ coincide for generic $v \in V$, this problem captures the moment polytope membership problem (Problem 1.11) for almost all vectors (all except those in a set of measure zero).

The study of Problem 1.11 in the noncommutative case focused on *Brascamp-Lieb polytopes* (which are affine slices of moment polytopes). [GGOW17] solved the moment polytope membership problem in time depending polynomially on the *unary* complexity of the target point. In [BFG+18], efficient algorithms were designed for the p-scaling problem (Problem 1.12) for tensor actions, extending the earlier work of [Fra18] for the simultaneous left-right action. The running times of both algorithms are $\operatorname{poly}(1/\varepsilon)$; for this reason both algorithms result in moment polytope algorithms depending exponentially on the binary bitsize of p, as in [GGOW17].

Regarding the approximate computation of the capacity (Problem 1.10), efficient algorithms were previously known only for the simultaneous left-right action. [GGOW16] gave an algorithm to approximate the capacity in time polynomial in all of the input description except ε, on which it had dependence $\operatorname{poly}(1/\varepsilon)$. The paper [AZGL+18] gave an algorithm that depended polynomially on the input description; it has running time dependence $\mathrm{poly}(\log(1/\varepsilon))$ on the error parameter ε.

In terms of algorithmic techniques, all prior works that were based on optimization methods fall into two categories. One is that of *alternating minimization* (which can be thought of as a large-step coordinate gradient descent, i.e., roughly speaking as a first order method). However, alternating minimization is limited in applicability to 'multilinear' actions of products of $T(n)$'s or $GL(n)$'s, where the action is linear in each component so that it is easy to optimize over one component when fixing all the others. This is true for all the actions described above and hence explains the applicability of alternating minimization (in fact, in all the above examples, one can even get a closed-form expression for the group element that has to be applied in each alternating step). The second category are geodesic analogues of *box-constrained Newton's methods* (second order). Recently, [AZGL+18] designed an algorithm tailored towards the specific case of the simultaneous left-right action (Example 1.5), but no second order algorithms were known for other group actions. However, many group actions of interest – from classical problems in invariant theory about symmetric forms to the important variant of Problem 2 in Section 1.2 for fermions – are not multilinear nor can otherwise be captured by the left-right action, and no efficient algorithms were known. All this motivates the development of new techniques.

In this paper, we show how these limitations can be overcome. Specifically, we provide both first and second order algorithms (geodesic variants of gradient descent and box-constrained Newton's method) that apply in great generality and identify the main structural parameters that control the running time of these algorithm. We now describe our contributions in more detail.

1.5 Algorithmic and structural results

We describe here our algorithmic and structural contributions to the theory of non-commutative optimization. In Section 1.5.1, we describe the main parameters that govern the running time of our algorithms. In Section 1.5.2, we describe the first order algorithm for $\mathrm{cap}(v)$ and the structural results we prove for its analysis. In Section 1.5.3, we describe a first order algorithm for the problem of membership in moment polytopes and the relevant structural results. In Section 1.5.4, we describe the second order algorithm for $\mathrm{cap}(v)$ and the techniques and ideas used in its analysis.

1.5.1 Essential parameters and structural results

In this section, we define the essential parameters related to the group action which, in addition to dictating the running times of our first and second order methods, control the relationships between the null cone, the norm of the moment map, and the capacity, i.e., between Problems 1.8 to 1.10.

We saw in Section 1.3 that for all actions of $T(n)$ on a vector space V, one can find a basis of V consisting of simultaneous eigenvectors of the matrices $\pi(g)$, $g \in T(n)$. While this is in general impossible for non-commutative groups, one can still decompose V into building blocks known as irreducible subspaces (or subrepresentations), as will be discussed in further detail in Section 2.3.

For $GL(n)$, these are uniquely characterized by nonincreasing sequences $\lambda \in \mathbb{Z}^n$; such sequences λ are in bijection with irreducible representations $\pi_\lambda\colon GL(n) \to GL(V_\lambda)$. We say that λ *occurs in* π if one of its irreducible subspaces is of type λ. If all the λ occuring in π have nonnegative entries, then the entries of the matrix $\pi(g)$ are polynomials in the entries of g. Such representations π are called *polynomial*, and if all λ occuring in π have sum exactly (resp. at most) d, then π is said to be a *homogeneous polynomial representation of degree (resp. at most)* d. We elaborate further on the representation theory of $\mathrm{GL}(n)$ in Section 1.6, Section 2.3, and Section 7.1.

Now we can define the complexity measure which captures the smoothness of the optimization problems of interest. Later on in Section 3.3 we discuss how to think of the following measure as a *norm* of the Lie algebra representation Π, hence the name *weight norm*.

Definition 1.14 (Complexity measure I: weight norm)**.** *We define the* weight norm $N(\pi)$ *of an action* π *of* $\mathrm{GL}(n)$ *by* $N(\pi) := \max\{\|\lambda\|_2 : \lambda$ *occurs in* $\pi\}$, *where* $\|\cdot\|_2$ *denotes the Euclidean norm.*

Another use of the weight norm is to provide a bounding ball for the moment polytope. As shown in Lemma 3.11, the moment polytope is contained in a Euclidean ball of radius $N(\pi)$. The weight norm is in turn controlled by the degree of a polynomial representation. More specifically, if π is a polynomial representation of $\mathrm{GL}(n)$ of degree at most d, then $N(\pi) \leqslant d$.

We now describe our second measure of complexity which will govern the running time bound for our second order algorithm. This parameter, which will be discussed further in Section 3.4, also features in Theorem 1.16 concerning quantitative non-commutative duality.

Definition 1.15 (Complexity measure II: weight margin)**.** *The* weight margin $\gamma(\pi)$ *of an action* π *of* $\mathrm{GL}(n)$ *is the minimum Euclidean distance between the origin and the convex hull of any subset of the weights of* π *that does not contain the origin.*

Our running time bound will depend inversely on the weight margin. Two interesting examples with large (inverse polynomial) weight margin are the left-right action (Example 1.5) and simultaneous conjugation. The existing second order algorithm for the left-right action relied on the large weight margin of the action [AZGL+18]. It is interesting that the simultaneous conjugation action (Example 1.7), the sole other interesting example of an action of a non-commutative group for which there are efficient algorithms for the null cone membership problem [RS05, FS13, DM18] (which have nothing to do with the weight margin), also happens to have large weight margin! On the other hand, the only generally applicable lower bound on the weight margin is $n^{-1}N(\pi)^{-n}$ (see Theorem 6.8), and indeed this exponential behavior is seen for the somewhat intractable 3-tensor action (Example 1.4), which has weight margin at most $2^{-n/3}$ and weight norm $\sqrt{3}$ (implicit in [Kra07]). For the convenience of the reader, we arrange in a tabular form the above information about the weight margin and weight norm for various paradigmatic group actions in Table 1.1 (using a definition of the weight margin and weight norm, given later in the paper, that naturally generalizes the one given above for $\mathrm{GL}(n)$):

Group action	**Weight margin** $\gamma(\pi)$	**Weight norm** $N(\pi)$
Matrix scaling[3]	$\geqslant n^{-3/2}$; [LSW98] and (6.4)	$\sqrt{2}$ (Example 6.3)
Simultan. left-right action (Example 1.5)	$\geqslant n^{-3/2}$; [Gur04a] and (6.4)	$\sqrt{2}$ (Example 6.3)
Quivers (Example 1.3)	$\geqslant (\sum_x n_x)^{-3/2}$ (Prop. 6.12)	$\sqrt{2}$ (Proposition 6.12)
Simultaneous conjugation (Example 1.7)	$\geqslant n^{-3/2}$ (Corollary 6.13)	$\sqrt{2}$ (Corollary 6.13)
3-tensor action (Example 1.4)	$\leqslant 2^{-n/3}$; implicit in [Kra07]	$\sqrt{3}$ (Example 6.4)
Polynomial $\mathrm{GL}(n)$-action of degree d	$\geqslant d^{-n}n^{-1}$ (Theorem 6.8)	$\leqslant d$ (Lemma 6.1)
Polynomial $\mathrm{SL}(n)$-action of degree d	$\geqslant d^{-n}n^{-3/2}$ (Theorem 6.9)	$\leqslant d$ (Remark 6.2)

Table 1.1: Weight margin and norm for various representations (see Section 6 for more).

As the moment map is the gradient of the geodesically convex function $\log\|v\|$, it stands to reason that as $\mu(v)$ tends to zero, $\|v\|$ tends to the capacity $\operatorname{cap}(v)$. However, in order to use this relationship to obtain efficient algorithms, we need this to hold in a precise quantitative sense. To this end, in Section 3.4 we show the following fundamental relation between the capacity and the norm of the moment map.

Theorem 1.16 (Noncommutative duality). *For $v \in V \setminus \{0\}$ we have*

$$1 - \frac{\|\mu(v)\|_F}{\gamma(\pi)} \leqslant \frac{\operatorname{cap}(v)^2}{\|v\|^2} \leqslant 1 - \frac{\|\mu(v)\|_F^2}{4N(\pi)^2}.$$

Equipped with these inequalities, it is easy to relate Problems 1.9 and 1.10.

Corollary 1.17. *An output g for the norm minimization problem on input (π, v, ε) is a valid output for the scaling problem on input $(\pi, v, N(\pi)\sqrt{8\varepsilon})$. If $\varepsilon/\gamma(\pi) < \frac{1}{2}$ then an output g for the scaling problem on input (π, v, ε) is a valid output for the norm minimization problem on input $(\pi, v, \frac{2\log(2)\varepsilon}{\gamma(\pi)})$.*

Because $0 \in \Delta(v)$ if and only if $\operatorname{cap}(v) > 0$, Theorem 1.16 and Corollary 1.17 immediately yield the accuracy to which we must solve the scaling problem or norm minimization problem to solve the null cone membership problem:

Corollary 1.18. *It holds that $0 \in \Delta(v)$ if and only if $\Delta(v)$ contains a point of norm smaller than $\gamma(\pi)$. In particular, solving the scaling problem with input $(\pi, v, \gamma(\pi)/2)$ or the norm minimization problem with $(\pi, v, \frac{1}{8}(\gamma(\pi)/2N(\pi))^2)$ suffices to solve the null cone membership problem for (π, v).*

Corollary 3.31 in Section 3.6 provides an analogue of the above corollary for the moment polytope membership problem.

1.5.2 First order methods: structural results and algorithms

As discussed above, in order to approximately compute the capacity in the commutative case, one can just run a Euclidean gradient descent on the convex formulation in Eq. (1.2). We will see that gradient descent method naturally generalizes to the non-commutative setting. It is worth mentioning that there are several excellent sources of the analysis of gradient descent algorithms for geodesically convex functions (in the general setting of Riemannian manifolds and not just the group setting that we are interested in); see e.g., [Udr94, AMS09, ZS16, ZRS16, SKM19, ZS18] and references therein. In this paper, our contribution is mostly in understanding the geometric properties (such as smoothness) of the optimization problems that we are concerned with, which allow us to carry out the classical analysis of Euclidean gradient descent in our setting.

The natural analogue of gradient descent for the optimization problem $\operatorname{cap}(v)$ is the following: start with $g_0 = I$ and repeat, for T iterations and a suitable step size η:

$$g_{t+1} = e^{-\eta\mu(\pi(g_t)v)} g_t.$$

Finally, return the group element g among $g_0, \ldots, g_{T-1}$, which minimizes $\|\mu(\pi(g)v)\|_F$. This algorithm is described in Algorithm 4.2. A natural geometric parameter which governs the

[3] This commutative example is modelled as follows: $G = ST(n) \times ST(n)$ acts on $\operatorname{Mat}(n)$ by $\pi(A, B)M = AMB$, where $ST(n)$ is the group of diagonal $n \times n$ matrices with unit determinant.

complexity (number of iterations T, step size η) of gradient descent is the *smoothness* of the function to be optimized. The smoothness parameter for actions of $T(n)$ in Eq. (1.1) can be shown to be $O(\max_{j\in[m]}\|\omega_j\|_2^2)$ (see, e.g., [SV19]), which is the square of the weight norm defined in Definition 1.14 for this action. We prove in Section 3 that, in general, the function $\log\|\pi(g)v\|$ is geodesically smooth, with a smoothness parameter which, analogously to the commutative case, is on the order of the square of the weight norm. We now state the running time for our geodesic gradient descent algorithm for Problem 1.9.

Theorem 1.19 (First order algorithm for scaling)**.** *Fix a representation* $\pi : \mathrm{GL}(n) \to \mathrm{GL}(V)$ *and a unit vector* $v \in V$ *such that* $\mathrm{cap}(v) > 0$ *(i.e., v is not in the null cone). Then Algorithm 4.2 with a number of iterations at most*

$$T = O\left(\frac{N(\pi)^2}{\varepsilon^2}\left|\log \mathrm{cap}(v)\right|\right)$$

outputs a group element $g \in G$ *satisfying* $\|\mu(\pi(g)v)\|_F \leqslant \varepsilon$*.*

This is proved in Section 4, where it is stated as Theorem 4.2 for general groups. Theorem 1.23 in Section 1.6 states concrete running time bounds in terms of the bit complexity of the input.

The analysis of Theorem 1.19 relies on the smoothness of the function $F_v(g) := \log\|\pi(g)v\|$, which implies that

$$F_v(e^H g) \leqslant F_v(g) + \mathrm{tr}\big[\mu\big(\pi(g)v\big)H\big] + N(\pi)^2\|H\|_F^2,$$

for all $g \in \mathrm{GL}(n)$ and for all Hermitian $H \in \mathrm{Herm}(n)$ (see Cor. 3.13).

1.5.3 First order method for moment polytope membership

Next, we describe our first order algorithm for the p-scaling problem. Theorem 1.19 solves the problem of minimizing the moment map (equivalent to capacity computation), hence can be used to determine if $0 \in \Delta(v)$. Can we reduce the general moment polytope membership problem, $p \in \Delta(v)$, to this case? This is straightforward in the commutative case, $G = T(n)$. It follows from the reasoning in Section 1.3.4 that, for $p \in \mathbb{R}^n$, we have $p \notin \Delta(v)$ iff

$$\mathrm{cap}_p(v)^2 := \inf_{x\in\mathbb{R}^n} \textstyle\sum_{j=1}^m |c_j|^2 e^{x\cdot(\omega_j - p)} = 0. \tag{1.5}$$

Thus, all we need to do is shift the relevant vectors by p. Is there an analog of this trick in the non-commutative world? There is! It is called, unsurprisingly, the *shifting trick* [Bri87]. Let us describe it here. A nice property about Eq. (1.5) is that (recall Eq. (1.3)) $\nabla_{x=0} \log\left(\sum_{j=1}^m |c_j|^2 e^{x\cdot(\omega_j-p)}\right) = \mu(v) - p$. How do we shift the moment map in the case of $\mathrm{GL}(n)$? It relies on the following two elementary properties of the moment map:

1. The moment map of the tensor product π of two representations π_1, π_2 of $\mathrm{GL}(n)$, which is defined as $\pi(g)(v \otimes w) := (\pi_1(g)v) \otimes (\pi_2(g)w)$, satisfies $\mu(v \otimes w) = \mu(v) + \mu(w)$.

2. There is a vector v_λ (known as a *highest weight vector*) in the vector space V_λ of the irreducible action π_λ such that $\mu(v_\lambda) = \mathrm{diag}(\lambda)$.

Now suppose $p \in \mathbb{Q}^n$ and let $\ell > 0$ be the least integer such that $\lambda := \ell p \in \mathbb{Z}^n$. Let $\lambda^* := (-\lambda_n, \ldots, -\lambda_1)$. Then one can see that the tensor product action of $\mathrm{GL}(n)$ on the space $\mathrm{Sym}^\ell(V) \otimes V_{\lambda^*}$ satisfies $\frac{1}{\ell}\mu\left(v^{\otimes \ell} \otimes v_{\lambda^*}\right) = \mu(v) + \mathrm{diag}(\lambda^*)/\ell = \mu(v) - \Lambda$, where Λ is the diagonal matrix with entries $\Lambda_{i,i} = p_{n-i+1}$, which has spectrum p. We have managed to shift the moment map! So we are led to the following optimization problem,

$$\mathrm{cap}_p(v)^\ell := \inf_{g \in G} \|(\pi(g)v)^{\otimes \ell} \otimes (\pi_{\lambda^*}(g)v_{\lambda^*})\|.$$

In the noncommutative case, the relation between this p-*capacity* and the moment polytope is slightly more subtle. While $\mathrm{cap}_p(v) > 0$ always guarantees that $p \in \Delta(v)$, these two conditions are in general *not* equivalent (unless $p = 0$, when $\mathrm{cap}_p(v)$ reduces to $\mathrm{cap}(v)$). However, what holds is that $p \in \Delta(v)$ if and only if $\mathrm{cap}_p(\pi(g)v) > 0$ for *generic* $g \in G$. We can thus solve the p-scaling problem by first applying a random group element and then applying an optimization algorithm to approximate $\mathrm{cap}_p(v)$.

We now outline our optimization algorithm for $\mathrm{cap}_p(v)$. The optimization problem defining $\mathrm{cap}_p(v)$ is defined in terms of actions on a space of exponential dimension. However, it turns out that the gradients can be explicitly computed and the geodesic gradient descent can be described explicitly as follows: start with $g_0 = I$ and repeat, for T iterations and suitable step size η:

$$g_{t+1} = e^{-\eta\left(\mu(\pi(g_t)v) - Q_t \Lambda Q_t^\dagger\right)} g_t,$$

where $g_t = Q_t R_t$ is the QR decomposition of g_t. Finally return group element g among $g_0, \ldots, g_{T-1}$, which minimizes $\|\mu(\pi(g_t)v) - Q_t \Lambda Q_t^\dagger\|_F$. This algorithm is stated precisely as Algorithm 4.3.

Theorem 1.20 (First order algorithm for p-scaling)**.** *Fix a representation $\pi : \mathrm{GL}(n) \to \mathrm{GL}(V)$, a unit vector $v \in V$, and a target point $p \in \mathbb{Q}^n$ such that $\mathrm{cap}_p(v) > 0$. Let $N^2 := N(\pi)^2 + \|p\|_2$. Then Algorithm 4.3 with a number of iterations at most*

$$T = O\left(\frac{N^2}{\varepsilon^2}\left|\log \mathrm{cap}_p(v)\right|\right)$$

outputs a group element $g \in G$ satisfying $\|\mathrm{spec}(\mu(\pi(g)v)) - p\|_2 \leqslant \varepsilon$.

This is proved in Section 4.3, where it is stated as Theorem 4.5 for general groups. A precise calculation of the smoothness of the function $g \mapsto \log\|\pi(g)v\| + \frac{1}{\ell}\log\|\pi_{\lambda^*}(g)v_{\lambda^*}\|$ (which underlies the p-capacity) features crucially in our analysis.

As described above, Theorem 1.20 preceded by a randomization step can be used to solve the p-scaling problem. Theorem 1.26 in Section 1.6 describes the performance of such a randomized algorithm for $G = \mathrm{GL}(n)$.

1.5.4 Second order methods: structural results and algorithms

Here we discuss our second order algorithm for Problem 1.10, the approximate norm minimization problem. As mentioned in Section 1.4, the paper [AZGL+18] (following the algorithms developed in [AZLOW17, CMTV17] for the commutative Euclidean case) developed a second order polynomial-time algorithm for approximating the capacity for the simultaneous left-right action (Example 1.5) with running time polynomial in the bit description of the approximation parameter ε. In Section 5, we generalize this algorithm to arbitrary groups and actions (Algorithm 5.1). It repeatedly optimizes quadratic Taylor expansions of the objective in a small neighbourhood. Such algorithms also go by the name "trust-region methods" in the Euclidean optimization literature [CGT00]. The running time of our algorithm will depend inversely on the weight margin defined in Definition 1.15.

Theorem 1.21 (Second-order algorithm for norm minimization)**.** *Fix a representation* $\pi : \mathrm{GL}(n) \to \mathrm{GL}(V)$ *and a unit vector* $v \in V$ *such that* $\operatorname{cap}(v) > 0$*. Put* $C := |\log \operatorname{cap}(v)|$, $\gamma := \gamma(\pi)$ *and* $N := N(\pi)$*. Then Algorithm 5.1 for a suitably regularized objective function outputs* $g \in G$ *satisfying* $\log \|\pi(g)v\| \leqslant \log \operatorname{cap}(v) + \varepsilon$ *with a number of iterations at most*

$$T = O\left(\frac{N\sqrt{n}}{\gamma}\left(C + \log\frac{n}{\varepsilon}\right)\log\frac{C}{\varepsilon}\right).$$

This is proved in Section 5, where it is restated precisely as Theorem 5.6. Theorem 1.24 in Section 1.6 specializes Theorem 1.21 to the group $\mathrm{SL}(n)$ by obtaining running time bounds in terms of the bit complexity of the input.

The two main structural parameters which govern the runtime of Algorithm 5.1 in general are the *robustness* (controlled by the weight norm) and a *diameter bound* (controlled by the weight margin). The robustness of a function bounds third derivatives in terms of second derivatives, similarly to the well-known notion of self concordance (however, in contrast to the latter, the robustness is not scale-invariant). As a consequence of the robustness, we show that the function $F_v(g) = \log\|\pi(g)v\|$ is sandwiched between two quadratic expansions in a small neighbourhood:

$$F(g) + \partial_{t=0}F(e^{tH}g) + \frac{1}{2e}\partial^2_{t=0}F(e^{tH}g) \leqslant F(e^H g) \leqslant F(g) + \partial_{t=0}F(e^{tH}g) + \frac{e}{2}\partial^2_{t=0}F(e^{tH}g)$$

for every $g \in \mathrm{GL}(n)$ and $H \in \mathrm{Herm}(n)$ such that $\|H\|_F \leqslant 1/(4N(\pi))$ (see Section 3).

Another ingredient in the analysis of Algorithm 5.1 is to prove the existence of "well-conditioned" approximate minimizers, i.e. $g_\star \in G$ with small condition number satisfying $\log\|\pi(g_\star)v\| \leqslant \log \operatorname{cap}(v) + \varepsilon$. The bound on the condition numbers of approximate minimizers helps us ensure that the algorithm's trajectory always lies in a compact region with the use of appropriate regularizers. As in [AZGL$^+$18], we obtain this "diameter bound" by designing a suitable gradient flow and bounding the (continuous) time it takes for it to converge (Proposition 5.5)! A crucial ingredient of this analysis is our Theorem 1.16 relating capacity and norm of the moment map.

Our gradient flow approach, which can be traced back to works in symplectic geometry [Kir84b], is the only one we know for proving diameter bounds in the non-commutative case. In contrast, in the commutative case several different methods are available (see, e.g., [SV14, SV19]). It is an important open problem to develop alternative methods for diameter bounds in the non-commutative case, which will also lead to improved running time bounds for algorithms like Algorithm 5.1.

1.6 Explicit time complexity bounds for $\mathrm{SL}(n)$ and $\mathrm{GL}(n)$

Moving beyond the number of oracle calls, we now describe the running time of our algorithms in terms of the bitsize needed to describe the vector v and the action π. For concreteness, we restrict to *homogeneous, polynomial* actions of $\mathrm{GL}(n)$, i.e., those for which there is a degree d such that entries of $\pi(g)$ are homogeneous polynomials of degree d in the matrix entries $g_{i,j}$. This important class includes the setting studied by Hilbert in his seminal paper [Hil93]. The results in this section extend readily to groups that are products of $\mathrm{GL}(n)$'s, a setting which captures all of the interesting examples discussed so far (tensor scaling, left-right action, simultaneous conjugation action, etc).

Up to isomorphism, the irreducible polynomial representations of $\mathrm{GL}(n)$ can be specified by *partitions* of length at most n, or nonincreasing vectors in $\mathbb{Z}^n_{\geqslant 0}$; the partition corresponding to an irreducible representation is called its *highest weight*. If λ is a partition of (sums to) d then the corresponding representation is homogeneous of degree d.

We must specify our input in such a way that the group action and moment map can be efficiently computed. To this end, if λ is a partition, we take $\pi_\lambda \colon \mathrm{GL}(n) \to \mathrm{GL}(m_\lambda)$ to be the irreducible representation of highest weight λ such that the standard basis of $\mathbb{C}^{m_\lambda}$ is a *Gelfand-Tsetlin basis*. The Gelfand-Tsetlin basis, described in Section 7.1, is a well-studied basis for irreducible representations in which the entries of π_λ are polynomials with rational coefficients that we can effectively bound.

A list of partitions $\lambda^1, \ldots, \lambda^s$ specifies the representation $\pi\colon \mathrm{GL}(n) \to \mathrm{GL}(m)$ on $V = \mathbb{C}^m$ given by $\pi := \oplus_{i=1}^s \pi_{\lambda^i}$, where $m := \sum_{i=1}^s m_{\lambda^i}$. Up to isomorphism, every finite-dimensional polynomial representation π of $\mathrm{GL}(n)$ can be obtained this way. If π is such a representation, the input size $\langle \pi \rangle$ of π is defined to be $\langle \lambda^1 \rangle + \cdots + \langle \lambda^s \rangle$ where $\langle \lambda^i \rangle$ is the total binary size of the entries of λ^i.

For a vector $v \in \mathbb{C}^m$ with coordinates in $\mathbb{Q} + i\mathbb{Q}$, $\langle v \rangle$ refers to the total binary size of its entries. In [Bür00, BCMW17] it is shown that, for π and v specified as above and for $g \in \mathrm{Mat}(n, \mathbb{Q} + i\mathbb{Q})$ specified in binary, the group action $\pi(g)v$ and moment map $\mu(v)$ can be computed in polynomial time. If ε is a rational number, $\langle \varepsilon \rangle$ refers to its size in binary.

We now define instances for the problems discussed in Section 1.4 for $\mathrm{GL}(n)$ and $\mathrm{SL}(n)$. We will assume that π is polynomial and homogeneous of degree d. We may assume that any target spectrum p for the moment polytope membership and p-scaling problems has nonnegative, rational entries adding to d, because every element of $\Delta(v)$ necessarily has this property. For the scaling (equivalently, norm minimization) and null cone membership problems, we consider the restriction of π to the smaller group $\mathrm{SL}(n)$. This is without loss of generality because, unless $d = 0$, the capacity for homogeneous actions of $\mathrm{GL}(n)$ is always zero (see discussion below Problem 1.8). In fact, the scaling problem for $\mathrm{SL}(n)$ is equivalent to the p-scaling problem for $\mathrm{GL}(n)$ for p a suitable multiple of the all-ones vector. This captures many natural scaling problems.

1. A tuple (π, v) is called an *instance of the null cone membership problem for* $\mathrm{SL}(n)$ if
   - $\pi\colon \mathrm{GL}(n) \to \mathrm{GL}(m)$ is a homogeneous, polynomial representation of $\mathrm{GL}(n)$ of degree $d > 0$, specified by a list of partitions,
   - $v \in V = \mathbb{C}^m$ is a Gaussian integer vector, i.e., its entries are in $\mathbb{Z} + i\mathbb{Z}$.

2. A tuple (π, v, ε) is called an *instance of the scaling problem for* $\mathrm{SL}(n)$ if (π, v) is an instance of the null cone membership problem for $\mathrm{SL}(n)$ and $\varepsilon > 0$ is a rational number.

3. A tuple (π, v, p) is an *instance of the moment polytope membership problem for* $\mathrm{GL}(n)$ if (π, v) is an instance of the null cone membership problem for $\mathrm{SL}(n)$ and $p \in \mathbb{Q}^n$ is a vector with entries $p_1 \geqslant \cdots \geqslant p_d \geqslant 0$ adding to d.

4. A tuple (π, v, p, ε) is an *instance of the* p*-scaling problem for* $\mathrm{GL}(n)$ if (π, v, p) is an instance of the moment polytope membership problem for $\mathrm{GL}(n)$ and $\varepsilon > 0$ is rational number.

Remark 1.22 (Degree versus dimension)**.** *We may assume that for our input representations* $\pi = \oplus_{i=1}^s \pi_{\lambda^i}$ *we have* $\lambda^i_n = 0$ *for some* $i \in [s]$*; this is without loss of generality because simultaneously translating each* λ^i

by an integer multiple of the all-ones vector simply shifts the entire moment polytope in $\mathbb{R}^n$ *by the same vector. If some* $\lambda^i_n = 0$*, then the bound* $d \leqslant m$ *follows from classical formulae for the dimensions of irreducible representations, which ensures that our bounds in the coming theorems are polynomial in* $\langle v \rangle, \langle \pi \rangle$*.*

By deriving capacity lower bounds for vectors of bounded bit complexity, we prove in Section 7 that Theorem 1.19 implies the following time bound for the scaling problem. Theorem 1.23 as well as all the other results below are proved in Section 7.4.

Theorem 1.23 (First order algorithm for scaling in terms of input size)**.** *Let* (π, v, ε) *be an instance of the scaling problem for* $\mathrm{SL}(n)$ *such that* $0 \in \Delta(v)$ *and every entry of* v *is bounded in absolute value by* M*. Let* d *denote the degree and* m *the dimension of* π*. Then, Algorithm 4.2 with a number of iterations at most*

$$T = O\left(\frac{d^3}{\varepsilon^2} mn^3 \log(Mmnd)\right)$$

returns a group element $g \in \mathrm{SL}(n)$ *such that* $\|\mu(\pi(g)v)\|_F \leqslant \varepsilon$*. In particular, there is a* $\mathrm{poly}(\langle \pi \rangle, \langle v \rangle, \varepsilon^{-1})$ *time algorithm to solve the scaling problem (Problem 1.9) for* $\mathrm{SL}(n)$*.*

We also show a concrete version of Theorem 1.21 for the norm minimization problem.

Theorem 1.24 (Second order algorithm for norm minimization in terms of input size)**.** *Let* (π, v, ε) *be an instance of the scaling problem for* $\mathrm{SL}(n)$ *such that* $0 \in \Delta(v)$ *and every entry of* v *is bounded in absolute value by* M*. Let* d *denote the degree,* m *the dimension, and* γ *the weight margin of* π*. Then, Algorithm 5.1 applied to a suitably regularized objective function and a number of iterations at most*

$$T = O\left(\frac{d\sqrt{n}}{\gamma}\left(mn^3 d \log(Mmnd) + \log\frac{1}{\varepsilon}\right)\log\left(\frac{mnd \log M}{\varepsilon}\right)\right)$$

returns a group element $g \in \mathrm{SL}(n)$ *such that* $\log\|\pi(g)v\| \leqslant \log \mathrm{cap}(v) + \varepsilon$*. In particular, there is an algorithm to solve the norm minimization problem (Problem 1.10) for* $\mathrm{SL}(n)$ *in time* $\mathrm{poly}(\langle \pi \rangle, \langle v \rangle, \gamma^{-1}, \log(\varepsilon^{-1}))$*, which is at most* $\mathrm{poly}(\langle \pi \rangle, \langle v \rangle^n, \log(\varepsilon^{-1}))$*.*

Corollary 1.18 implies that both the first and second order algorithm result in a null cone membership algorithm with polynomial dependence on γ^{-1}; the tradeoffs are discussed in Remark 4.4 in Section 4.

Corollary 1.25 (Algorithm for null cone membership problem in terms of input size)**.** *There is an algorithm to solve the null cone membership problem (Problem 1.8) for* $\mathrm{SL}(n)$ *in time* $\mathrm{poly}(\langle \pi \rangle, \langle v \rangle, \gamma^{-1})$*, which is at most* $\mathrm{poly}(\langle \pi \rangle, \langle v \rangle^n)$*.*

In the important setting when the group is fixed (i.e., n is constant), the above corollary asserts that our second order algorithm solves the null cone problem for $\mathrm{SL}(n)$ in deterministic polynomial time. Prior to this result, the only known polynomial time algorithms for this class of null cone problems were given by the use of quantifier elimination (which is impractical) and, more recently, by Mulmuley in [Mul17, Theorem 8.5] through a purely algebraic approach. Mulmuley constructs a circuit which encodes a generating set of invariants for the ring of invariants of the corresponding action, and then invokes previous results on polynomial identity testing to obtain an algorithm for the null cone problem.

Finally, we give a randomized algorithm for the p-scaling problem based on Theorem 1.20. Here it is natural to consider the full group $\mathrm{GL}(n)$ rather than $\mathrm{SL}(n)$ as in the scaling problem.

Theorem 1.26 (First-order randomized algorithm for p-scaling in terms of input size)**.** *Let (π, v, p, ε) be an instance of the moment polytope problem for $\mathrm{GL}(n)$ such that $p \in \Delta(v)$ and every entry of v is bounded in absolute value by M. Let d denote the degree and m the dimension of π. Then, with probability at least $1/2$, Algorithm 7.1 with a number of iterations at most*

$$T = O\left(\frac{d^3}{\varepsilon^2} mn^5 \log(Mmnd)\right).$$

returns a group element $g \in \mathrm{GL}(n)$ such that $\|\mathrm{spec}(\mu(\pi(g)v) - p\|_2 \leqslant \varepsilon$. In particular, there is a randomized algorithm to solve the p-scaling problem (Problem 1.12) for $\mathrm{GL}(n)$ in time $\mathrm{poly}(\langle\pi\rangle, \langle v\rangle, \langle p\rangle, \varepsilon^{-1})$ and using $\mathrm{poly}(\langle\pi\rangle, \langle v\rangle)$ bits of randomness.

Similarly as for the null cone problem, our first order algorithm for p-scaling also implies a randomized algorithm for moment polytope membership.

Corollary 1.27 (Randomized algorithm for moment polytope membership in terms of input size)**.** *There is a randomized algorithm to solve the moment polytope membership problem (Problem 1.11) for $\mathrm{GL}(n)$ in time $\mathrm{poly}(\langle\pi\rangle, \langle v\rangle^n, 2^{n\langle p\rangle})$ and using $\mathrm{poly}(\langle\pi\rangle, \langle v\rangle)$ bits of randomness.*

We remark that our algorithms readily extend to representations of $G = \mathrm{GL}(n_1) \times \cdots \times \mathrm{GL}(n_k)$ and $G = \mathrm{SL}(n_1) \times \cdots \times \mathrm{SL}(n_k)$ at the expense of running times polynomial in $n_1 \cdots n_k$.

1.7 Organization of the paper

In Section 2, we discuss preliminaries from group and representation theory. In Section 3, we present our main structural results about the geometry of non-commutative optimization including smoothness, robustness, non-commutative duality, and gradient flow. Sections 4 and 5 contain the description and analysis of our first and second order algorithms, respectively, for null cone and moment polytope membership and capacity computation. Section 6 contains useful bounds on weight norms and weight margins. In Section 7, we design concrete algorithms and time complexity bounds for representations of $\mathrm{SL}(n)$ and $\mathrm{GL}(n)$ based on a priori lower bounds on the capacity. We conclude in Section 8 with a discussion of intriguing open problems. In Appendix A we supply the proof of a technical lemma for lifting coefficient bounds.

2 Preliminaries

In this section, we fix our basic notation and conventions and explain our basic group and representation theoretic setup. Throughout this article, we will be working with representations of continuous groups on finite-dimensional vector spaces. To make our article more accessible, we spell out explicitly all definitions in the important case when $G = \mathrm{GL}(n)$ (see Table 2.1 below). Thus, Sections 2.2 and 2.3 can be skipped on a first reading.

2.1 Notation and conventions

Throughout the paper, log denotes the natural logarithm. We abbreviate $[m] := \{1, \ldots, m\}$ for $m \in \mathbb{N}$.

All vector spaces are assumed to be finite-dimensional. If V is a complex vector space, let $L(V)$ denote the space of linear maps from V to V, and $\mathrm{GL}(V) \subseteq L(V)$ the group of invertible linear maps from V to V. The identity operator in $L(V)$ is denoted by I. Now assume that V is equipped with a Hermitian inner product $\langle\cdot,\cdot\rangle$ (by convention linear in the second argument). We caution that even when $V = \mathbb{C}^m$ this need *not* be the standard inner product. Then $A^\dagger$ denotes the adjoint of an operator $A \in L(V)$. Moreover, $U(V) \subset L(V)$ denotes the group of unitary operators (i.e., $U^\dagger U = UU^\dagger = I$), $\mathrm{Herm}(V) \subseteq L(V)$ the space of Hermitian operators (i.e., $A = A^\dagger$), and $\mathrm{PD}(V) \subseteq L(V)$ the set of positive definite operators in $L(V)$. Given an operator $X \in L(V)$, we write $\|X\|_F := (\operatorname{tr} X^\dagger X)^{1/2}$ for the Frobenius norm and $\|X\|_{\mathrm{op}} := \max_{\|v\|=1} \|Xv\|$ for the operator norm.

In this paper we work with matrix subgroups of $\mathrm{GL}(n)$, so it is useful to make distinct notation for each of the notions in the previous paragaph for $\mathbb{C}^n$ with its standard inner product. For $v, w \in \mathbb{C}^n$ we define $v \cdot w := \sum_{i=1}^n \overline{v_i} w_i$ and $\|v\|_2 := (v \cdot v)^{1/2} = (\sum_{i=1}^n |v_i|^2)^{1/2}$. Let $\mathrm{Mat}(n) \cong L(\mathbb{C}^n)$ denote the set of complex $m \times m$ matrices, and denote by $\mathrm{Herm}(n) \subseteq \mathrm{Mat}(n)$ the Hermitian matrices, by $\mathrm{PD}(n) \subseteq \mathrm{Mat}(n)$ the set of positive definite Hermitian matrices, by $\mathrm{GL}(n) \subseteq \mathrm{Mat}(n)$ the *general linear group* consisting of the invertible matrices, and by $U(n)$ the *unitary group* consisting of unitary matrices. For $A \in \mathrm{Mat}(n)$, $A^\dagger$ denotes the conjugate transpose of A, and we also use I for the identity matrix in $\mathrm{Mat}(n)$. We also write $\mathrm{Mat}(m, R)$ for the $m \times m$-matrices over a commutative ring R (e.g., the integers or Gaussian integers).

2.2 Groups

We now define the groups that our algorithms deal with in full generality and explain some of their main structural properties required in the analysis of the algorithms.

A subgroup G of $\mathrm{GL}(n)$ is called *symmetric*[4] if it is Zariski-closed and it holds that $g^\dagger \in G$ for every element $g \in G$. Here, Zariski-closed means that G is a subset of $\mathrm{GL}(n)$ defined by polynomial equations in the matrix entries $g_{i,j}$. For example, $\mathrm{SL}(n) = \{g \in \mathrm{GL}(n) : \det(g) = 1\}$. It can be shown that any complex reductive algebraic group is of this form [Wal17]. We will also demand that G is connected (in the standard topology, which is induced by any of the matrix norms defined above).

Next, define $K := G \cap U(n)$, the set of unitary matrices in G, which forms a maximal compact subgroup of G. Define $\mathrm{Lie}(K) := \{X \in \mathrm{Mat}(n) : e^{tX} \in K \forall t \in \mathbb{R}\}$, and likewise for G. Then $\mathrm{Lie}(K)$ is a real Lie algebra and $\mathrm{Lie}(G)$ is a complex Lie algebra. This means that they are a real and complex vector space, respectively, and that they are closed with respect to the Lie bracket $[X, Y] = XY - YX$. Moreover, $\mathrm{Lie}(K)$ is a subset of the *skew-Hermitian matrices* $i\,\mathrm{Herm}(n)$, so $i\,\mathrm{Lie}(K) \subseteq \mathrm{Herm}(n)$ and $\mathrm{Lie}(G) = \mathrm{Lie}(K) \oplus i\,\mathrm{Lie}(K)$, which means the Lie algebra of G is the *complexification* of that of K.

It is a crucial property of the functions we are optimizing that they are invariant under left multiplication by K. Hence, we will be interested in the Riemannian manifold $K\backslash G$, the set whose elements are cosets Kg of the subgroup $K \subset G$. We may identify $K\backslash G$ with the set $P := \exp(i\,\mathrm{Lie}(K))$ via the map $Kg \mapsto g^\dagger g$, because any element $g \in G$ has a unique *polar decomposition* $g = ke^H$ where $k \in K$ and $H \in i\,\mathrm{Lie}(K)$;[5][6] in particular, $G = KP$. The geodesics in $K\backslash G$ take the form $Ke^{tH}g$, which under our identification with P take the form $g^\dagger e^{2tH} g$; in particular these are geodesics with respect to a commonly studied Riemannian metric on $\mathrm{PD}(n)$ [FK94]. In Section 3.3 we will see that the optimization problem of interest satisfies convexity properties along such geodesics.

[4] Not to be confused with the *symmetric group* defined as the group of permutations of a finite set.

[5] Since $P = \exp(i\mathrm{Lie}(K)) \subseteq \exp(\mathrm{Herm}(n)) = \mathrm{PD}(n)$ and $K = G \cap U(n)$, the term polar decomposition is justified.

[6] In fact, the map $K \times i\,\mathrm{Lie}(K) \to G$, $(k, H) \mapsto ke^H$ is a diffeomorphism. Thus, $i\,\mathrm{Lie}(K)$ is diffeomorphic to $K\backslash G$ via $H \mapsto Ke^H$, which is naturally a *symmetric space* with non-positive sectional curvature (see, e.g., [Hel79, Woo10]).

Notation or concept	definition for $\mathrm{GL}(n)$
G	$\mathrm{GL}(n)$, invertible $n \times n$-matrices
$K \subseteq G$	$\mathrm{U}(n)$, unitary $n \times n$-matrices
$B \subseteq G$	upper triangular invertible $n \times n$-matrices
$N \subseteq B$	upper triangular $n \times n$-matrices with 1s on diagonal
P	$\mathrm{PD}(n)$, positive definite $n \times n$-matrices
T_G	$\mathrm{T}(n)$, diagonal invertible $n \times n$-matrices
T_K	diagonal unitary $n \times n$-matrices
$G = KP$	polar decomposition
$G = KB$	QR decomposition
$\mathrm{Lie}(G)$	$\mathrm{Mat}(n)$, complex $n \times n$-matrices
$\mathrm{Lie}(K)$	$\mathrm{Herm}(n)$, Hermitian $n \times n$-matrices
$\mathrm{Lie}(T_G)$	complex diagonal $n \times n$-matrices
$\mathrm{Lie}(T_K)$	purely imaginary diagonal $n \times n$-matrices
$i\,\mathrm{Lie}(T_K)$	real diagonal $n \times n$-matrices, usually identified with $\mathbb{R}^n$
$C(G) \subseteq i\,\mathrm{Lie}(T_K)$	$C(n) := \{p \in \mathbb{R}^n : p_1 \geqslant \cdots \geqslant p_n\}$
$s\colon i\,\mathrm{Lie}(T_K) \to C(G)$	spec, sorted eigenvalues of a Hermitian $n \times n$-matrix
p^* for $p \in C(G)$	$p^* = (-p_n, \ldots, -p_1)$
ω weight	$\omega \in \mathbb{Z}^n$, corresponding to irreducible representation of $\mathrm{T}(n)$ given by $\mathrm{diag}(x) \mapsto \prod_{j=1}^n x_j^{\omega_j}$
λ highest weight	$\lambda \in \mathbb{Z}^n \cap C(n) = \{\lambda \in \mathbb{Z}^n : \lambda_1 \geqslant \cdots \geqslant \lambda_n\}$
$\pi\colon G \to \mathrm{GL}(V)$	arbitrary representation
$\pi_\lambda\colon G \to \mathrm{GL}(V_\lambda)$	irreducible representation with highest weight λ

Table 2.1: Summary for $\mathrm{GL}(n)$

Let T_K be a maximal connected commutative subgroup of K. Then $T_G := \exp(\mathrm{Lie}(T_K) + i\,\mathrm{Lie}(T_K)) = T_K \exp(i\,\mathrm{Lie}(T_K))$ is a maximal connected commutative subgroup of G. We record the following generalization of the singular value decomposition, known as the *Cartan decomposition*:

$$G = KT_GK = K\exp(i\,\mathrm{Lie}(T_K))K \tag{2.1}$$

The group T_G is itself a symmetric subgroup of $\mathrm{GL}(n)$ and so the theory of this paper is also applicable to T_G. In fact, T_G is always isomorphic to the group of diagonal $r \times r$ matrices in $\mathrm{GL}(r)$ for some integer r, called the rank of G. As discussed in the introduction, this commutative case corresponds to geometric programming and enjoys a global convexity property that is simpler than the non-commutative case. At the same time, the subgroup T_G plays an important role in the representation theory of G, as we explain below.

Finally, let $B \subseteq G$ be a *Borel subgroup*, i.e., a maximal connected solvable subgroup, that contains T_G. Here, solvable means that if we inductively define $B^{(k)} := \{ghg^{-1}h^{-1} : g, h \in B^{(k-1)}\}$, with $B^{(0)} := B$, then we eventually reach the trivial subgroup, i.e., $B^{(k)} = \{I\}$ for some k. Moreover, define $N = \{b \in B : (b - I)^n = 0\}$. Then we have the following generalization of the QR decomposition, known as the *Iwasawa decomposition*:

$$G = KB = KT_G N = K \exp(i \operatorname{Lie}(T_K))N \tag{2.2}$$

Indeed, if $G = GL(n)$ then B consists of the upper-triangular invertible $n \times n$-matrices and N consists of the upper-triangular matrices with all ones on the diagonal. The final decomposition in (2.2) amounts to writing an invertible matrix as a product of a unitary matrix, a diagonal matrix with positive diagonal entries, and an upper triangular matrix with all ones on the diagonal. As a consequence: Any element $H \in \operatorname{Lie}(K)$ can be decomposed as $H = D + R + R^\dagger$, where $D \in \operatorname{Lie}(T_K)$ and $R \in \operatorname{Lie}(N)$, and this decomposition is orthogonal with respect to the Hilbert-Schmidt inner product (i.e., $\|H\|_F^2 = \|D\|_F^2 + 2\|R\|_F^2$).

2.3 Representations

In this section, we briefly discuss the basics of representation theory. The point which is crucial for us is that every representation can be associated with a set of integer vectors and the properties of these vectors will govern the running time of our algorithms.

Let $G \subseteq GL(n)$ be a symmetric subgroup as defined in Section 2.2. Let $\pi\colon G \to GL(V)$ be a rational representation of G. That is, π is a group homomorphism, i.e., $\pi(gh) = \pi(g)\pi(h)$ for all $g, h \in G$, and in any basis of V the matrix entries of $\pi(g) \in GL(V)$ are polynomials in the matrix entries $g_{i,j}$ and in $\det(g)^{-1}$. There always exists a K-invariant inner product on V; we let $\langle\cdot,\cdot\rangle$ denote such an inner product. That is, $\langle\cdot,\cdot\rangle$ is an inner product such that $\pi(K) \subset U(V)$. Even though we will often work with $V = \mathbb{C}^m$, the inner product $\langle\cdot,\cdot\rangle$ need not be the standard inner product; for instance, if $\operatorname{Sym}^2(\mathbb{C}^2)$ is identified with $\mathbb{C}^3$ by the monomial basis, then the standard inner product is not invariant under the action of $U(2)$ on $\operatorname{Sym}^2(\mathbb{C}^2)$.

We now define a number of objects associated to the representation π. Consider the complex-linear map $\Pi\colon \operatorname{Lie}(G) \to L(V) = \operatorname{Lie}(GL(V))$ given by

$$\Pi(H) := \partial_{t=0}\pi\big(e^{tH}\big). \tag{2.3}$$

This is the *Lie algebra representation* corresponding to π; in particular, the identity $\Pi([X,Y]) = [\Pi(X),\Pi(Y)]$ holds for all $X, Y \in \operatorname{Lie}(G)$. It holds that $e^{\Pi(X)} = \pi(e^X)$ for every $X \in \operatorname{Lie}(G)$. Furthermore, $\Pi(\operatorname{Lie}(K)) \subseteq i\operatorname{Herm}(V)$, so $\Pi(i\operatorname{Lie}(K)) \subseteq \operatorname{Herm}(V)$, and $\Pi(X^\dagger) = \Pi(X)^\dagger$ for every $X \in \operatorname{Lie}(G)$. A representation is called *trivial* if $\pi(g) = I$ for every $g \in G$.

A representation is called *irreducible* if it contains no invariant subspace other than $\{0\}$ and V itself, i.e., there exists no subspace $\{0\} \subsetneq W \subsetneq V$ such that $\pi(G)W \subseteq W$. Any representation of G can be decomposed into a direct sum of irreducible representations. This means that there exist irreducible representations $\pi_k\colon G \to GL(m_k)$, $\sum_k m_k = m$, and a unitary $u \in U(V)$, such that $u^\dagger\pi(g)u = \bigoplus_k \pi_k(g)$ for all $g \in G$. That is, up to a base change, the representation π can be decomposed into diagonal blocks, each of which corresponds to an irreducible representation.

If we restrict the representation to the commutative subgroup T_G then this decomposition is particularly simple, since it amounts to a joint diagonalization of the pairwise commuting operators $\{\pi(h) : h \in T_G\}$ or $\{\Pi(H) : H \in \operatorname{Lie}(T_G)\}$. Thus, there exists a decomposition $V = \bigoplus_{\omega\in\Omega(\pi)} V_\omega$, labeled by a set $\Omega(\pi) \subseteq i\operatorname{Lie}(T_K)$, such that

$$\pi(e^H)v_\omega = e^{\operatorname{tr}[H\omega]}v_\omega \qquad \text{and} \qquad \Pi(H)v_\omega = \operatorname{tr}[H\omega]v_\omega \tag{2.4}$$

for all $v_\omega \in V_\omega$ and $H \in \mathrm{Lie}(T_G)$. The vectors ω are called *weights*, the spaces V_ω are called *weight spaces*, and its elements v_ω are called *weight vectors*. Note that each $\mathbb{C}v_\omega$ is a one-dimensional irreducible representation of T_G. Since the exponential function is $2\pi i$-periodic, it is not hard to see that the set of possible weights forms a lattice isomorphic to $\mathbb{Z}^r$, where r is the rank as defined above. We note that since we take the weights to be elements of $i\,\mathrm{Lie}(T_K) \subseteq \mathrm{Herm}(n)$, it makes sense to take their Frobenius norm $\|\omega\|_F$ etc.

Returning to G, its irreducible representations can be labeled by a subset of the weights of T_G, called the *highest weights*. We denote the irreducible representation with highest weight λ by $\pi_\lambda \colon G \to \mathrm{GL}(V_\lambda)$. The space V_λ contains a one-dimensional invariant subspace for the Borel subgroup B, spanned by a (unique up to phase) unit vector v_λ. The vector v_λ is called a highest weight vector. It is a weight vector of weight λ and N-invariant. The latter means that

$$\pi(b)v_\lambda = v_\lambda \qquad \text{and} \qquad \pi(R)v_\lambda = 0 \tag{2.5}$$

for all $b \in N$ and $R \in \mathrm{Lie}(N)$. In general, the highest weights of the irreducible representations that appear in π form a subset of the set of weights $\Omega(\pi)$. The set of all possible highest weights spans a convex cone known as the *positive Weyl chamber*, denoted $C(G) \subseteq i\,\mathrm{Lie}(T_K)$. There is an involution $p \mapsto p^*$ on $C(G)$ such that, for every highest weight λ, λ^* is the highest weight of the dual representation, i.e., $\pi_{\lambda^*} \cong \pi_\lambda^*$. For any $H \in i\,\mathrm{Lie}(K)$, the intersection $\{kHk^\dagger : k \in K\} \cap C(P)$ is a single point, which we denote by $s(H)$. The function s generalizes the function spec taking a matrix to its spectrum, only with one technical difference: the image of s is a matrix in $C(P)$ rather than $\mathbb{R}^n$ as for spec; for instance, if $G = \mathrm{GL}(n)$, then $s(H) = \mathrm{diag}(\mathrm{spec}(H))$. We often use the identity $s(-p^*) = p$.

3 Geometry of non-commutative optimization

In this section, we first define the main optimization problem of interest, which is a norm minimization problem over group orbits. We then discuss the geometric properties of the objective function. While it is well-known that this function is in some sense log-convex, we will prove stronger convexity properties that will be instrumental for the algorithms discussed in the sequel. Throughout the paper, we work in the setup introduced in Section 2, with $\pi \colon G \to \mathrm{GL}(V)$ a representation of a symmetric subgroup $G \subseteq \mathrm{GL}(n)$ and $\langle \cdot, \cdot \rangle$ a K-invariant inner product on V.

3.1 Capacity and moment map

The norm minimization problem is concerned with solving the following optimization problem, the value of which we call the capacity.

Definition 3.1 (Capacity). *The* capacity *of a vector $v \in V$ is defined as the infimum of the norm on its G-orbit. Formally,*

$$\mathrm{cap}(v) := \inf_{g \in G} \|\pi(g)v\| = \min_{w \in \overline{\pi(G)v}} \|w\|.$$

In the second formula, the closure can be taken with respect to the standard topology (i.e., the one defined by the norm). The capacity is manifestly G-invariant, i.e., $\mathrm{cap}(\pi(g)v) = \mathrm{cap}(v)$ for all $v \in V$ and $g \in G$.

In geometric invariant theory, vectors are called *unstable* if $\operatorname{cap}(v) = 0$ and otherwise *semistable*. The set of unstable vectors forms the so-called *null cone* (this is a cone in the sense of algebraic geometry, namely closed under multiplication by arbitrary scalars). These are important in geometric invariant theory, defined by Mumford [Mum65], and going back to ideas introduced by Hilbert in his work on invariant theory [Hil93].

We are interested in the log-convexity properties of the objective function, so we define, for $0 \neq v \in V$, the function, also known as the *Kempf-Ness function*,

$$F_v \colon G \to [0, \infty), \quad g \mapsto \log\|\pi(g)v\| = \frac{1}{2}\log\|\pi(g)v\|^2. \tag{3.1}$$

It is useful to observe that this function is right-G-equivariant and left-K-invariant in the sense that

$$F_v(kgh) = \log\|\pi(kgh)v\| = \log\|\pi(g)\pi(h)v\| = F_{\pi(h)v}(g), \tag{3.2}$$

for all $k \in K$, $g, h \in G$, and $0 \neq v \in V$. Here we used that $\pi(K) \subseteq U(V)$, so group elements in K do not change the norm. The equivariance property shows that it suffices to study the local properties of F_v in a neighborhood of the identity element $g = I$. The invariance property implies that we can also think of F_v as a function on right cosets $K\backslash G$ or, equivalently, on $P = \exp(i \operatorname{Lie}(K))$ (recall that P is a subset of the set of Hermitian matrices).

The following definition captures the gradient of F_v at $g = I$:

Definition 3.2 (Moment map)**.** *The* moment map *is the function* $\mu : V \setminus \{0\} \to i \operatorname{Lie}(K)$ *defined by the property that, for all* $H \in i \operatorname{Lie}(K)$,

$$\operatorname{tr}\big[\mu(v)H\big] = \partial_{t=0} F_v(e^{tH}) = \frac{\langle v, \Pi(H)v\rangle}{\|v\|^2}.$$

Here, we recall that Π is the Lie algebra representation defined in Eq. (2.3). Since it is linear in H, the function $\mu(v)$ is well-defined. It is also a moment map in the sense of symplectic geometry (for the K-action on the projective space over V), which will have import implications in Section 3.6. We note that $\mu(\lambda v) = \mu(v)$ for $\lambda \in \mathbb{C}^* = \mathbb{C} \setminus \{0\}$.

Remark 3.3. *In the literature, the moment map is often defined as a function to the dual of* $\operatorname{Lie}(K)$ *or of* $i \operatorname{Lie}(K)$*. For us it is convenient to identify the dual with* $i \operatorname{Lie}(K)$ *so that we can think of the moment map concretely as computing gradient vectors rather than derivatives, which are naturally covectors.*

If G is commutative (i.e., $G = T_G$ and $K = T_K$) then one can write down a more concrete formula for the moment map. Write $v = \sum_{\omega \in \Omega(\pi)} v_\omega$, with v_ω contained in the weight space V_ω (cf. Section 2.3). Since weight spaces are pairwise orthogonal, it follows that the moment map is given by the convex combination

$$\mu(v) = \sum_{\omega \in \Omega(\pi)} \frac{\|v_\omega\|^2}{\|v\|^2} \omega, \tag{3.3}$$

which is a point in the convex hull of the points ω with $v_\omega \neq 0$ (the 'support' of the vector v).

3.2 Geodesic convexity

The group G is not a Euclidean space, but rather a manifold with an interesting topology and curvature. Therefore, the usual notions of convexity do not apply. However, it is well-known that the Kempf-Ness function is convex along certain curves, which have the interpretation of geodesics (e.g., [Wal17, Woo10]). We now show how to appropriately generalize definitions of convex optimization to this scenario. Next, we prove some quantitative results that have not been discussed in the literature but which will be crucial to our algorithms.

Definition 3.4 (Good geodesic)**.** *A* good geodesic *is a curve* $\gamma \colon \mathbb{R} \to G$ *of the form* $\gamma(t) = e^{tH} g$ *where* $H \in i\,\mathrm{Lie}(K)$ *and* $g \in G$. *We say that* γ *has* unit speed *if* $\|H\|_F = 1$.

Such curves are indeed geodesics in G with respect to a natural right-invariant metric. While not all geodesics in G are of this form, the induced curves $\tilde{\gamma} \colon \mathbb{R} \to K\backslash G$ defined by $\tilde{\gamma}(t) := K\gamma(t)$ are general geodesics in $K\backslash G$. Likewise, $\hat{\gamma} \colon \mathbb{R} \to P$ defined by $\hat{\gamma}(t) := \gamma(t)^\dagger \gamma(t) = g^\dagger e^{2tH} g$ is a general geodesic in $P = \exp(i\,\mathrm{Lie}(K))$. Thus, for left-$K$-invariant functions, good geodesics provide the appropriate curves with respect to which we will define our generalized notions of convexity. (One could also develop the entire formalism based on the function $\langle v, \pi(p) v \rangle$ for $p \in P$, but this would lead to less natural formulations of the algorithms below.)

Definition 3.5 (Convex, smooth, robust)**.** *Let* $F \colon G \to \mathbb{R}$ *be a function that is left* K*-invariant in the sense that* $F(kg) = F(g)$ *for all* $k \in K$, $g \in G$. *Assume that* F *is sufficiently differentiable such that all the derivatives below exist. We say that*

- F *is* (geodesically) convex *if*

$$\partial_t^2 F(\gamma(t)) \geqslant 0$$

  for every good geodesic $\gamma(t) = e^{tH} g$ *and* $t \in \mathbb{R}$. *That is,* F *is convex along all good geodesics (as a function of* t*).*

- F *is* L-smooth *for some* $L > 0$ *if*

$$\left|\partial_t^2 F(\gamma(t))\right| \leqslant L \|H\|_F^2$$

  for every good geodesic $\gamma(t) = e^{tH} g$ *and* $t \in \mathbb{R}$. *That is, it is* L*-smooth along all good geodesics with unit speed (as a function of* t*).*

- F *is* R-robust *for some* $R > 0$ *if*

$$\left|\partial_t^3 F(\gamma(t))\right| \leqslant R \|H\|_F \, \partial_t^2 F(\gamma(t))$$

  for every good geodesic $\gamma(t) = e^{tH} g$ *and* $t \in \mathbb{R}$.

Any robust function is convex (the right-hand side contains the second derivative, not its absolute value).[7] Just like in the Euclidean case, smooth convex functions and robust functions have local models that are useful for optimization. To state these concisely, it is useful to introduce the following notions:

[7] We note that the notion of robustness is similar but different from the notion of self-concordance (which plays a crucial role in the analysis of Newton's method and interior point methods in the Euclidean world [NN94]) which requires that $\left|\partial_t^3 F(\gamma(t))\right| \leqslant R \left(\partial_t^2 F(\gamma(t))\right)^{3/2}$ for every good geodesic $\gamma(t) = e^{tH} g$ and $t \in \mathbb{R}$ (generalizing the Euclidean definition to the geodesic world). One difference is that self-concordance is scale invariant whereas robustness is not.

Definition 3.6 (Geodesic gradient and Hessian)**.** *Let* $F\colon G \to \mathbb{R}$ *be a function that is left* K*-invariant in the sense that* $F(kg) = F(g)$ *for all* $k \in K$, $g \in G$. *Assume that* F *is sufficiently differentiable such that all the derivatives below exist. The* geodesic gradient *at* $g \in G$ *is defined as the vector* $\nabla F(g) \in i\,\mathrm{Lie}(K)$ *defined by*

$$\mathrm{tr}\big[\nabla F(g)H\big] = \partial_{t=0} F(e^{tH}g) \tag{3.4}$$

for all $H \in i\,\mathrm{Lie}(K)$. *The* geodesic Hessian *at* $g \in G$ *is the symmetric tensor* $\nabla^2 F(g) \in \mathrm{Sym}^2(i\,\mathrm{Lie}(K))$ *given by*

$$\mathrm{tr}\big[\nabla^2 F(g)(H \otimes H)\big] = \partial^2_{t=0} F(e^{tH}g) \tag{3.5}$$

for all $H \in i\,\mathrm{Lie}(K)$.

In other words, $\nabla F(g)$ and $\nabla^2 F(g)$ are the gradient and Hessian of the function $f_g\colon i\,\mathrm{Lie}(K) \to \mathbb{R}$ defined by $f_g(H) := F(e^H g)$ at $H = 0$.

Smoothness implies that a function is universally upper-bounded by a quadratic expansion.

Lemma 3.7. *Let* $F\colon G \to \mathbb{R}$ *be a convex and* L*-smooth function as defined in Definition 3.5. Then,*

$$F(g) + \mathrm{tr}\big[\nabla F(g)H\big] \leqslant F(e^H g) \leqslant F(g) + \mathrm{tr}\big[\nabla F(g)H\big] + \frac{L}{2}\|H\|_F^2$$

Proof. Consider the function $f(t) := F(e^{tH}g)$. By Taylor's approximation and the mean value theorem, we know that

$$f(1) = f(0) + f'(0) + \frac{1}{2}f''(\zeta)$$

for some $\zeta \in [0,1]$. By Eq. (3.4), $f'(0) = \mathrm{tr}[\nabla F(g)H]$. Finally, convexity and L-smoothness mean that

$$0 \leqslant f''(t) \leqslant L\|H\|_F^2$$

for all $t \in \mathbb{R}$. Thus the claim follows. □

Similarly, robustness implies upper *and* lower bounds in terms of local quadratic expansions.

Lemma 3.8. *Let* $F\colon G \to \mathbb{R}$ *be an* R*-robust function as defined in Definition 3.5. Then,*

$$F(g) + \partial_{t=0}F(e^{tH}g) + \frac{1}{2e}\partial^2_{t=0}F(e^{tH}g) \leqslant F(e^H g) \leqslant F(g) + \partial_{t=0}F(e^{tH}g) + \frac{e}{2}\partial^2_{t=0}F(e^{tH}g)$$

for every $g \in G$ *and* $H \in i\,\mathrm{Lie}(K)$ *such that* $\|H\|_F \leqslant 1/R$.

Proof. Consider the function $f(t) := F(e^{tH}g)$. Since F is R-robust, it holds that $|f'''(t)| \leqslant R\|H\|_F f''(t)$. Then the claim follows from [AZGL+18, Proposition B.1], which asserts that if $f\colon \mathbb{R} \to \mathbb{R}$ satisfies $|f'''(t)| \leqslant \rho f''(t)$ for all $t \in \mathbb{R}$ then

$$f(0) + f'(0)t + \frac{1}{2e}f''(0)t^2 \leqslant f(t) \leqslant f(0) + f'(0)t + \frac{e}{2}f''(0)t^2$$

for all $|t| \leqslant \frac{1}{\rho}$. □

3.3 Smoothness and robustness of the log-norm function

We now return to the log-norm or Kempf-Ness function F_v defined in Eq. (3.1), the logarithm of the objective function that defines the capacity (Definition 3.1). Note that the moment map from Definition 3.2 is nothing but its geodesic gradient at $g = I$. More generally,

$$\mu(\pi(g)v) = \nabla F_v(g). \tag{3.6}$$

We will prove that the left-K-invariant function F_v is convex and prove bounds on its smoothness and robustness. This will, unsurprisingly, depend on the properties of the Lie algebra representation. In particular, we will see that it depends on the following norm:

Definition 3.9 (Weight norm)**.** *We define the* weight norm *of the representation* π *by*

$$N(\pi) := \max_{H \in i\,\mathrm{Lie}(K), \|H\|_F = 1} \|\Pi(H)\|_{op}.$$

That is, the weight norm is an induced norm of the Lie algebra representation Π *defined in Eq. (2.3), where we equip the Lie algebra with the Frobenius norm and the linear operators on* V *with the usual operator norm.*

The weight norm can be computed explicitly in terms of representation-theoretic data, which justifies the name. For this, we borrow the following result from [BCMW17, Proof of Lemma 14]:

Proposition 3.10. *The weight norm can be computed as*

$$N(\pi) = \max\{\|\omega\|_F : \omega \in \Omega(\pi)\} = \max\{\|\lambda\|_F : \pi_\lambda \subseteq \pi\}.$$

In the first formula, we maximize over all weights of the representation V *and in the second formula over all irreducible representations* π_λ *that appear in* π *(cf. Section 2.3).*

Next, we show that the moment map (i.e., the gradient of F_v) is universally bounded by the weight norm:

Lemma 3.11 (Bound on gradient)**.** *For every* $v \in V \setminus \{0\}$*, we have that* $\|\mu(v)\|_F \leqslant N(\pi)$*.*

Proof. Using Definition 3.2 with $H = \mu(v) \in i\,\mathrm{Lie}(K)$, we obtain

$$\|\mu(v)\|_F^2 = \mathrm{tr}[\mu(v)\mu(v)] = \frac{\langle v, \Pi(\mu(v))v\rangle}{\|v\|^2} \leqslant \|\Pi(\mu(v))\|_{op} \leqslant N(\pi)\|\mu(v)\|_F,$$

from which the claim follows. □

Now we are ready to prove the desired convexity properties.

Proposition 3.12 (Convexity and smoothness)**.** *For any* $v \in V \setminus \{0\}$*, the function* F_v *defined in Eq. (3.1) is convex and* $2N(\pi)^2$*-smooth.*

Proof. Consider a good geodesic $\gamma(t) = e^{tH}g$, so $H \in i\,\mathrm{Lie}(K)$ and $g \in G$. Define $\tilde{H} := \Pi(H)$, $w(t) := \pi(e^{tH})gv = e^{t\tilde{H}}gv$, and $f(t) := F_v(\gamma(t)) = \frac{1}{2}\log\|w(t)\|^2$. Further, define unit vectors $u(t) := \frac{w(t)}{\|w(t)\|}$. Then, $w'(t) = \tilde{H}w(t)$, and after a short calculation we obtain that $u'(t) = \big(\tilde{H} - \langle u(t), \tilde{H}u(t)\rangle\, I\big)u(t)$ and

$$f'(t) = \langle u(t), \tilde{H}u(t)\rangle\,,$$

$$f''(t) = 2\left(\|\tilde{H}u(t)\|^2 - \langle u(t), \tilde{H}u(t)\rangle^2\right). \tag{3.7}$$

By the Cauchy-Schwarz inequality, $f''(t) \geqslant 0$, which proves convexity. Moreover,

$$|f''(t)| = f''(t) \leqslant 2\|\tilde{H}u(t)\|^2 \leqslant 2\|\tilde{H}\|_{op}^2 \leqslant 2N(\pi)^2\|H\|_F^2,$$

where we used that the $u(t)$ are unit vectors and Definition 3.9. □

A simple corollary shows that F_v is universally upper-bounded by a quadratic expansion.

Corollary 3.13. *For any $v \in V \setminus \{0\}$, the function F_v defined in Eq. (3.1) satisfies*

$$F_v(g) + \operatorname{tr}\left[\mu\left(\pi(g)v\right)H\right] \leqslant F_v(e^H g) \leqslant F_v(g) + \operatorname{tr}\left[\mu\left(\pi(g)v\right)H\right] + N(\pi)^2\|H\|_F^2$$

for every $g \in G$ and $H \in i\,\mathrm{Lie}(K)$.

Proof. This follows from Lemma 3.7, Proposition 3.12, and Eq. (3.6). □

Proposition 3.14 (Robustness)**.** *For every $v \in V \setminus \{0\}$, the function F_v defined in Eq. (3.1) is $4N(\pi)$-robust.*

Proof. We continue the calculation in the proof of Proposition 3.12. On the one hand, we can rewrite Eq. (3.7) as

$$f''(t) = 2\left\langle u(t), \left(\tilde{H} - \langle u(t), \tilde{H}u(t)\rangle I\right)^2 u(t)\right\rangle = 2\left\|\left(\tilde{H} - \langle u(t), \tilde{H}u(t)\rangle I\right)u(t)\right\|^2.$$

On the other hand, we obtain by taking another derivative that

$$\begin{aligned} f'''(t) &= 4\langle u(t), \tilde{H}^3 u(t)\rangle - 12\langle u(t), \tilde{H}u(t)\rangle\langle u(t), \tilde{H}^2 u(t)\rangle + 8\langle u(t), \tilde{H}u(t)\rangle^3 \\ &= 4\left\langle u(t), \left(\tilde{H} - \langle u(t), \tilde{H}u(t)\rangle I\right)^3 u(t)\right\rangle. \end{aligned}$$

By the Cauchy-Schwarz inequality and the triangle inequality, we obtain

$$|f'''(t)| \leqslant 2\left\|\tilde{H} - \langle u(t), \tilde{H}u(t)\rangle I\right\|_{op} f''(t) \leqslant 4\|\tilde{H}\|_{op} f''(t) \leqslant 4N(\pi)\|H\|_F f''(t).$$

This proves the claim. □

Remark 3.15 (Cumulant generating functions, tightness of Proposition 3.14)**.** *The preceding two propositions can also be established by interpreting $f(t)$ as a cumulant generating function. Without loss of generality, assume that $\pi(g)v$ is a unit vector. Consider the spectral decomposition $\tilde{H} = \sum_{\omega\in\Omega}\omega P_\omega$, and define a random variable X by $\Pr(X = 2\omega) = \|P_\omega\pi(g)v\|^2$. Then, $f(t) = \frac{1}{2}\log E[e^{tX}]$ is half the cumulant generating function of X, so we can interpret the k^{th} derivative of $f(t)$ at $t = 0$ as half the k^{th} cumulant of X. For $k = 2, 3$ the k^{th} cumulant is nothing but the k^{th} central moment, which when re-expressed in terms of $\tilde{H}$ yields the claim. Likewise, the function $n(t)$ in the proof of Proposition 3.16 below has a pleasant interpretation in terms of a moment generating function. More inequalities between higher order derivatives can be obtained via this connection but it is not clear if they are useful.*

This discussion also implies that, for any given $\varepsilon > 0$, there exists a representation π of, e.g., $G = GL(1)$ that is not $(\frac{1}{2}N(\pi) - \varepsilon)$-robust. Showing this amounts to finding a distribution on $[-2\beta, 2\beta] \cap \mathbb{Z}$ whose second and third central moments differ by a factor of β, which is quite simple. Thus, Proposition 3.14 is tight up to a constant factor.

A calculation similar to Proposition 3.14 shows that the norm-square function is robust. We state this in the following proposition (which has similar corollaries as the above).

Proposition 3.16 (Robustness)**.** *For every $0 \neq v \in V$, the function $N_v(g) := \|\pi(g)v\|^2$ is left-K-invariant, convex, and $2N(\pi)$-robust.*

Proof. Since $\pi(K) \subseteq U(V)$, N_v is clearly left-K-invariant. To prove robustness, fix $H \in i\,\mathrm{Lie}(K)$ and $g \in G$ as before. Define $\tilde{H} := \Pi(H)$, $w(t) := e^{t\tilde{H}}gv$, and $n(t) := \|w(t)\|^2$. Then, $w'(t) = \tilde{H}w(t)$ and

$$\begin{aligned} n'(t) &= 2\left\langle w(t), \tilde{H}w(t)\right\rangle, \\ n''(t) &= 4\left\langle w(t), \tilde{H}^2 w(t)\right\rangle = 4\left\|\tilde{H}w(t)\right\|^2, \\ n'''(t) &= 8\left\langle w(t), \tilde{H}^3 w(t)\right\rangle = 8\left\langle \tilde{H}w(t), \tilde{H}\tilde{H}w(t)\right\rangle \end{aligned}$$

Thus, by the Cauchy-Schwarz inequality,

$$|n'''(t)| \leqslant 8\|\tilde{H}\|_{op}\|\tilde{H}w(t)\|^2 = 2\|\tilde{H}\|_{op}n''(t) \leqslant 2N(\pi)\|H\|_F n''(t).$$

We conclude that $N_v(g)$ is $2N(\pi)$-robust. □

3.4 Noncommutative duality theory

As discussed in the preceding section, the log-norm function is geodesically convex in the sense of Definition 3.5. In particular, it follows that critical points of the norm function are global minima, and it is not hard to see that, within each orbit closure, minima are unique up to the action of K. These are basic and important results of geometric invariant theory [Mum65, KN79]. For example, the well-known Kempf-Ness theorem [KN79] asserts that

$$\mathrm{cap}(v) = \inf_{g \in G}\|\pi(g)v\| = \min_{w \in \overline{\pi(G)v}}\|w\| > 0 \iff \inf_{g \in G}\|\mu(\pi(g)v)\|_F = \min_{w \in \overline{\pi(G)v}}\|\mu(w)\|_F = 0, \tag{3.8}$$

From the perspective of optimization theory, this means that we can think of computing the capacity (i.e., minimizing the norm in an orbit closure) and minimizing the moment map (i.e., mimizing the gradient of the log-norm) as two *dual problems* – a point of view that was initially taken in [BGO$^+$17, BFG$^+$18].

In the following, we will prove two results that show that the norm of a vector v is close to its minimum (the capacity) if and only if the moment map is small. From the perspective of optimization theory, they relate the primal gap and dual gap of the two optimization problems. This improves over non-commutative duality theory developed in [BGO$^+$17, BFG$^+$18] and systematizes and generalizes results for matrix and operator scaling [LSW98, GGOW16] to arbitrary group representations.

We first prove the most difficult part, which is to show that if the gradient is small then the norm of v is close to its minimum. The argument is inspired by the proof of the analogous statement in [LSW98] for matrix scaling. To state our quantitative bound, we need the following definition:

Definition 3.17 (Weight margin; precise statement of Definition 1.15)**.** *We define the* weight margin *of the representation π by*

$$\gamma(\pi) := \min\left\{d(0, \mathrm{conv}(\Gamma)) \;:\; \Gamma \subseteq \Omega(\pi),\, \mathrm{conv}(\Gamma) \not\ni 0\right\},$$

where $\Omega(\pi)$ denotes the set of weights of the representation π and $d(0, \operatorname{conv}(\Gamma)) := \min\{\|x\|_F : x \in \operatorname{conv}(\Gamma)\}$ the minimal distance from the convex hull of Γ to the origin.

Theorem 3.18 (Lower bound from Theorem 1.16)**.** *For all $v \in V \setminus \{0\}$,*

$$\frac{\operatorname{cap}(v)}{\|v\|} \geqslant \sqrt{1 - \frac{\|\mu(v)\|_F}{\gamma(\pi)}}.$$

where $\gamma(\pi)$ is the weight margin defined in Definition 3.17.

Proof. Both sides are invariant under rescaling, so we may assume that $\|v\| = 1$. Thus we want to prove that

$$\operatorname{cap}^2(v) \geqslant 1 - \frac{\|\mu(v)\|_F}{\gamma(\pi)}. \tag{3.9}$$

We first prove the result in the case that G is commutative. If we expand v into weight vectors, $v = \sum_{\omega \in \Omega(\pi)} v_\omega$, then $p_\omega := \|v_\omega\|^2$ is a probability distribution. The squared capacity and moment map can then be computed as (cf. Eqs. (2.4) and (3.3))

$$\operatorname{cap}^2(v) = \inf_{g \in T_G} \|\pi(g)v\|^2 = \inf_{H \in i \operatorname{Lie}(T_K)} \sum_{\omega \in \operatorname{supp}(p)} p_\omega e^{2 \operatorname{tr}[\omega H]}, \tag{3.10}$$

$$\mu(v) = \sum_{\omega \in \operatorname{supp}(p)} p_\omega \omega, \tag{3.11}$$

where $\operatorname{supp}(p) := \{\omega : p_\omega > 0\}$. If $0 \notin \operatorname{conv}(\operatorname{supp}(p))$ then $\|\mu(v)\|_F \geqslant \gamma(\pi)$ by definition of the weight margin (Definition 3.17), so Eq. (3.9) holds because the capacity is nonnegative. Now assume that $0 \in \operatorname{conv}(\operatorname{supp}(p))$. We claim and will prove below that in this case there exist probability distributions p' and p'' on $\operatorname{supp}(p)$, as well as $\lambda \in [0, 1]$, such that

$$\sum_{\omega \in \operatorname{supp}(p)} p'_\omega \omega = 0, \quad p = (1-\lambda)p' + \lambda p'', \quad \text{and if } \lambda > 0 \text{ then } 0 \notin \operatorname{conv}(\operatorname{supp}(p'')). \tag{3.12}$$

Once we have such a distribution, observe from Eq. (3.11) that $\mu(v) = \lambda \sum_{\omega \in \operatorname{supp}(p'')} p''_\omega \omega$, which implies that

$$\lambda \leqslant \frac{\|\mu(v)\|_F}{\gamma(\pi)}$$

by definition of the weight margin (for $\lambda = 0$, this inequality holds trivially). Next, Jensen's inequality applied to the convex function $f(\omega) := e^{2 \operatorname{tr}[\omega H]}$ shows that

$$\sum_\omega p'_\omega e^{2 \operatorname{tr}[\omega H]} = \sum_\omega p'_\omega f(\omega) \geqslant f\left(\sum_\omega p'_\omega \omega\right) = f(0) = 1$$

for any fixed $H \in i \operatorname{Lie}(T_K)$. Hence we can lower bound the formula for the capacity in Eq. (3.10) by

$$\operatorname{cap}^2(v) \geqslant \inf_{H \in i \operatorname{Lie}(T_K)} (1-\lambda) \sum_\omega p'_\omega e^{2 \operatorname{tr}[\omega H]} \geqslant 1 - \lambda \geqslant 1 - \frac{\|\mu(v)\|_F}{\gamma(\pi)},$$

which is precisely what we wanted to show, i.e., Eq. (3.9).

To conclude the proof in the commutative case, we still need to prove that a decomposition as in Eq. (3.12) always exists. We will show this by induction on the size of $\operatorname{supp}(p)$. If $|\operatorname{supp}(p)| = 1$, the statement is clear — we may take $\lambda = 0$ and $p' = p$. For $|\operatorname{supp}(p)| > 1$, let q be a probability distribution with $\operatorname{supp}(q) \subseteq \operatorname{supp}(p)$ and $\sum_\omega q_\omega \omega = 0$; such a q exists by assumption. Choose α to be the largest number such that $p - \alpha q$ is still a nonnegative vector (that is, $\alpha = \min_{\omega \in \operatorname{supp}(q)} \frac{p_\omega}{q_\omega}$). If $\alpha = 1$ then $p = q$ and we are done, since we may again take $\lambda = 0$ and $p' = p$. Otherwise, $\alpha < 1$, so we can write $p - \alpha q = (1-\alpha) r$, where r is a probability distribution with $|\operatorname{supp}(r)| < |\operatorname{supp}(p)|$. If $0 \notin \operatorname{conv}(\operatorname{supp}(r))$ then this yields a decomposition of the desired form with $\lambda = \alpha$, $p' = q$, and $p'' = r$. Finally, if $0 \in \operatorname{conv}(\operatorname{supp}(r))$, then by induction $r = (1-\beta) r' + \beta r''$, where $\sum_\omega r'_\omega \omega = 0$ and $0 \notin \operatorname{conv}(\operatorname{supp}(r''))$ if $\beta > 0$. Then:

$$p = \alpha q + (1-\alpha) r = \alpha q + (1-\alpha)(1-\beta) r' + (1-\alpha)\beta r'',$$

so we may take p' as the normalization of $\alpha q + (1-\alpha)(1-\beta) q''$ and p'' as r''. This concludes the proof that a decomposition as in Eq. (3.12) always exists and, thereby, the proof in the commutative case.

Finally, consider the case that G is a general group. In the following we will consider both G and its maximal torus T_G, so we denote the capacity and moment map over a group H by cap_H and μ_H, respectively. By the Cartan decomposition $G = K T_G K$ from Eq. (2.1), it follows that

$$\operatorname{cap}_G(v) = \inf_{k \in K} \inf_{t \in T} \|\pi(t)\pi(k)v\| = \inf_{k \in K} \operatorname{cap}_{T_G}(\pi(k)v).$$

On the other hand, for every $k \in K$,

$$\|\mu_G(v)\|_F = \|\mu_G(\pi(k)v)\|_F \geqslant \|\mu_{T_G}(\pi(k)v)\|_F.$$

The last identity holds because $\mu_{T_G}(\pi(k)v)$ is the orthogonal projection of $\mu_G(\pi(k)v)$ onto $i\operatorname{Lie}(T_K) \subseteq i\operatorname{Lie}(K)$.[8] Thus, by Eq. (3.9) for the commutative group T_G,

$$\operatorname{cap}^2_G(v) = \inf_{k \in K} \operatorname{cap}^2_{T_G}(\pi(k)v) \geqslant 1 - \sup_{k \in K} \frac{\|\mu_{T_G}(\pi(k)v)\|_F}{\gamma(\pi)} \geqslant 1 - \frac{\|\mu_G(v)\|_F}{\gamma(\pi)},$$

which proves the claim. □

The next result states that, conversely, if $\|v\|$ is close to its infimum then the gradient is small.

Theorem 3.19 (Upper bound from Theorem 1.16)**.** *For all $v \in V \setminus \{0\}$,*

$$\frac{\operatorname{cap}(v)^2}{\|v\|^2} \leqslant 1 - \frac{\|\mu(v)\|_F^2}{4N(\pi)^2}.$$

where $N(\pi)$ is the weight norm defined in Definition 3.9.

Proof. Consider the function $F_v(g) = \log\|\pi(g)v\|$ defined in Eq. (3.1). If we apply the right-hand inequality in Corollary 3.13 with $g = I$ (the identity element) and $H = -\frac{\mu(v)}{2N(\pi)^2}$, we obtain

$$F_v\left(e^H\right) - F_v(I) \leqslant -\operatorname{tr}\left[\mu(v)\frac{\mu(v)}{2N(\pi)^2}\right] + N(\pi)^2 \left\|\frac{\mu(v)}{2N(\pi)^2}\right\|_F^2 = -\frac{\|\mu(v)\|_F^2}{4N(\pi)^2}.$$

[8] This is because, by definition of the moment map, $\operatorname{tr}[\mu_G(\pi(k)v)H] = \operatorname{tr}[\mu_{T_G}(\pi(k)v)H]$ for all $H \in i\operatorname{Lie}(T_K)$.

Since $F_v\left(e^H\right) - F_v(I) \geqslant \log \operatorname{cap}(v) - \log\|v\|$, we get that

$$\frac{\operatorname{cap}(v)^2}{\|v\|^2} \leqslant e^{-\frac{\|\mu(v)\|_F^2}{2N(\pi)^2}} \leqslant 1 - \frac{\|\mu(v)\|_F^2}{4N(\pi)^2},$$

where the second inequality follows from the fact that $e^{-x} \leqslant 1 - x/2$ for all $x \in [0, 1]$ and $\|\mu(v)\|_F \leqslant N(\pi)$ (Lemma 3.11). □

Theorems 3.18 and 3.19 together establish Theorem 1.16 announced in the introduction. They strengthen the classical Kempf-Ness result, Eq. (3.8), which can be obtained as a direct consequence. Indeed, if $\operatorname{cap}(v) > 0$ then there exists a sequence $g_k \in G$ such that $\|\pi(g_k)v\| \to \operatorname{cap}(v)$, so $\mu(\pi(g_k)v) \to 0$ by Theorem 3.19. Conversely, if $g_k \in G$ is a sequence such that $\mu(\pi(g_k)v) \to 0$, then $\operatorname{cap}(v)/\|\pi(g_k)v\| > 0$ for k sufficiently large by Theorem 3.18, and so $\operatorname{cap}(v) > 0$. In both arguments we used that the capacity is G-invariant, i.e., $\operatorname{cap}(v) = \operatorname{cap}(\pi(g)v)$ for every $g \in G$.

Remark 3.20. *In the language of moment polytopes, $\gamma(\pi)$ can be interpreted as the minimal distance between the origin and any moment polytope that does not contain the origin, when the group action is restricted to the commutative subgroup T_G (cf. Section 3.6). In other words, the weight margin is the largest constant $C > 0$ with the following property: If $\|\mu(v)\|_F < C$ then v is not in the null cone for the T_G-action.*

One can define a similar measure in terms of the moment polytopes for the action of G, called gap constant *in [BFG+18]. It is an interesting question whether a bound as in Theorem 3.18 holds with this improved constant.*

3.5 Gradient flow

In view of the convexity properties of the log-norm, it is natural to minimize it by using gradient descent. Indeed, this is the perspective that gives rise to our first-order algorithm presented in Section 4. Since minimizing the log-norm function is dual to minimizing the moment map, it is natural to also study gradient flows for minimizing the moment map. Kirwan first observed the remarkable properties of the gradient flow for the norm square of the moment map in [Kir84a]. In the context of tensor scaling and quantum marginals, Kirwan's flow was first proposed as an algorithmic tool in [WDGC13, Wal14]. [KLLR18, AZGL+18] studied the gradient flow for the (unsquared) norm of the moment map in the context of operator scaling. We will now explain how this analysis can be carried out in complete generality, which also leads to straightforward proofs.

For this, it will be convenient to consider a differently normalized version of the moment map. Namely, define $\tilde{\mu} : V \to i\operatorname{Lie}(K)$ by $\tilde{\mu}(v) = \|v\|^2\,\mu(v)$. Then, by Definition 3.2,

$$\operatorname{tr}\left[\tilde{\mu}(v)H\right] = \langle v, \Pi(H)v\rangle \tag{3.13}$$

for all $v \in V$ and $H \in i\operatorname{Lie}(K)$. From this, we recognize that $\tilde{\mu}(v)$ is the gradient of $f_v(g) = \frac{1}{2}\|\pi(g)v\|^2$ in the same way that $\mu(v)$ is the gradient of the function F_v defined in Eq. (3.1). (Recall that the gradient $\nabla f(v)$ of a function $f\colon V \to \mathbb{R}$ at $v \in V$ is defined by $\operatorname{Re}\langle\nabla f(v), w\rangle = D_w f(v)$ for all $w \in V$, where $D_w f(v) = \partial_{s=0} f(v + sw)$ denotes the partial derivative in direction w.) We note that $\tilde{\mu}(\lambda v) = \|\lambda\|^2\mu(v)$ for $\lambda \in \mathbb{C}^*$.

Remark 3.21. *From the point of view of geometric invariant theory, $\tilde{\mu}$ is a moment map on the vector space V rather than on projective space $\mathbb{P}(V)$.*

Consider the gradient of the infinitely differentiable, real-valued function

$$v \mapsto \|\tilde{\mu}(v)\|_F$$

defined on the open subset $U := \{v \in V : v \neq 0, \mu(v) \neq 0\}$ of V. We now derive a concrete formula by a slight variation of [Kir84a, Lemma 6.6].

Lemma 3.22. *For $v \in U$ we have*

$$\nabla\|\tilde{\mu}\|_F(v) = 2\frac{\Pi(\tilde{\mu}(v))v}{\|\tilde{\mu}(v)\|_F} = 2\frac{\Pi(\mu(v))v}{\|\mu(v)\|_F}.$$

Proof. First we note that for $v \in U$,

$$\nabla\|\tilde{\mu}\|_F(v) = \frac{\nabla\|\tilde{\mu}\|_F^2(v)}{2\|\tilde{\mu}(v)\|_F}. \tag{3.14}$$

Next, we compute the right-hand side gradient by

$$\operatorname{Re}\langle\nabla\|\tilde{\mu}\|_F^2(v), w\rangle = D_w\|\tilde{\mu}\|_F^2(v) = D_w \operatorname{tr}\left[\tilde{\mu}(v)^2\right] = 2\operatorname{tr}\left[\tilde{\mu}(v)D_w\tilde{\mu}(v)\right]. \tag{3.15}$$

However, differentiating both sides of Eq. (3.13) shows that

$$\operatorname{tr}\left[HD_w\tilde{\mu}(v)\right] = D_w\langle v, \Pi(H)v\rangle = \langle w, \Pi(H)v\rangle + \langle v, \Pi(H)w\rangle = 2\operatorname{Re}\langle\Pi(H)v, w\rangle$$

for every $H \in i\operatorname{Lie}(K)$. In particular, this holds for $H = \tilde{\mu}(v)$. Plugging this into Eq. (3.15) yields

$$\operatorname{Re}\langle\nabla\|\tilde{\mu}\|_F^2(v), w\rangle = 4\operatorname{Re}\langle\Pi(\tilde{\mu}(v))v, w\rangle$$

for all $w \in V$, so that $\nabla\|\tilde{\mu}\|_F^2(v) = 4\Pi(\tilde{\mu}(v))v$. Now the claim follows from Eq. (3.14). □

Definition 3.23 (Gradient flow). *For $v \in U$ we consider in U the ordinary differential equation*

$$v'(t) = -\nabla\|\tilde{\mu}\big(v(t)\big)\|_F, \quad v(0) = v. \tag{3.16}$$

We denote by $[0, T_v) \to U$, $t \mapsto v(t)$ its unique solution on its maximal interval of definition, where $T_v \leqslant \infty$.

The existence and uniqueness of the solution follow from a standard ODE result (the Picard-Lindelöf theorem, see, e.g., Theorem 2.2 in [CL55]), since the vector field $v \mapsto \nabla\|\tilde{\mu}\big(v\big)\|_F$ is C^1 and hence locally Lipschitz continuous on U. As a consequence of Lemma 3.22, we can write the flow in Definition 3.23 as

$$v'(t) = -2\Pi\left(\frac{\mu(v(t))}{\|\mu(v(t))\|_F}\right)v(t) \in \Pi(\operatorname{Lie}(G))v(t). \tag{3.17}$$

It follows that the flow $v(t)$ actually stays in the orbit $\pi(G)v(0)$ at all times $t \in [0, T_v)$. We record this and additional useful properties of the flow in the following proposition:

Proposition 3.24 (Properties of the flow). *For $0 \leqslant t < T_v$ we have:*

1. $\partial_t\|\tilde{\mu}(v(t))\|_F = -\|v'(t)\|^2$.
2. $\partial_t\|v(t)\|^2 = -4\|\tilde{\mu}(v(t))\|_F$.

3. $\partial_t^2 \|v(t)\|^2 = 4\|v'(t)\|^2$.

4. *The ordinary differential equation in* G,

$$g'(t) = -2\frac{\mu(v(t))}{\|\mu(v(t))\|_F} g(t), \quad g(0) = I, \tag{3.18}$$

has a solution $g : [0, T_v) \to G$*, which satisfies* $v(t) = \pi(g(t))v$*. In particular,* $v(t) \in \pi(G)v$ *for* $t \in [0, T_v)$.

5. *Suppose* T_v *is finite. Then the limit* $v(T_v) := \lim_{t\uparrow T_v} v(t)$ *exists and it satisfies*

$$\|v(T_v)\| = \mathrm{cap}(v).$$

Proof. Item 1 is true for any gradient flow. To see Item 2, note that

$$\partial_t \|v(t)\|^2 = 2\,\langle v'(t), v(t)\rangle = -4\,\mathrm{Re}\,\frac{\langle \Pi(\mu(v(t)))v(t), v(t)\rangle}{\|\mu(v(t))\|_F} = -4\frac{\mathrm{tr}[\tilde{\mu}(v(t))\mu(v(t))]}{\|\mu(v(t))\|_F} = -4\|\tilde{\mu}(v(t))\|_F,$$

where the second equality is Eq. (3.17) and the third is Eq. (3.13). Item 3 follows from Item 1 and Item 2.

Item 4 follows because Eq. (3.18) is a linear ODE in the entries of $g(t)$ with continuous coefficients and hence has a solution on $[0, T_v)$. Observe that $\pi(g(t))v$ also solves Eq. (3.17), and hence $v(t) = \pi(g(t))v$ by the uniqueness of $v(t)$.

Finally, for showing Item 5, we assume $T_v < \infty$. By Item 2, $t \mapsto \|v(t)\|$ is monotonically decreasing, hence $\|v(t)\| \leqslant \|v\|$ for all $0 \leqslant t < T_v$ and the solution is bounded. In the situation of a finite time and bounded solution, a standard ODE argument (e.g., see [CL55, Theorem 4.1]) tells us that the limit $v(T_v) := \lim_{t\uparrow T_v} v(t)$ exists, but it does not lie in the domain of definition U.

Observe that $\lim_{t\uparrow T_v}\|v(t)\| = \|v(T_v)\|$. If $\|v(T_v)\| > 0$, then because $v(T_v)$ is outside the domain of definition of U we must have $\mu(v(T_v)) = 0$. In particular, $\lim_{t\uparrow T_v} \mu(v(t)) = 0$, which implies $\|v(T_v)\| = \lim_{t\uparrow T_v}\|v(t)\| = \mathrm{cap}(v)$ by Theorem 1.16. On the other hand, if $v(T_v) = 0$, then $\lim_{t\uparrow T_v}\|v(t)\| = 0$ and hence $\mathrm{cap}(v) = 0$[9] because $v(t) \in \pi(G)v$ for $t < T_v$ by Item 4. This proves Item 5. □

Next, we show that the flow converges quickly to an approximate minimizer of the norm-square function.

Theorem 3.25 (Convergence of gradient flow)**.** *Let* $v(t)$ *denote the gradient flow from Definition 3.23 with initial vector* $v = v(0)$*. For every* $\varepsilon > 0$*, there is some* $T \leqslant \frac{1}{4\gamma(\pi)}\log(\|v\|^2/\varepsilon)$ *such that* $T < T_v$ *and for every* $T \leqslant t < T_v$ *we have*

$$\|v(t)\|^2 \leqslant \mathrm{cap}^2(v) + \varepsilon,$$

where $\gamma(\pi)$ *is the weight margin defined in Definition 3.17.*

[9] Though we do not need it here, we will see later in the proof of Proposition 5.5 that T_v is, in fact, never finite if $\mathrm{cap}(v) = 0$.

Proof. By the second fact in Proposition 3.24, we have for $0 \leqslant t < T_v$

$$\partial_t \left(\|v(t)\|^2 - \operatorname{cap}^2(v)\right) = -4\|\tilde{\mu}(v(t))\|_F.$$

Moreover, by Theorem 3.18, we have $\operatorname{cap}^2(v(t)) \geqslant \|v(t)\|^2 - \|\tilde{\mu}(v(t))\|_F/\gamma(\pi)$, hence

$$\partial_t \left(\|v(t)\|^2 - \operatorname{cap}^2(v)\right) \leqslant -4\gamma(\pi)\left(\|v(t)\|^2 - \operatorname{cap}^2(v(t))\right) \leqslant -4\gamma(\pi)\left(\|v\|^2 - \operatorname{cap}^2(v)\right),$$

where we used $\|v(t)\| \leqslant \|v\|$ and that the capacity is G-invariant. Hence we obtain for $0 \leqslant t < T_v$ that

$$\|v(t)\|^2 - \operatorname{cap}^2(v) \leqslant e^{-4\gamma(\pi)t}\left(\|v\|^2 - \operatorname{cap}^2(v)\right) \leqslant e^{-4\gamma(\pi)t}\|v\|^2.$$

If $\frac{1}{4\gamma(\pi)}\log(\|v\|^2/\varepsilon) < T_v$, the assertion follows by taking $T := \frac{1}{4\gamma(\pi)}\log(\|v\|^2/\varepsilon)$. Otherwise, T_v is finite and we have $\|v(T_v)\| = \operatorname{cap}(v)$ by Item 5. In this case, any $T < T_v$ sufficiently close to T_v will do. □

The next corollary gives an analogous bound for the log-norm. We will use it in Section 5 to derive a diameter bound for our second order algorithm.

Corollary 3.26. *Let $v(t)$ denote the gradient flow from Definition 3.23 with initial vector $v = v(0)$ and assume $\operatorname{cap}(v) > 0$. For every $\varepsilon > 0$, there is some $T \leqslant \frac{1}{4\gamma(\pi)}\log\left(\|v\|^2/(2\operatorname{cap}^2(v)\varepsilon)\right)$ such that $T < T_v$ and for every $T \leqslant t < T_v$, we have*

$$\log\|v(t)\| \leqslant \log\operatorname{cap}(v) + \varepsilon.$$

Proof. By Theorem 3.25 and our choice of T, for $T \leqslant t < T_v$ we have

$$\|v(t)\|^2 \leqslant \operatorname{cap}^2(v) + 2\operatorname{cap}^2(v)\varepsilon = (1+2\varepsilon)\operatorname{cap}^2(v),$$

so the claim follows by taking logarithms and using the estimate $\log(1+x) \leqslant x$. □

3.6 Moment polytopes

In this section, we discuss the optimization problem underlying the moment polytope membership problem. We first explain the general definition of the moment polytope. For a nonzero vector $v \in V$, we define the *moment polytope of* v by

$$\Delta(v) := \overline{\{\mu(w) : w \in G\cdot v\}} \cap C(G) = \overline{\{s(\mu(w)) : w \in G\cdot v\}},$$

where $C(G)$ is the positive Weyl chamber defined in Section 2.3. The equality follows because the moment map is K-equivariant, which means that, $\mu(\pi(k)w) = k\mu(w)k^\dagger$ for all $w \in V$ and $k \in K$. If $G = GL(n)$ then $K = U(n)$, $C(G) = C(n)$, and $s = \operatorname{spec}$, so the moment polytope is precisely the set of all spectra (eigenvalues ordered non-increasingly) of moment map images obtained from scalings of v; this is the definition that we gave in Section 1.3.3. A point in $C(G)$ is called rational if an integer multiple of it is a (highest) weight. We remark that $\Delta(v)$ is a moment polytope in the sense of symplectic geometry of the orbit closure of v in the projective space $\mathbb{P}(V)$. It is a nontrivial fact that $\Delta(v)$ is a convex polytope with rational vertices [NM84, Bri87].

Now let $p \in C(G)$ be an arbitrary rational point and let $\ell > 0$ be an integer such that $\lambda := \ell p$ is a highest weight. In Section 1.5.3 we motivated the following p-*capacity*,

$$\operatorname{cap}_p(v) := \inf_{g \in G} \|(\pi(g)v)^{\otimes \ell} \otimes (\pi_{\lambda^*}(g)v_{\lambda^*})\|^{1/\ell}, \tag{3.19}$$

where λ^* denotes the highest weight of the dual representation as defined in Section 2.3. Clearly,

$$\operatorname{cap}_p(v) = \operatorname{cap}(v^{\otimes \ell} \otimes v_{\lambda^*})^{1/\ell}, \tag{3.20}$$

where the right-hand side capacity is computed using the representation $\rho \colon G \to GL(W)$ on the vector space $W = \operatorname{Sym}^{\ell}(V) \otimes V_{\lambda^*}$ defined by $\rho(g) = \pi(g)^{\otimes \ell} \otimes \pi_{\lambda^*}(g)$. The relevance of the p-capacity is due to the 'shifting trick' from [NM84, Bri87], which shows that $p \in \Delta(v)$ iff $0 \in \Delta(w)$ for some vector of the form $w = (\pi(h)v)^{\otimes \ell} \otimes v_{\lambda^*}$. Moreover, if the latter condition holds for some $h \in G$ then it holds for generic $h \in G$. Now, by the Kempf-Ness theorem, $0 \in \Delta(w)$ if and only if $\operatorname{cap}(w) > 0$, as explained in Section 1.3.4. Since $\operatorname{cap}(w)$ is nothing but the p-capacity of $\pi(h)v$, we obtain the following important equivalence:

$$p \in \Delta(v) \iff \operatorname{cap}_p(\pi(h)v) > 0 \text{ for some } h \in G \iff \operatorname{cap}_p(\pi(h)v) > 0 \text{ for generic } h \in G. \tag{3.21}$$

Thus, membership in the moment polytope can be reduced to p-capacities by a suitable randomization step ($v \mapsto \pi(h)v$ for random h). We will later state an effective version of this observation that shows, for the case $G = GL(n)$, how much randomness suffices (Theorem 7.15).

In the remainder of this section we will focus on the p-capacity. We first analyze the smoothness and robustness of the logarithm of the objective function underlying the p-capacity (3.19), i.e.,

$$F_{v,p} \colon G \to \mathbb{R}, \quad F_{v,p}(g) = \log \|\pi(g)v\| + \frac{1}{\ell} \log \|\pi_{\lambda^*}(g)v_{\lambda^*}\| \tag{3.22}$$

(this is nothing but $F_{v^{\otimes \ell} \otimes v_{\lambda^*}}$, the log-norm function of the vector $v^{\otimes \ell} \otimes v_{\lambda^*}$, divided by ℓ). Clearly, $F_{v,p}$ can be written as a linear combination of two log-norm functions (3.1):

$$F_{v,\lambda} = F_v + \frac{1}{\ell} F_{v^*_\lambda}. \tag{3.23}$$

By Proposition 3.12, F_v is convex and $2N(\pi)^2$-smooth, while $F_{v^*_\lambda}$ is $2\|\lambda\|_F^2$-smooth by Proposition 3.12. Thus, we find that the smoothness of $F_{v,p}$ can be upper bounded by $2N(\pi)^2 + 2\ell\|p\|_F^2$. While correct, this bound does not lead to efficient algorithms since ℓ depends exponentially on the bitsize of p. Fortunately, it is excessively pessimistic since it does not use the fact that v_{λ^*} is a highest weight vector. We will now derive a better bound that does not depend on ℓ:

Proposition 3.27. *The function* F_{v_λ} *is* $2\|\lambda\|_F$*-smooth. As a consequence, the function* $F_{v,p}$ *is* $2N^2$*-smooth, where* $N^2 := N(\pi)^2 + \|p\|_F$.

Proof. For $g \in G$ and $H \in i\operatorname{Lie}(K)$, consider the function

$$h(t) := F_{v_\lambda}(e^{tH}g) = \log\|e^{t\Pi_\lambda(H)}\pi_\lambda(g)v_\lambda\|.$$

We would like to show that

$$h''(t) \leqslant 2\|\lambda\|_F\|H\|_F^2 \tag{3.24}$$

for all t. It suffices to prove Eq. (3.24) for $t = 0$, since we can always replace g by $e^{tH}g$. We will now argue that we can also restrict to $g = I$. Using the Iwasawa decomposition to write $g = kb$ for some $k \in K$ and $b \in B$, we have $\pi_\lambda(g)v_\lambda = z\pi_\lambda(k)v_\lambda$ for some $z \in \mathbb{C}^*$, because v_λ is a highest weight vector. Thus:

$$\begin{aligned} h(t) &= \log\|e^{t\Pi_\lambda(H)}\pi_\lambda(g)v_\lambda\| = \log\|e^{t\Pi_\lambda(H)}\pi_\lambda(k)v_\lambda\| + \log|z| \\ &= \log\|\pi_\lambda(k^{-1})e^{t\Pi_\lambda(H)}\pi_\lambda(k)v_\lambda\| + \log|z| = \log\|e^{t\Pi_\lambda(k^{-1}Hk)}v_\lambda\| + \log|z|, \end{aligned}$$

where we used that the norm is K-invariant. The additive constant does not impact derivatives and $k^{-1}Hk \in i\,\mathrm{Lie}(K)$. We may thus assume that $g = I$. Then, Eq. (3.7) shows that

$$\frac{1}{2}h''(0) = \langle \Pi_\lambda(H)v_\lambda, \Pi_\lambda(H)v_\lambda\rangle - \langle v_\lambda, \Pi_\lambda(H)v_\lambda\rangle^2 .$$

Since $H \in i\,\mathrm{Lie}(K)$, we can decompose it as $H = D+R+R^\dagger$, where $D \in i\,\mathrm{Lie}(T_K)$ and $R \in \mathrm{Lie}(N)$. Then we know from Eq. (2.4) that $\Pi_\lambda(D)v_\lambda$ is a real scalar multiple of v_λ, from Eq. (2.5) that $\Pi_\lambda(R)v_\lambda = 0$, and that $\Pi_\lambda(R^\dagger) = \Pi_\lambda(R)^\dagger$. Using this, we can simplify as follows:

$$\begin{aligned} &\langle \Pi_\lambda(H)v_\lambda, \Pi_\lambda(H)v_\lambda\rangle - \langle v_\lambda, \Pi_\lambda(H)v_\lambda\rangle^2 = \langle \Pi_\lambda(R)^\dagger v_\lambda, \Pi_\lambda(R^\dagger)v_\lambda\rangle \\ &= \langle v_\lambda, \Pi_\lambda(R)\Pi_\lambda(R^\dagger)v_\lambda\rangle = \langle v_\lambda, [\Pi_\lambda(R), \Pi_\lambda(R^\dagger)]v_\lambda\rangle = \langle v_\lambda, \Pi_\lambda([R, R^\dagger])v_\lambda\rangle . \end{aligned}$$

In the second line we used once more that $\Pi_\lambda(R)v_\lambda = 0$ and that Π_λ is a Lie algebra representation. We obtain

$$\frac{1}{2}h''(0) \leqslant \|\Pi_\lambda([R, R^\dagger])\|_{op} \leqslant N(\pi_\lambda)\|[R, R^\dagger]\|_F \leqslant 2N(\pi_\lambda)\|R\|_F^2 \leqslant 2\|\lambda\|_F\|R\|_F^2 \leqslant \|\lambda\|_F\|H\|_F$$

by definition of the weight norm, submultiplicativity of the Frobenius norm, Proposition 3.10, and, finally, $2\|R\|_F^2 \leqslant \|H\|_F^2$, which holds since the decomposition $H = D + R + R^\dagger$ is orthogonal with respect to the Hilbert-Schmidt inner product. We have thus shown Eq. (3.24) for $t = 0$, concluding the proof. □

We now compute the geodesic gradient of the objective function (3.22). By Eqs. (3.6) and (3.23),

$$\nabla F_{v,p}(g) = \mu(\pi(g)v) + \frac{1}{\ell}\mu_{\lambda^*}(\pi_{\lambda^*}(g)v_{\lambda^*}),$$

where we write μ_λ^* for the moment map associated with the irreducible representation π_{λ^*}. The latter can computed readily. Write $g = kb$ according to the Iwasawa decomposition $G = KB$ from Eq. (2.2). Using that v_{λ^*} is a B-eigenvector and the K-equivariance of the moment map, we find that $\mu_{\lambda^*}(\pi_{\lambda^*}(g)v_{\lambda^*}) = k\lambda^*k^\dagger$. Thus we obtain the following formula for the gradient of $F_{v,p}$ at $g = kb$:

$$\nabla F_{v,p}(g) = \mu(\pi(g)v) + kp^*k^\dagger = \mu(\pi(g)v) - k(-p^*)k^\dagger. \tag{3.25}$$

Since $s(-kp^*k^\dagger) = s(-p^*) = p$, we note that the gradient vanishes if and only if $s(\mu(\pi(g)v)) = p$, i.e., $\pi(g)v$ maps to the desired point p in the moment polytope. We will use this formula in our first-order algorithm for non-uniform scaling (4.3). It also implies that $F_{v,p}$ is universally upper-bounded by the following quadratic expansion, generalizing Corollary 3.13.

Corollary 3.28. *For any $v \in V \setminus \{0\}$ and rational $p \in C(G)$, the function $F_{v,p}$ defined in Eq. (3.22) satisfies*

$$F_{v,p}(g) + \operatorname{tr}\big[(\mu(\pi(g)v) + p^*)\,H\big] \leqslant F_{v,p}(e^H g) \leqslant F_{v,p}(g) + \operatorname{tr}\big[(\mu(\pi(g)v) + p^*)\,H\big] + N^2\|H\|_F^2$$

for every $g \in G$ and $H \in i\operatorname{Lie}(K)$.

Proof. This follows from Lemma 3.7, Proposition 3.27, and Eq. (3.25). □

Next, we derive noncommutative duality results that generalize Theorems 3.18 and 3.19. Recall that any point p in the moment polytope necessarily satisfies $\|p\|_F \leqslant N(\pi)$ by Lemma 3.11. Thus the condition in the following two results is without loss of generality.

Theorem 3.29. *For any $v \in V \setminus \{0\}$ and rational $p \in C(G)$ with $\|p\|_F \leqslant N(\pi)$,*

$$\frac{\operatorname{cap}_p(v)^2}{\|v\|^2} \leqslant 1 - \frac{\|\mu(v) + p^*\|_F^2}{4N^2},$$

where $N^2 := N(\pi)^2 + \|p\|_F$.

Proof. This follows by adapting the proof of Theorem 3.19 to use Corollary 3.28 in place of Corollary 3.13. We apply the second inequality in Corollary 3.28 with $g = I$ and $H = -\frac{\mu(v)+p^*}{2N^2}$. Then,

$$F_{v,p}(e^H) - F_{v,p}(I) \leqslant \operatorname{tr}\big[(\mu(v) + p^*)\,H\big] + N^2\|H\|_F^2 = -\frac{\|\mu(v) + p^*\|_F^2}{4N^2}$$

and we can proceed as in the proof of Theorem 3.19 to see that

$$\frac{\operatorname{cap}_p(v)^2}{\|v\|^2} \leqslant e^{2\left(F_{v,p}(e^H) - F_{v,p}(I)\right)} \leqslant e^{-\frac{\|\mu(v)+p^*\|_F^2}{2N^2}} \leqslant 1 - \frac{\|\mu(v) + p^*\|_F^2}{4N^2}.$$

For the last inequality, use that $e^{-x} \leqslant 1 - x/2$ for all $x \in [0, 1]$. □

Theorem 3.30. *Let $v \in V \setminus \{0\}$ and let $p \in C(G)$ be rational with $\|p\|_F \leqslant N(\pi)$. Let $\ell > 0$ be an integer such that $\lambda := \ell p$ is a highest weight. Then,*

$$\frac{\operatorname{cap}_p(v)^\ell}{\|v\|^\ell} \geqslant \sqrt{1 - \frac{\ell\|\mu(v) + p^*\|_F}{\gamma(\rho)}},$$

where $\gamma(\rho)$ is the weight margin of the representation $\rho\colon G \to GL(W)$ on $W = \operatorname{Sym}^\ell(V) \otimes V_{\lambda^*}$. In particular, if $\|s(\mu(v)) - p\|_F < \gamma(\rho)/\ell$ then $p \in \Delta(v)$.

Proof. Use Eq. (3.20) to write $\operatorname{cap}_p(v)^\ell$ as the capacity of the vector $w = v^{\otimes \ell} \otimes v_{\lambda^*}$ with respect to the representation ρ. In view of (3.25), the corresponding moment map is given by $\mu(w) = \ell\mu(v) + \lambda^*$. Thus, the first claim is a consequence of Theorem 3.18. The second claim follows from the first, since $\|s(\mu(v)) - p\|_F < \gamma(\rho)/\ell$ means that there exists $k \in K$ such that $\|\mu(\pi(k)v) + p^*\|_F < \gamma(\rho)/\ell$. Then, $\operatorname{cap}_p(\pi(k)v) > 0$ and so $p \in \Delta(\pi(k)v) = \Delta(v)$ by Eq. (3.21). □

Theorem 3.30 shows that we can reduce the moment polytope membership problem to the p-scaling problem for some suitable choice of $\varepsilon > 0$, generalizing Corollary 1.18.

Corollary 3.31. *Let $v \in V \setminus \{0\}$ and let $p \in C(G)$ be rational with $\|p\|_F \leqslant N(\pi)$. Let $\ell > 0$ be an integer such that $\lambda := \ell p$ is a highest weight. Then, $p \in \Delta(v)$ if and only if $\Delta(v)$ contains a point of distance smaller than $\gamma(\rho)/\ell$ to p, where $\gamma(\rho)$ is the weight margin of the representation ρ on $W = \operatorname{Sym}^\ell(V) \otimes V_{\lambda^*}$. In particular, solving the p-scaling problem with input $(\pi, v, p, \gamma(\rho)/2\ell)$ suffices to solve the moment polytope membership problem for (π, v, p).*

Finally, we show that the p-capacity is log-concave in the parameter p. For this, it is convenient to generalize its definition from rational p to all of $C(G)$. Recall the Iwasawa decomposition (2.2) in the form $G = K \exp(i \operatorname{Lie}(T_K))N$ (which generalizes the decomposition of a matrix in $\mathrm{GL}(n)$ into a product of a unitary matrix, a diagonal matrix with positive diagonal entries, and an upper triangular matrix with ones on its diagonal). For rational $p = \lambda/\ell$ and $g = k \exp(H)b$, where $k \in K$, $H \in i \operatorname{Lie}(T_K)$, and $b \in N$, we have

$$\begin{aligned}\|(\pi(g)v)^{\otimes \ell} \otimes (\pi_{\lambda^*}(g)v_{\lambda^*})\| &= \|(\pi(\exp(H)b)v)^{\otimes \ell} \otimes (\pi_{\lambda^*}(\exp(H)b)v_{\lambda^*})\| \\ &= e^{\operatorname{tr}[\lambda^* H]}\|\pi(\exp(H)b)v\|^\ell,\end{aligned}$$

where we first used that the inner product is K-invariant and then that v_{λ^*} is a highest weight vector (hence invariant under the action of N, see Eq. (2.5)) of weight λ^* (hence transforms as (2.4)). Thus:

$$\operatorname{cap}_p(v) = \inf_{H \in i \operatorname{Lie}(T_K), b \in N} e^{\operatorname{tr}[p^* H]}\|\pi(\exp(H)b)v\|. \tag{3.26}$$

We will take this formula as the definition of the p-capacity for general $p \in C(G)$.

Proposition 3.32 (Log-concavity in p)**.** *For $0 \neq v \in V$, the function $C(G) \mapsto \mathbb{R} \cup \{-\infty\}$ given by* $p \mapsto \log \operatorname{cap}_p(v)$ *is* concave. *In particular, $\Delta^+(v) := \{p \in C(G) : \operatorname{cap}_p(v) > 0\}$ is a convex subset of the moment polytope $\Delta(v) \subseteq C(G)$.*

Proof. This follows directly from Eq. (3.26). Indeed, note that

$$\log \operatorname{cap}_p(v) = \inf_{H \in i \operatorname{Lie}(T_K), b \in N} \left(\operatorname{tr}[p^* H] - \log\|\pi(\exp(H)b)v\|\right).$$

Since $p \mapsto p^*$ is linear and the expression inside the infimum is affine in p^*, the log-capacity is manifestly concave in p. □

4 First-order algorithms

In this section, we state and analyze a first-order method for scaling and norm minimization, elaborating on the discussion in Sections 1.5.2 and 1.5.3. We first state a general *geodesic gradient descent* algorithm (Algorithm 4.1) and analyze it for arbitrary convex smooth left-K-invariant functions as defined in Section 3.2. Our algorithm for the scaling problem problem (Algorithm 4.2) is then obtained by specializing this algorithm to the log-norm function defined in Eq. (3.1). This is natural since norm minimization and scaling are dual to each other, as explained in Sections 1 and 3.4. In Section 4.3, we extend our first-order algorithm to the p-scaling problem (Algorithm 4.3).

```
Input:
  • Oracle access to the geodesic gradient $\nabla F$ of a left-$K$-invariant convex function $F\colon G \to \mathbb{R}$ (see Definitions 3.5 and 3.6),
  • a step size $\eta > 0$,
  • a number of iterations $T$.
Output: A group element $g \in G$.
.
Algorithm:
  1. Set $g_0 = I$ (identity element of the group $G$).
  2. For $t = 0, \dots, T-1$: Set $g_{t+1} := e^{-\eta\, \nabla F(g_t)} g_t$.
  3. Return $\arg\min_{g \in \{g_0, \dots, g_{T-1}\}} \|\nabla F(g)\|_F^2$
```

Algorithm 4.1: Geodesic first-order minimization algorithm (cf. Theorem 4.1).

4.1 General first-order optimization algorithm

We now state our general first-algorithm geodesic optimization algorithm and its analysis.

Theorem 4.1. *Let $F\colon G \to \mathbb{R}$ be a left-K-invariant function the sense that $F(kg) = F(g)$ for all $k \in K$, $g \in G$. Moreover, suppose that F is geodesically convex and L-smooth in the sense of Definition 3.5 for some $L > 0$, and that $F_{\text{inf}} := \inf_{g \in G} F(g) > 0$. For every $\varepsilon > 0$, Algorithm 4.1 with step size $\eta = 1/L$ and*

$$T \geqslant \frac{2L}{\varepsilon^2}\left(F(I) - F_{\text{inf}}\right)$$

iterations returns a group element $g \in G$ such that $\|\nabla F(g)\|_F \leqslant \varepsilon$.

Proof. Suppose to the contrary that $\|\nabla F(g_t)\|_F > \varepsilon$ for all $t = 0, \dots, T-1$. Then we find, using Lemma 3.7 with $H = -\eta \nabla F(g_t)$ for the first inequality, that

$$\begin{aligned} F(g_{t+1}) - F(g_t) &= F(e^{-\eta \nabla F(g_t)} g_t) - F(g_t) \\ &\leqslant -\eta\, \mathrm{tr}[\nabla F(g_t) \nabla F(g_t)] + \frac{L}{2}\eta^2 \|\nabla F(g_t)\|_F^2 \\ &= \left(\frac{L}{2}\eta^2 - \eta\right) \|\nabla F(g_t)\|_F^2 = -\frac{1}{2L}\|\nabla F(g_t)\|_F^2 < -\frac{\varepsilon^2}{2L} \end{aligned}$$

for $t = 0, \dots, T-1$. By a telescoping sum, we obtain the upper bound in

$$F_{\text{inf}} - F(I) \leqslant F(g_T) - F(g_0) < -\frac{T\varepsilon^2}{2L},$$

hence that $T < \frac{2L}{\varepsilon^2}\left(F(I) - F_{\text{inf}}\right)$. In view of our choice of T, this is the desired contradiction. □

4.2 Application to scaling and norm minimization problem

As in Section 2, let $\pi \colon G \to GL(V)$ be a representation of a symmetric subgroup $G \subseteq GL(n)$ on the vector space V with K-invariant inner product $\langle \cdot, \cdot \rangle$. We now specialize Algorithm 4.1 to the function $g \mapsto \log\|\pi(g)v\|$. The resulting algorithm is Algorithm 4.2:

Input:

- Oracle access to the moment map restricted to a group orbit, i.e., to the map $g \mapsto \mu(\pi(g)v)$,
- a number of iterations T.

Output: A group element $g \in G$.

.

Algorithm:

```
1. Set $g_0 = I$. Set a step size $\eta = \frac{1}{2N(\pi)^2}$.
2. For $t = 0, \ldots, T-1$: Set $g_{t+1} := e^{-\eta\, \mu(\pi(g_t)v)} g_t$.
3. **Return** $\arg\min_{g \in \{g_0, \ldots, g_{T-1}\}} \|\mu(\pi(g)v)\|_F^2$
```

Algorithm 4.2: Algorithm for the scaling problem (cf. Theorem 4.2).

The following theorem gives rigorous guarantees for Algorithm 4.2 in terms of the capacity of v and the weight norm of the Lie algebra representation (Definition 3.9).

Theorem 4.2 (First order algorithm for scaling; general version of Theorem 1.19)**.** *Let $v \in V$ be a vector with $\operatorname{cap}(v) > 0$. For every $\varepsilon > 0$, Algorithm 4.2 with*

$$T \geqslant \frac{4N(\pi)^2}{\varepsilon^2} \log\left(\frac{\|v\|}{\operatorname{cap}(v)}\right)$$

iterations returns a group element $g \in G$ such that $\|\mu(\pi(g)v)\|_F \leqslant \varepsilon$.

Proof. Since the gradient of the log-norm function F_v defined in Eq. (3.1) is computed by the moment map (Eq. (3.6)), we can interpret Algorithm 4.2 as the specialization of Algorithm 4.1 to F_v. Note that $F_v(I) = \log\|v\|$ and $F_{\text{inf}} = \log \operatorname{cap}(v)$. Moreover, F_v is convex and $2N(\pi)^2$-smooth by Proposition 3.12. Thus the claim follows from from Theorem 4.1. □

By Corollary 1.17, Theorem 4.2 implies that the first order algorithm also computes an approximation to capacity, however the runtime becomes inversely proportional to the weight margin:

Corollary 4.3 (First order algorithm for norm minimization)**.** *Let $v \in V$ be a vector with $\operatorname{cap}(v) > 0$. For every $\varepsilon > 0$, Algorithm 4.2 with step size $\eta = 1/(2N(\pi)^2)$ and*

$$T \geqslant \frac{4N(\pi)^2}{\gamma(\pi)^2 \varepsilon^2} \log\left(\frac{\|v\|}{\operatorname{cap}(v)}\right)$$

iterations returns an ε-approximate minimizer for log-capacity, i.e., a group element $g \in G$ such that $\log\|\pi(g)v\| - \log \operatorname{cap}(v) \leqslant \varepsilon$.

Remark 4.4. *Comparing Corollary 4.3 (first order) with Theorem 5.6 (second order), it is clear that the second order algorithm is better in terms of the dependence on the approximation parameter (with the dependence on weight margin and norm similar) if the goal is to approximate the capacity. However, the first order algorithm can be better if the goal is solve the scaling problem (Theorem 4.2) because of the non-dependence on the weight margin in this case.*

4.3 Application to p-scaling and moment polytopes

We now explain how to generalize Algorithm 4.2 to the optimization problem underlying p-scaling. Since the latter is characterized by the p-capacity (3.19), as explained in Section 3.6, this is achieved by replacing the log-norm objective function by its 'shifted' variant (3.22), namely

$$F_{v,p} \colon G \to \mathbb{R}, \quad F_{v,p}(g) = \log \|\pi(g)v\| + \frac{1}{\ell} \log \|\pi_{\lambda^*}(g)v_{\lambda^*}\|.$$

We state our first-order optimization algorithm in Algorithm 4.3.

Input:

- Oracle access to the moment map restricted to a group orbit, i.e., to the map $g \mapsto \mu(\pi(g)v)$,
- a rational point $p \in C(G)$,
- a number of iterations T.

Output: A group element g.

Algorithm:

1. Set $g_0 = I$. Set a step size $\eta = \frac{1}{2N^2}$, where $N^2 := N(\pi)^2 + \|p\|_F$.
2. For $t = 0, \ldots, T-1$: Set $g_{t+1} = e^{-\eta\,\left(\mu(\pi(g_t)v) + k_t p^* k_t^\dagger\right)} g_t$, where $g_t = k_t b_t$ according to the Iwasawa decomposition $G = KB$. (If $G = \mathrm{GL}(n)$ then this is the QR decomposition.)
3. **Return** g_t, where $t = \arg\min_{t=0,\ldots,T-1} \|\mu(\pi(g_t)v) + k_t p^* k_t^\dagger\|_F$

Algorithm 4.3: Algorithm for the p-scaling problem (cf. Theorem 4.5)

The following theorem gives rigorous guarantees on its performance

Theorem 4.5 (First order algorithm for p-scaling; general version of Theorem 1.20)**.** *Let $p \in C(G)$ and let $v \in V$ be a vector with $\operatorname{cap}_p(v) > 0$. Set $N^2 := N(\pi)^2 + \|p\|_F$. For every $\varepsilon > 0$, Algorithm 4.3 with*

$$T \geqslant \frac{4N^2}{\varepsilon^2} \log \left(\frac{\|v\|}{\operatorname{cap}_p(v)} \right)$$

iterations returns $g \in G$ such $\|\mu(\pi(g)v) - k(-p^*)k^\dagger\|_F \leqslant \varepsilon$, where $g = kb$ according to $G = KB$. In particular, $\|s(\mu(\pi(g)v)) - p\|_F \leqslant \varepsilon$.

Proof. Similarly to the proof of Theorem 4.2, we observe that Algorithm 4.3 is obtained by specializing Algorithm 4.1 to the function $F_{\nu,p}$, whose geodesic gradient is given by Eq. (3.25). Note that $F_{\nu,p},(I) = \log\|\nu\|$ and $\inf_{g\in G} F_{\nu,p}(g) = \log \operatorname{cap}_p(\nu)$. Since $F_{\nu,p}$ is moreover convex and 2N-smooth by Proposition 3.27, the first claim follows from Theorem 4.1. The second claim follows since $s(-k_t p^* k_t^\dagger) = p$, as mentioned above, and the map s is a contraction [Wal14, Lemma 4.10]. □

5 Second-order algorithms

In this section, we state our second order algorithm for $\operatorname{cap}(\nu)$. We first give a general algorithm for minimizing geodesically convex functions in Section 5.1. Next, in Section 5.2, we specialize to the norm minimization problem and derive our running time bounds using the analysis of gradient flow from Section 3.5. Again, we work in the setup introduced in Section 2, with $\pi\colon G \to GL(V)$ a representation of a symmetric subgroup $G \subseteq GL(n)$ and $\|\cdot\|$ a K-invariant norm on V.

5.1 General second-order optimization algorithm

Our starting point is the following second-order optimization algorithm.

```
Input:
  • Oracle access to the geodesic gradient $\nabla F$ and Hessian $\nabla^2 F$ of a left-K-invariant convex
    function $F\colon G \to \mathbb{R}$ (see Definitions 3.5 and 3.6),
  • a robustness parameter $R \geqslant 1$,
  • a number of iterations $T$.
Output: An element $g \in G$.
Algorithm:
  1. Set $g_0 = I$.
  2. For $t = 0, \dots, T-1$:
     (a) Compute the geodesic gradient $V := \nabla F(g_t)$ and Hessian $Q := \nabla^2 F(g_t)$ at $g_t$.
     (b) Solve the following (Euclidean) convex quadratic optimization problem:
         $$H_t := \arg\min \left\{ \operatorname{tr}[VH] + \frac{1}{2e}\operatorname{tr}[Q(H\otimes H)] \;:\; H \in i\,\mathrm{Lie}(K),\ \|H\|_F \leqslant \frac{1}{R} \right\}$$
     (c) Set $g_{t+1} := e^{H_t/e^2} g_t$.
  3. Return $g_T$.
```

Algorithm 5.1: Geodesic second-order minimization algorithm (cf. Theorem 5.1).

Algorithm 5.1 generalizes the algorithm of [AZGL+18] to arbitrary geodesically convex left-K-invariant functions $F\colon G \to \mathbb{R}$ (equivalently, on the symmetric space $K\backslash G \cong P$, cf. Sections 2.2

and 3.2). It is of the "box-constrained" form, where progress is made in steps by optimizing a simple function in a bounded region (determined by the robustness of the target function). If F is convex in the sense of Definition 3.5 then the geodesic Hessians $\nabla^2 F$ are positive definite, so the optimization problem in step 2, (b) of Algorithm 5.1 is an ordinary convex quadratic optimization problem on the real vector space $i\,\mathrm{Lie}(K)$, which can be solved using standard methods.

We now state our technical result about Algorithm 5.1. The group element $g_\star \in G$ should be thought of as a 'well-conditioned' approximate minimizer.

Theorem 5.1. *Let $F\colon G \to \mathbb{R}$ be a function that is left-K-invariant in the sense that $F(kg) = F(g)$ for all $k \in K$, $g \in G$. Moreover, suppose that F is R-robust in the sense of Definition 3.5 for some $R \geqslant 1$ (in particular, F is geodesically convex). Finally, let $g_\star \in G$ and let $D \geqslant 1$ be a 'diameter' constant such that*

$$D \geqslant \max_{F(g)\leqslant F(I)} \frac{1}{2}\left\|\log\left((g_\star g^{-1})^\dagger (g_\star g^{-1})\right)\right\|_F. \tag{5.1}$$

Then, Algorithm 5.1 with robustness parameter R and, if $F(g_\star) < F(I)$, with

$$T \geqslant e^2 DR \log\left(\frac{F(I) - F(g_\star)}{\varepsilon}\right)$$

iterations (otherwise, any number of iterations works) returns an element $g \in G$ such that $F(g) \leqslant F(g_\star) + \varepsilon$.

Proof. The proof is a straightforward generalization of the argument in [AZGL+18]. We prove the following assertions for $t \geqslant 0$:

1. $F(g_t) \leqslant F(I)$
2. $F(g_t) - F(g_\star) \leqslant \left(1 - \frac{1}{e^2 DR}\right)^t (F(I) - F(g_\star))$.

These two statements clearly imply the theorem: If $F(g_\star) < F(I)$ then the second assertion shows that $F(g_T) - F(g_\star) \leqslant \varepsilon$ for our choice of T. Otherwise, the first shows that $F(g_T) \leqslant F(g_\star)$ for any T.

We now prove the two statements by induction on $t \geqslant 0$: For $t = 0$, they are evident since $g_0 = I$. Now suppose they hold up to some t. Define

$$H_\star := \frac{1}{2}\log\left((g_\star g_t^{-1})^\dagger (g_\star g_t^{-1})\right).$$

Since $F(g_t) \leqslant F(I)$, by definition of D we have $\|H_\star\|_F \leqslant D$. Furthermore,

$$F(e^{H_\star} g_t) = F\left(\sqrt{(g_\star g_t^{-1})^\dagger (g_\star g_t^{-1})}\, g_t\right) = F(g_\star g_t^{-1} g_t) = F(g_\star), \tag{5.2}$$

where the second equality follows by the left-K-invariance of F and the polar decomposition $G = KP$. Now consider the quadratic approximation from Lemma 3.8, which asserts that, for all $\|H\|_F \leqslant 1/R$,

$$q_-(H) \leqslant F(e^H g_t) - F(g_t) \leqslant q_+(H) \tag{5.3}$$

where

$$q_+(H) := \partial_{s=0} F(e^{sH} g_t) + \frac{e}{2}\partial^2_{s=0} F(e^{sH} g_t),$$

$$q_-(H) := \partial_{s=0} F(e^{sH} g_t) + \frac{1}{2e} \partial^2_{s=0} F(e^{sH} g_t).$$

Note that H_t in Algorithm 5.1 is precisely the minimizer of q_- subject to $\|H\|_F \leqslant 1/R$. If we define $H_\diamond := H_\star/(DR)$ then $\|H_\diamond\|_F \leqslant 1/R$; together with the lower bound in Eq. (5.3), we find that

$$q_-(H_t) \leqslant q_-(H_\diamond) \leqslant F(e^{H_\diamond} g_t) - F(g_t). \tag{5.4}$$

Since F is geodesically convex in the sense of Definition 3.5, the function $h(s) := F(e^{sH_\star} g_t)$ is convex in $s \in \mathbb{R}$. In particular, since $DR \geqslant 1$, and using Eq. (5.2),

$$F(e^{H_\diamond} g_t) - F(g_t) = h\left(\tfrac{1}{DR}\right) - h(0) \leqslant \frac{1}{DR}\left(h(1) - h(0)\right) = \frac{1}{DR}\left(F(g_\star) - F(g_t)\right). \tag{5.5}$$

If we combine Eqs. (5.4) and (5.5), we get

$$q_-(H_t) \leqslant -\frac{1}{DR}\left(F(g_t) - F(g_\star)\right). \tag{5.6}$$

This shows that our choice of H_t makes significant progress in decreasing the *quadratic approximation* of F. It remains to show that we actually decrease F itself. Here we use that $q_-(H) = e^2 q_+(H/e^2)$ for all H. Using the upper bound in Eq. (5.3) and noting that $\|H_t/e^2\|_F \leqslant 1/R$, we find that

$$e^2\left(F(g_{t+1}) - F(g_t)\right) = e^2\left(F(e^{H_t/e^2} g_t) - F(g_t)\right) \leqslant e^2 q_+(H_t/e^2) = q_-(H_t). \tag{5.7}$$

We have $q_-(H_t) \leqslant 0$ by definition of H_t. Thus, Eq. (5.7) implies that $F(g_{t+1}) \leqslant F(g_t) \leqslant F(I)$, the latter by the induction hypothesis. This establishes the first assertion that we wanted to show. Moreover, if we combine Eqs. (5.6) and (5.7) then we obtain

$$e^2\left(F(g_{t+1}) - F(g_t)\right) \leqslant -\frac{1}{DR}\left(F(g_t) - F(g_\star)\right),$$

which can be rearranged as

$$F(g_{t+1}) - F(g_\star) \leqslant \left(1 - \frac{1}{e^2 DR}\right)\left(F(g_t) - F(g_\star)\right).$$

Using the induction hypothesis, this establishes the second assertion, concluding the induction. □

5.2 Application to norm minimization and scaling problem

We now describe our second order algorithm for the norm minimization problem and analyze its complexity. The algorithm is simply Algorithm 5.1 run on an appropriate function F, namely the log-norm function (3.8) plus a suitable regularizer, which we now define.

Definition 5.2 (Regularizer). *The* regularizer *for the group* G *is the function* reg: $G \to (0, \infty)$ *defined as*

$$\operatorname{reg}(g) := \|g\|_F^2 + \|g^{-1}\|_F^2 = \operatorname{tr}\left[g^\dagger g\right] + \operatorname{tr}\left[(g^\dagger g)^{-1}\right].$$

The function reg provides a convenient upper bound on the condition number $\kappa_F(g) := \|g\|_F \|g^{-1}\|_F$ of $g \in G$. Indeed, by the AM-GM inequality, we have

$$\kappa_F(g) \leqslant \frac{1}{2}\left(\|g\|_F^2 + \|g^{-1}\|_F^2\right) = \frac{1}{2}\operatorname{reg}(g).$$

Note that the singular values of g are always between $\operatorname{reg}(g)^{-1/2}$ and $\operatorname{reg}(g)^{1/2}$. Furthermore, reg is minimized at the identity, so we have $\operatorname{reg}(g) \geqslant 2n$ for all $g \in G$. We first analyze the smoothness and robustness of the function reg.

Lemma 5.3. *The function* $\operatorname{reg} : G \to (0, \infty)$ *is left-*K*-invariant and 2-robust.*

Proof. Since $K \subseteq U(n)$, we have that $(kg)^\dagger kg = g^\dagger g$ for every $k \in K$ and $g \in G$, so reg is clearly left-K-invariant. To see that it is 2-robust, we prove that each term is individually 2-robust.

Consider the representation $\pi\colon G \to GL(V)$ obtained by letting G act on $V = \operatorname{Mat}(n)$ by left multiplication, i.e., $\pi(g)M = gM$. Equip V with the Hilbert-Schmidt inner product, which induces the Frobenius norm, and note that K acts unitarily. Then, $\operatorname{tr}\left[g^\dagger g\right] = \|\pi(g)I\|_F^2$, so $g \mapsto \operatorname{tr}\left[g^\dagger g\right]$ is nothing but the norm-square function for this representation and $v = I$ the identity matrix. By Proposition 3.16 we obtain that $g \mapsto \operatorname{tr}\left[g^\dagger g\right]$ is $2N(\pi)$-robust. Finally, Proposition 3.10 shows that $N(\pi) = 1$. This can also be calculated directly: The Lie algebra representation is given by $\Pi(H)M = HM$ for $H \in \operatorname{Lie}(G)$ and $M \in \operatorname{Mat}(n)$. Since the Frobenius norm is submultiplicative, it holds that $\|\Pi(H)M\|_F \leqslant \|H\|_F \|M\|_F$. It follows that $\|\Pi(H)\|_{op} \leqslant \|H\|_F$, and hence $N(\pi) = 1$.

Similarly, the map $g \mapsto \operatorname{tr}\left[(g^\dagger g)^{-1}\right]$ can be interpreted as the norm-square function for the representation $\pi\colon G \to GL(V)$ defined by $\pi(g)M = Mg^{-1}$. Again, this representation has weight norm 1, so we may conclude that reg is 2-robust. □

Next, we show that by adding a suitable multiple of the regularizer, Algorithm 5.1 can be made efficient for the norm minimization problem – assuming that there exists a well-conditioned approximate minimizer. We consider the following objective function for $v \in V$, $\kappa > 0$, and $\varepsilon > 0$:

$$F_{v,\kappa,\varepsilon} : G \to \mathbb{R}, \quad F_{v,\kappa,\varepsilon}(g) := F_v(g) + \frac{\varepsilon}{\kappa}\operatorname{reg}(g) = \log\|\pi(g)v\| + \frac{\varepsilon}{\kappa}\operatorname{reg}(g) \tag{5.8}$$

Proposition 5.4. *Let* $\varepsilon > 0$, $\kappa > 0$, *and* $C \geqslant \log(\|v\|/\operatorname{cap}(v))$.

1. *The function* $F_{v,\kappa,\varepsilon}$ *is left-*K*-invariant and* $4N$*-robust, where* $N := \max\{N(\pi), 1/2\}$.
2. *Suppose there exists an element* $g_\star \in G$ *with* $\log\|\pi(g_\star)v\| \leqslant \log\operatorname{cap}(v) + \varepsilon$ *and* $\operatorname{reg}(g_\star) \leqslant \kappa$. *Then, Algorithm 5.1 applied to* $F_{v,\kappa,\varepsilon}$ *with robustness parameter* $R = 4N$ *and*

$$T \geqslant 8e^2 N\sqrt{n}\left(\log\kappa + \log\left(1 + \frac{C}{\varepsilon}\right)\right)\log\left(\frac{C}{\varepsilon}\right)$$

iterations returns a group element $g \in G$ *such that* $\log\|\pi(g)v\| \leqslant \log\operatorname{cap}(v) + 3\varepsilon$.

Proof. We abbreviate $F := F_{v,\kappa,\varepsilon}$. By Proposition 3.14 and Lemma 5.3, F is left-K-invariant and a sum of a $4N(\pi)$-robust function and a 2-robust function, hence $4N$-robust. This shows the first claim.

To prove the second claim, we apply Theorem 5.1 to the function F. The theorem asserts that, if we run Algorithm 5.1 with robustness parameter $R = 4N$ and a suitable number of iterations, we obtain a group element $g \in G$ such that $F(g) \leqslant F(g_\star) + \varepsilon$. The latter implies that

$$\log\|\pi(g)v\| \leqslant F(g) \leqslant F(g_\star) + \varepsilon = \log\|\pi(g_\star)v\| + \frac{\varepsilon}{\kappa}\operatorname{reg}(g_\star) + \varepsilon \leqslant \log\operatorname{cap}(v) + 3\varepsilon,$$

as desired. It suffices to bound the number of iterations. Only the case that $F(g_\star) < F(I)$ is of interest. Here, Theorem 5.1 asserts that

$$T \geqslant e^2 D R \log\left(\frac{F(I) - F(g_\star)}{\varepsilon}\right)$$

iterations suffice. We first note that, since $\operatorname{reg}(g)$ is minimal at $g = I$,

$$F(I) - F(g_\star) \leqslant \log\|v\| - \log\|\pi(g_\star)v\| \leqslant \log\frac{\|v\|}{\operatorname{cap}(v)} \leqslant C. \tag{5.9}$$

To find a suitable diameter bound D, recall that we need $D \geqslant 1$ as well as (cf. Eq. (5.1))

$$D \geqslant \max_{F(g) \leqslant F(I)} \frac{1}{2}\left\|\log\left((g_\star g^{-1})^\dagger (g_\star g^{-1})\right)\right\|_F. \tag{5.10}$$

The condition $F(g) < F(I)$ implies that

$$\operatorname{reg}(g) \leqslant \operatorname{reg}(I) + \frac{\kappa}{\varepsilon}\left(\log\|v\| - \log\|\pi(g)v\|\right) \leqslant 2n + \frac{\kappa}{\varepsilon}\log\frac{\|v\|}{\operatorname{cap}(v)} \leqslant \kappa\left(1 + \frac{C}{\varepsilon}\right).$$

In the last step, we used that $\kappa \geqslant \operatorname{reg}(g_\star) \geqslant 2n$ (recall that reg is bounded from below by $2n$). As mentioned earlier, this implies that the singular values of g are between $\kappa^{-1/2}(1 + \frac{C}{\varepsilon})^{-1/2}$ and $\kappa^{1/2}(1 + \frac{C}{\varepsilon})^{1/2}$. We also have $\operatorname{reg}(g_\star) \leqslant \kappa$, which implies that the singular values of $g_\star$ are between $\kappa^{-1/2}$ and $\kappa^{1/2}$. It follows that the singular values of $g_\star g^{-1}$ are between $\kappa^{-1}(1 + \frac{C}{\varepsilon})^{-1/2}$ and $\kappa(1 + \frac{C}{\varepsilon})^{1/2}$, so

$$\frac{1}{2}\left\|\log\left((g_\star g^{-1})^\dagger (g_\star g^{-1})\right)\right\|_F \leqslant \frac{\sqrt{n}}{2}\log\left(\kappa^2\left(1 + \frac{C}{\varepsilon}\right)\right).$$

Thus, we choose

$$D := \sqrt{n}\log\left(\kappa^2\left(1 + \frac{C}{\varepsilon}\right)\right),$$

so that Eq. (5.10) and $D \geqslant 1$ are satisfied. Together with Eq. (5.9), we find that, indeed,

$$\begin{aligned} e^2 D R \log\left(\frac{F(I) - F(g_\star)}{\varepsilon}\right) &\leqslant e^2\sqrt{n}\log\left(\kappa^2\left(1 + \frac{C}{\varepsilon}\right)\right) 4N \log\left(\frac{C}{\varepsilon}\right) \\ &\leqslant 8e^2 N\sqrt{n}\left(\log\kappa + \log\left(1 + \frac{C}{\varepsilon}\right)\right)\log\left(\frac{C}{\varepsilon}\right) \end{aligned}$$

iterations suffice. □

Finally, we show that there exist well-conditioned approximate minimizers. Our diameter bounds are described in terms of the *weight margin* $\gamma(\pi)$ defined in Definition 3.17, which is the closest the convex hull of a subset of weights can be to the origin without containing it. The main mathematical tool in our proof is the gradient flow analyzed in Section 3.5.

Proposition 5.5 (Diameter bound). *Let* $v \in V$ *be a vector with* $\operatorname{cap}(v) > 0$ *and let* $\varepsilon > 0$. *Then there exists a group element* $g_\star \in G$ *such that* $\log\|\pi(g_\star)v\| \leqslant \log\operatorname{cap}(v) + \varepsilon$ *and*

$$\operatorname{reg}(g_\star) \leqslant 2n \left(\frac{\|v\|^2}{2\operatorname{cap}^2(v)\varepsilon}\right)^{\frac{1}{\gamma(\pi)}}.$$

Proof. Let $v\colon [0, T_v) \to V$ denote the solution of the gradient flow on its maximal domain of definition (Definition 3.23). Recall from Item 4 of Proposition 3.24 that $v(t) = \pi(g(t))v$ for $t \in [0, T_v)$, where $g\colon [0, T_v) \to G$ is a solution to the ordinary differential equation

$$g'(t) = -2\frac{\mu(\pi(g(t))v)}{\|\mu(\pi(g(t))v)\|_F} g(t) = -2\frac{\mu(v(t))}{\|\mu(v(t))\|_F} g(t), \quad g(0) = I. \tag{5.11}$$

By Corollary 3.26, there is some $T < T_v$ with

$$T \leqslant \frac{1}{4\gamma(\pi)} \log \frac{\|v\|^2}{2\operatorname{cap}^2(v)\varepsilon} \tag{5.12}$$

such that, for $g_\star := g(T)$, we have

$$\log\|\pi(g_\star)v\| = \log\|v(T)\| \leqslant \log\operatorname{cap}(v) + \varepsilon.$$

It remains to verify the bound on $\operatorname{reg}(g_\star)$. We first bound $\varphi(t) := \operatorname{tr}\left[g(t)^\dagger g(t)\right]$ by an ODE argument. By taking derivatives, using Eq. (5.11),

$$\begin{aligned}\partial_t \operatorname{tr}\left[g(t)^\dagger g(t)\right] &= \operatorname{tr}\left[g'(t)^\dagger g(t)\right] + \operatorname{tr}\left[g(t)^\dagger g'(t)\right] = -\frac{4}{\|\mu(v(t))\|_F} \operatorname{tr}\left[g(t)^\dagger \mu(v(t)) g(t)\right] \\ &\leqslant 4\frac{\|\mu(v(t))\|_{op}}{\|\mu(v(t))\|_F} \operatorname{tr}\left[g(t)^\dagger g(t)\right] \leqslant 4\operatorname{tr}\left[g(t)^\dagger g(t)\right].\end{aligned}$$

For the first inequality we used the general fact that $|\operatorname{tr}[AB]| \leqslant \|A\|_{op} \operatorname{tr}[B]$ for any Hermitian matrix A and positive semidefinite matrix B. Thus, we have shown that $\varphi'(t) \leqslant 4\varphi(t)$, which implies that $\operatorname{tr}\left[g(t)^\dagger g(t)\right] = \varphi(t) \leqslant \varphi(0)e^{4t} = ne^{4t}$. Similarly,

$$\begin{aligned}\partial_t \operatorname{tr}\left[\left(g(t)^\dagger g(t)\right)^{-1}\right] &= -\operatorname{tr}\left[\left(g(t)^\dagger g(t)\right)^{-1} \partial_t \left(g(t)^\dagger g(t)\right) \left(g(t)^\dagger g(t)\right)^{-1}\right] \\ &= -\operatorname{tr}\left[\left(g(t)^\dagger g(t)\right)^{-1} \left((g'(t))^\dagger g(t) + g(t)^\dagger g'(t)\right) \left(g(t)^\dagger g(t)\right)^{-1}\right] \\ &= \frac{4}{\|\mu(v(t))\|_F} \operatorname{tr}\left[\left(g(t)^\dagger g(t)\right)^{-1} g(t)^\dagger \mu(v(t)) g(t) \left(g(t)^\dagger g(t)\right)^{-1}\right] \\ &= \frac{4}{\|\mu(v(t))\|_F} \operatorname{tr}\left[g(t)^{-1} \mu(v(t)) g(t)^{-\dagger}\right] \\ &\leqslant 4\frac{\|\mu(v(t))\|_{op}}{\|\mu(v(t))\|_F} \operatorname{tr}\left[g(t)^{-1} g(t)^{-\dagger}\right] \leqslant 4\operatorname{tr}\left[\left(g(t)^\dagger g(t)\right)^{-1}\right],\end{aligned}$$

so $\operatorname{tr}\left[(g(t)^\dagger g(t))^{-1}\right] \leqslant ne^{4t}$. Together, evaluating at $t = T$, and using Eq. (5.12), we obtain

$$\operatorname{reg}(g_\star) \leqslant 2ne^{4T} \leqslant 2ne^{\frac{1}{\gamma(\pi)} \log \frac{\|v\|^2}{2\operatorname{cap}^2(v)\varepsilon}} = 2n\left(\frac{\|v\|^2}{2\operatorname{cap}^2(v)\varepsilon}\right)^{\frac{1}{\gamma(\pi)}},$$

completing the proof. □

We thus obtain the main theorem of this section – a second order optimization algorithm for minimizing the norm, i.e., approximating the capacity.

Theorem 5.6 (Second order algorithm for norm minimization; general statement of Theorem 1.21)**.** *Let $v \in V$ be a vector with $\operatorname{cap}(v) > 0$. Let $0 < \varepsilon < 1/2$ and $C \geqslant \log(\|v\|/\operatorname{cap}(v))$. Set $\gamma := \min\{\gamma(\pi), 1\}$, $N := \max\{N(\pi), 1/2\}$, and $\kappa := 2n\left(e^{2C}/2\varepsilon\right)^{1/\gamma}$. Then, Algorithm 5.1 applied to the function $F_{v,\kappa,\varepsilon}$ from Eq. (5.8), robustness parameter $R = 4N$, and*

$$T \geqslant 24e^2 \frac{N\sqrt{n}}{\gamma}\left(\log\frac{n}{\varepsilon} + C\right)\log\frac{C}{\varepsilon}$$

iterations returns a group element $g \in G$ such that $\log\|\pi(g)v\| \leqslant \log\operatorname{cap}(v) + 3\varepsilon$.

Proof. According to Proposition 5.5, there exists $g_\star \in G$ such that $\log\|\pi(g_\star)v\| \leqslant \log\operatorname{cap}(v) + \varepsilon$ and

$$\operatorname{reg}(g_\star) \leqslant 2n\left(\frac{\|v\|^2}{2\operatorname{cap}^2(v)\varepsilon}\right)^{\frac{1}{\gamma(\pi)}} \leqslant 2n\left(\frac{e^{2C}}{2\varepsilon}\right)^{\frac{1}{\gamma(\pi)}} \leqslant 2n\left(\frac{e^{2C}}{2\varepsilon}\right)^{\frac{1}{\gamma}} = \kappa,$$

where the second inequality holds by the assumption on C. The last inequality follows from $\gamma(\pi) \geqslant \gamma$ and $e^{2C}/2\varepsilon \geqslant 1$, which holds by the assumption $\varepsilon \leqslant 1/2$.

Now apply Item 2 of Proposition 5.4 with the element $g_\star$, which asserts that Algorithm 5.1 applied to the function $F_{v,\kappa,\varepsilon}$ and robustness parameter $R = 4N$ returns the desired group element $g \in G$ in a number of iterations at most

$$8e^2N\sqrt{n}\left(\log 2n + \frac{1}{\gamma}\log\frac{e^{2C}}{2\varepsilon} + \log\left(1 + \frac{C}{\varepsilon}\right)\right)\log\frac{C}{\varepsilon} \leqslant 24e^2\frac{N\sqrt{n}}{\gamma}\left(\log\frac{n}{\varepsilon} + C\right)\log\frac{C}{\varepsilon},$$

where the inequality is obtained using $\gamma \leqslant 1$ and $\log(1 + C/\varepsilon) \leqslant \log(1/\varepsilon) + C$. □

By Corollary 1.17, the second order algorithm described in Theorem 1.21 can also be used to address the scaling problem.

6 Bounds on weight norms and weight margins

In this section we prove general upper bounds on weight norms and lower bounds on weight margins. We focus on $G = GL(n_1) \times \cdots \times GL(n_k)$ and its subgroups, which capture many interesting applications. Note that G can be realized as a symmetric subgroup of $GL(L)$, where $L := \sum_{i=1}^k n_i$. Thus its weights can be viewed as elements of the lattice $\mathbb{Z}^L$ (see Section 2.3 and Table 2.1). Under this identification, the Frobenius norm $\|\omega\|_F$ of a weight ω equals its Euclidean norm $\|\omega\|_2$. We define the *weight matrix* $M(\pi)$ of a representation π as the integer matrix of format $|\Omega(\pi)| \times L$ whose rows are labeled and given by $\Omega(\pi)$, the weights of π (each weight appears once, irrespective of its multiplicity in π). Note that, by the definition of the norm $N(\pi)$, we have $\|\omega\|_2 \leqslant N(\pi)$ for every row ω of $M(\pi)$.

We first state a simple bound on the weight norm of polynomial representations.

Lemma 6.1 (Weight norm of homogeneous representations)**.** *Let π be a polynomial representation of $GL(n_1) \times \cdots \times GL(n_k)$ that is homogeneous of degree d. Then, $N(\pi) \leqslant d$.*

Proof. By Proposition 3.10, it suffices to bound the Euclidean norm of any highest weight $\lambda \in \mathbb{Z}^L$ corresponding to a homogeneous polynomial representation of degree d. The coefficients of such a highest weight are non-negative and their sum is d. Thus, $\|\lambda\|_2 \leqslant \|\lambda\|_1 = \sum_{i=1}^{L} \lambda_i = d$. □

Remark 6.2. *The weight norm of a representation never increases when we restrict it to a subgroup. In particular, the bound in Lemma 6.1 also holds when we restrict to subgroups of* $\mathrm{GL}(n_1) \times \cdots \times \mathrm{GL}(n_k)$*, such as products of special linear groups or of tori.*

We now compute the weight norm for several important applications.

Example 6.3 (Weight norm for matrix and operator scaling)**.** *Consider the action of* $G = \mathrm{GL}(n) \times \mathrm{GL}(n)$ *on* $\mathrm{Mat}(n)$ *by* $\pi(g,h)M := gMh^T$*. Clearly,* π *is polynomial and homogeneous of degree* 2*, so* $N(\pi) \leqslant 2$*. In fact,* $N(\pi) = \sqrt{2}$ *since it is an irreducible representation of highest weight* $\lambda = e_1 + e_{n+1}$*. The same holds for the simultaneous left-right action on* $\mathrm{Mat}(n)^k$*, since it is a sum of* k *copies of this representation.*

Matrix scaling corresponds to restricting to $\mathrm{ST}(n) \times \mathrm{ST}(n)$ *and* $k = 1$*, whereas operator scaling as in Example 1.5 is obtained by restricting to* $\mathrm{SL}(n) \times \mathrm{SL}(n)$*. Thus, Remark 6.2 shows that* $N(\pi) \leqslant \sqrt{2}$ *also holds for these representations.*

Example 6.4 (Weight norm for tensor scaling)**.** *Consider the action of* $G = \mathrm{GL}(n_1) \times \cdots \times \mathrm{GL}(n_k)$ *on* $V = \mathbb{C}^{n_1} \otimes \cdots \otimes \mathbb{C}^{n_k}$ *by* $\pi(g_1, \ldots, g_k)X := (g_1 \otimes \cdots \otimes g_k)X$ *(generalizing Example 1.4). Clearly,* π *is polynomial and homogeneous of degree* k*, hence* $N(\pi) \leqslant k$ *by Lemma 6.1. In fact,* π *is irreducible with highest weight* $\lambda = e_1 + e_{n_1+1} + \cdots + e_{L-n_k+1}$*, hence* $N(\pi) = \sqrt{k}$ *by Proposition 3.10. The same holds for tuples of* k*-tensors.*

In Proposition 6.12 we prove that the weight norm for group representations associated with quivers satisfies $N(\pi) = \sqrt{2}$ (unless the representation is trivial).

In the remainder of this section we will compute bounds on the weight margin. Our main technical tool for deriving weight margin bounds is the *gap* of its weight matrix $M(\pi)$ (defined below). The gap can be seen as a condition measure of $M(\pi)$. For several representations of interest, the weight matrix is *totally unimodular* (including for quivers, see Proposition 6.12). Such matrices turn out to have a large gap, which thus implies a large weight margin.

Let $A \in \mathbb{R}^{r \times L}$ be such that $r \leqslant L$ and denote by $\sigma_{\min}(A)$ the smallest of the r singular values of A. It is well known that $\sigma_{\min}(A)$ measures the distance of A to the set of matrices of rank less than r. In particular, $\sigma_{\min}(A) > 0$ iff A has (full) rank r. We consider now matrices $M \in \mathbb{R}^{s \times L}$ with the property that all of its $r \times L$ submatrices M_I are either singular or have large $\sigma_{\min}(M_I)$. In order to make this quantitative, we define the *gap* of M as follows.

Definition 6.5 (Gap)**.** *The* gap $\sigma(M)$ *of a matrix* $M \in \mathbb{R}^{s \times L}$ *is the minimum of* $\sigma_{\min}(M_I)$ *over all* $1 \leqslant r \leqslant \min(s, L)$ *and all* $r \times L$ *submatrices* M_I *of* M *which have rank* r*.*

We can lower bound the weight margin of a representation in terms of the gap of its weight matrix.

Proposition 6.6 (Weight margin lower bound in terms of gap)**.** *Let* π *be a rational representation of* $G = \mathrm{GL}(n_1) \times \cdots \times \mathrm{GL}(n_k)$ *and put* $L := \sum_{i=1}^{k} n_i$*. Then,* $\gamma(\pi) \geqslant \sigma(M(\pi))L^{-\frac{1}{2}}$*.*

Proof. Recall that we view $\Omega(\pi) \subseteq \mathbb{Z}^L$. Let $\Gamma \subseteq \Omega(\pi)$ be such that 0 is not contained in the polytope $P := \mathrm{conv}(\Gamma) \subseteq \mathbb{R}^L$. Put $\sigma := \sigma(M(\pi))$. We need to show that $d(0, P) \geqslant \sigma L^{-\frac{1}{2}}$ with respect to the Euclidean distance d. It is easy to see that there is a face F of P with $d(0, P) = d(0, F)$ such that $0 \notin \operatorname{aff}(F)$. Let ν be the point in $\operatorname{aff}(F)$ closest to the origin. Since $d(0, F) \geqslant d(0, \operatorname{aff}(F)) = \|\nu\|_2$, it suffices to prove that $\|\nu\|_2 \geqslant \sigma L^{-\frac{1}{2}}$.

Since the face F is the convex hull of a subset of Γ, there exist $\omega_1, \ldots, \omega_r \in \Gamma$ that are affinely independent and affinely span $\operatorname{aff}(F)$. Since $0 \notin \operatorname{aff}(F)$, it holds that $\dim \operatorname{aff}(F) = r - 1 \leqslant n - 1$ and the $\omega_1, \ldots, \omega_r$ are in fact linearly independent. Thus, the matrix $A \in \mathbb{R}^{r \times L}$ with rows $\omega_1, \ldots, \omega_r$ has rank r. Therefore, $\sigma_{\min}(A) \geqslant \sigma$ by the definition of the gap of $M(\pi)$. Note that ν is in the row span of A. Since ν is the point in $\operatorname{aff}(F)$ closest to the origin, we have that $(\omega_i - \nu) \cdot \nu = 0$ and hence $\omega_i \cdot \nu = \|\nu\|_2^2$ for all $i \in [r]$. Thus, $x := \|\nu\|_2^{-2} \nu$ satisfies $Ax = \mathbf{1}$, where $\mathbf{1} \in \mathbb{R}^r$ is the all-ones vector, and is in the row span of A. With Lemma 6.7 below we conclude that

$$\|x\|_2 \leqslant \sigma^{-1} \|\mathbf{1}\|_2 = \sigma^{-1} r^{1/2} \leqslant \sigma^{-1} L^{1/2}.$$

Therefore, $\|\nu\|_2 = \|x\|_2^{-1} \geqslant \sigma L^{-1/2}$, which completes the proof. □

Lemma 6.7. *Let $A \in \mathbb{R}^{r \times L}$ be of (full) rank r such that all of its singular values are at least $\varepsilon > 0$. Let $y \in \mathbb{R}^r$. Then the unique $x \in \mathbb{R}^L$ in the row span of A such that $Ax = y$ satisfies $\|x\|_2 \leqslant \varepsilon^{-1} \|y\|_2$.*

Proof. It is easily checked that $x = A^T (AA^T)^{-1} y$. We have

$$\|x\|_2^2 = x^T x = y^T (AA^T)^{-T} AA^T (AA^T)^{-1} y = y^T (AA^T)^{-T} y \leqslant \|(AA^T)^{-T}\|_{\text{op}} \|y\|_2^2.$$

The smallest nonzero singular value of A equals $\|(AA^T)^{-1}\|_{\text{op}}^{-1/2} \geqslant \varepsilon$. Therefore, $\|x\|_2^2 \leqslant \varepsilon^{-2} \|y\|_2^2$ as claimed. □

As a consequence, we obtain a general bound that applies to all representations.

Theorem 6.8 (General weight margin lower bound)**.** *Let π be a representation of $G = \mathrm{GL}(n_1) \times \cdots \times \mathrm{GL}(n_k)$ and put $L := \sum_{i=1}^k n_i$. Then, $\gamma(\pi) \geqslant N(\pi)^{-L} L^{-1}$.*

Proof. By Proposition 6.6, it is sufficient to show that $\sigma(M(\pi)) \geqslant N(\pi)^{-L} L^{-\frac{1}{2}}$. Consider a submatrix $A \in \mathbb{R}^{r \times L}$ of $M(\pi)$ of rank r. It is sufficient to prove that $\sigma_{\min}(A) \geqslant N(\pi)^{-L} r^{-\frac{1}{2}}$. Using the formula $\sigma_{\min}(A) = \|(AA^T)^{-1}\|_{\text{op}}^{-1/2}$ and noting that $\|(AA^T)^{-1}\|_{\text{op}} \leqslant \operatorname{tr}\left[(AA^T)^{-1}\right]$, we see that it is sufficient to prove that

$$\operatorname{tr}\left[(AA^T)^{-1}\right] \leqslant r\, N(\pi)^{2L}. \tag{6.1}$$

Since the $r \times r$ matrix AA^T has integer entries of magnitude at most $N(\pi)^2$, we note that

$$\operatorname{tr}\left[AA^T\right] = \|A\|_F^2 \leqslant r N(\pi)^2.$$

The AM-GM inequality, together with the fact that A has integer entries, implies that

$$1 \leqslant \det(AA^T) \leqslant \left(\frac{1}{r} \operatorname{trace}(AA^T)\right)^r \leqslant N(\pi)^{2r}. \tag{6.2}$$

The same bounds hold for the determinants of the principal minors of AA^T. Thus, using the formula for the inverse of a matrix in terms of its adjugate, we see that each diagonal entry of $(AA^T)^{-1}$ is upper bounded by $N(\pi)^{2(r-1)}$. Therefore,

$$\operatorname{tr}\left[(AA^T)^{-1}\right] \leqslant r N(\pi)^{2r}$$

which is the desired inequality (6.1). □

A slight modification to the proof of Theorem 6.8 also gives weight margin lower bounds for representations of special linear groups. For simplicity we only consider a single $\mathrm{SL}(n)$-factor.

Theorem 6.9 (General weight margin lower bound for SL)**.** *Let π be a representation of $\mathrm{GL}(n)$ that is homogeneous of degree d. Denote by π_0 its restriction to $\mathrm{SL}(n)$. Then, $\gamma(\pi_0) \geqslant N(\pi)^{-n} n^{-3/2}$.*

Proof. The weights of the $\mathrm{SL}(n)$-representation π_0 are given by the orthogonal projections of the weights of the $\mathrm{GL}(n)$-representation π onto the subspace of $\mathbb{R}^n$ consisting of vectors that sum to zero. Since $\sum_{i=1}^n \omega_i = d$ for any weight ω of π, this orthogonal projection can be calculated as $\omega - (d, \dots, d)/n$. Thus, the weight matrix $M(\pi_0)$ is given by $M(\pi) - \frac{d}{n} \mathbf{1}_{|\Omega(\pi)|} \mathbf{1}_n^T$. We claim that $M(\pi_0)$ has gap at least $N(\pi)^{-n} n^{-3/2}$.

To prove this, let A be an $r \times r$ submatrix of $M(\pi_0)$ of rank r and proceed as in the proof of Theorem 6.8. Because $M(\pi_0)$ is not integral, the first inequality in Eq. (6.2) does not hold, but we may write $A = B + \frac{d}{n} \mathbf{1}_r \mathbf{1}_r^T$ where B is integral (it is the corresponding submatrix of $M(\pi)$). Thus, by the rank-one update formula for the determinant, $\det(AA^T) = |\det(A)|^2 = |\det(B) + \frac{d}{n} \mathbf{1}_r^T \operatorname{adj}(B) \mathbf{1}_r|^2$ where $\operatorname{adj}(B)$ denotes the adjugate matrix of B. As A is nonsingular, $|\det(A)|^2 > 0$, but $\det(B)$ and $\operatorname{adj}(B)$ are integral, so $|\det(A)| \geqslant 1/n$. Now we have

$$\frac{1}{n^2} \leqslant \det(AA^T) \leqslant \left(\frac{1}{r} \operatorname{trace}(AA^T)\right)^r \leqslant N(\pi_0)^{2r}.$$

Following the rest of the proof of Theorem 6.8, we find that each diagonal entry of $(AA^T)^{-1}$ is bounded by $n^2 N(\pi)^{2(r-1)}$ which yields $\sigma_{\min}(A) \geqslant N(\pi)^{-n} r^{-\frac{1}{2}} n^{-1}$. This implies that $M(\pi_0)$ has gap at least $N(\pi)^{-n} n^{-3/2}$. □

When the weight matrix is totally unimodular one can prove much better bounds. Recall that an integer matrix is called *totally unimodular* if all its subdeterminants are 0, 1, or −1. We first show that totally unimodular matrices have a large gap.

Lemma 6.10. *A totally unimodular matrix $M \in \mathbb{R}^{s \times L}$ satisfies $\sigma_{\min}(M) \geqslant L^{-1}$.*

Proof. It suffices to show that a totally unimodular matrix $A \in \mathbb{R}^{r \times L}$ of rank $r \leqslant L$ satisfies $\sigma(A) \geqslant r^{-1}$. The smallest singular value of A equals the minimum of $\|A^T y\|_2 / \|y\|_2$ over all nonzero $y \in \mathbb{R}^r$. Let A' be an invertible $r \times r$ submatrix of A. By the total unimodularity of A, we have that $\det(A') = \pm 1$. For $y \in \mathbb{R}^r$, we put $x := A^T y \in \mathbb{R}^L$. Then $x' := (A')^T y$ is a subvector of x and hence $\|x'\|_2 \leqslant \|x\|_2$. It is thus sufficient to show that $\|x'\|_2 \geqslant r^{-1} \|y\|_2$. Solving $x' = (A')^T y$ by Cramer's rule we get that, for all $i \in [r]$,

$$y_i = \frac{\det(B_i)}{\det(A')} = \pm \det(B_i),$$

where B_i is the matrix obtained by replacing the i^{th} column of $(A')^T$ with the vector x'. By performing the Laplace expansion with respect to the i^{th} column of B_i, using the total unimodularity of A, we get $|y_i| = |\det(B_i)| \leqslant \|x'\|_1$. Hence, $\|y\|_2 \leqslant \sqrt{r} \|x\|_1 \leqslant r \|x'\|_2$, as claimed. □

The next corollary follows immediately from Proposition 6.6 and Lemma 6.10.

Corollary 6.11 (Weight margin lower bound for totally unimodular weight matrices)**.** *Let π be a rational representation of $G = \mathrm{GL}(n_1) \times \cdots \times \mathrm{GL}(n_k)$ and put $L := \sum_{i=1}^k n_i$. If the weight matrix $M(\pi)$ is totally unimodular then $\gamma(\pi) \geqslant L^{-\frac{3}{2}}$.*

We now discuss an important class of examples where the weight matrix is totally unimodular.

Proposition 6.12 (Weight margin and norm for quivers). *Consider a quiver Q with vertex set Q_0 and edge set Q_1, and let $\mathbf{n} = (n_x)_{x \in Q_0}$ be a vector of natural numbers. Let π denote the representation of $G = \prod_{x \in Q_0} \mathrm{GL}(n_x)$ on $V = \bigoplus_{a\colon x \to y \in Q_1} \mathrm{Mat}(n_y, n_x)$ associated with the quiver Q and dimension vector $\mathbf{n}$, as in Example 1.3. Then, π has a totally unimodular weight matrix. In particular, $\gamma(\pi) \geqslant (\sum_{x \in Q_0} n_x)^{-3/2}$. Moreover, $N(\pi) = \sqrt{2}$ unless the representation is trivial (there are only self-loops and all $n_x = 1$).*

Proof. We will use the well-known fact that the incidence matrix of a directed graph is totally unimodular (see, e.g., [Sch86, §19.3, Example 2]). Recall that the *incidence matrix* of a directed graph with vertex set $[L]$ and m edges is defined as the $m \times L$ matrix with one row $e_i - e_j$ for each edge (i, j). Here, $e_i \in \mathbb{R}^n$ denotes the i^{th} standard basis vector.

Set $L := \sum_{x \in Q_0} n_x$. It will be convenient to identify $\mathbb{Z}^L \cong \bigoplus_{x \in Q_0} \mathbb{Z}^{n_x}$ and denote the standard basis vectors by $e_{x,i}$ for $x \in Q_0$ and $i \in [n_x]$. For $a\colon x \to y \in Q_1$, $i \in [n_y]$, and $j \in [n_x]$, let $E^a_{i,j}$ denote the vector in $V = \bigoplus_{a\colon x \to y \in Q_1} \mathrm{Mat}(n_y, n_x)$ obtained by putting the basis matrix $E_{i,j}$ (with all entries zero except for a one in position i, j) in the place correspond to the edge a and zero matrices elsewhere. Then, $E^a_{i,j}$ is a weight vector of weight $e_{y,i} - e_{x,j}$, and by varying a, i, j, we obtain a basis of weight vectors. Thus,

$$\Omega(\pi) = \left\{ e_{y,i} - e_{x,j} \;:\; i \in [n_y],\, j \in [n_x],\, a\colon x \to y \in Q_1 \right\}. \tag{6.3}$$

Clearly, the corresponding weight matrix equals the incidence matrix of a directed graph with vertex set $\{(x, j) : x \in Q_0, j \in [n_x]\} \cong [L]$, where we put an edge $(x, j) \to (y, i)$ whenever there exists an edge $a\colon x \to y \in Q_1$ and, if $x = y$, $i \neq j$. Using the before-mentioned fact we conclude that the weight matrix is unimodular. The bound on the weight margin then follows from Corollary 6.11. The statement about the weight norm is obvious from Eq. (6.3). □

Proposition 6.12 imply that the following three important families of representations have inverse polynomial weight margins.

Corollary 6.13. *The following three families of representations have totally unimodular weight matrix.*

1. *Simultaneous conjugation of $\mathrm{GL}(n)$ on $\mathrm{Mat}(n)^k$ (Example 1.7). Thus, $\gamma(\pi) \geqslant n^{-3/2}$.*
2. *Generalized Kronecker quiver actions of $\mathrm{GL}(n)^2$ on $\mathrm{Mat}(n)^k$ (Example 1.6). Thus, $\gamma(\pi) \geqslant (2n)^{-3/2}$.*
3. *The action of $\mathrm{GL}(n)^3$ on $\mathrm{Mat}(n)^2$ underlying Horn's problem (Example 1.2). Thus, $\gamma(\pi) \geqslant (3n)^{-3/2}$.*

Moreover, each of these representations are either trivial or have $N(\pi) = \sqrt{2}$.

For further illustration, we discuss the weight margin for operator and matrix scaling. The weights for the action of $\mathrm{GL}(n) \times \mathrm{GL}(n)$ from Example 6.3 are given by $\{e_i + e_{n+j} : i, j \in [n]\}$. To model operator and matrix scaling we need to restrict the group action to $G = \mathrm{SL}(n) \times \mathrm{SL}(n)$ and its maximal commutative subgroup $T_G = \mathrm{ST}(n) \times \mathrm{ST}(n)$, respectively. The set of weights is the same in both cases (since the very notion of a weight always refers to the maximal commutative subgroup) and can be obtained as the orthogonal projection of the weights for the $\mathrm{GL}(n) \times \mathrm{GL}(n)$-action onto the subspace of $\mathbb{R}^{2n}$ of vectors whose first n and last n components both sum to zero. Namely:

$$\Omega(\pi) = \left\{ e_i + e_{n+j} - \tfrac{1}{n} \textstyle\sum_{k=1}^{2n} e_k \;\middle|\; i, j \in [n] \right\}$$

We claim that the weight margin, which is thus likewise the same for matrix and for operator scaling, can be lower-bounded as follows:

$$\gamma(\pi) \geqslant n^{-3/2} \tag{6.4}$$

This bound can be obtained directly by modifying the proof of Lemma 6.10 analogously to how the proof of Theorem 6.8 is modified to obtain Theorem 6.9, but the bound for the commutative case was already present in the literature. Indeed, recall that in the commutative case we know that the weight margin is the largest constant $C > 0$ with the following property: If $\|\mu(v)\|_F < C$ then v is not in the null cone (cf. Remark 3.20). In [LSW98, Lemma 5.2] it is shown that, for matrix scaling, this statement holds for $C = n^{-3/2}$. Thus, Eq. (6.4) follows. For operator scaling, this bound appears in [Gur04a, Prop. 2.4], but in different language.

7 Explicit algorithms for $\mathrm{GL}(n)$ and $\mathrm{SL}(n)$

In this section we specialize our results and design concrete scaling algorithms for homogeneous polynomial actions of $\mathrm{GL}(n)$ and $\mathrm{SL}(n)$ with explicit running time bounds. Our algorithms take as input vectors in a Gelfand-Tsetlin basis, and we review its construction in Section 7.1. In Section 7.2 we prove upper bounds on the coefficients of invariant polynomials. We use these in Section 7.3 to obtain a priori lower bounds for capacities. The capacity lower bounds in turn are used in Section 7.4 to give explicit running time bounds for our algorithms.

7.1 Construction of the Gelfand-Tsetlin basis

Up to isomorphism, the irreducible representations $\pi_\lambda \colon \mathrm{GL}(n) \to \mathrm{GL}(V_\lambda)$ of $\mathrm{GL}(n)$ are labeled by their highest weight, which are integers vectors $\lambda \in \mathbb{Z}^n$ with $\lambda_1 \geqslant \cdots \geqslant \lambda_n$ (cf. Section 2.3).

If $\lambda_n \geqslant 0$ then λ can be identified with a *partition*. In this case, π_λ is a polynomial representation. That is, the matrix entries of $\pi_\lambda(g)$ in any basis of the representation space V_λ are homogeneous polynomial functions of degree $d = \sum_{i=1}^n \lambda_i$ in the matrix entries $g_{i,j}$ of the group element $g \in \mathrm{GL}(n)$. The restriction to polynomial irreducible representations is essentially without loss of generality. This is because $\pi_\lambda \otimes \det^k \cong \pi_{(\lambda_1+k,\ldots,\lambda_n+k)}$, which means that shifting the highest weight by the all-ones vectors amounts to tensoring with powers of the determinant (which itself is a one-dimensional irreducible representation of degree n). Moreover, the irreducible representation of $\mathrm{SL}(n)$ are parametrized by highest weights λ modulo this shift.

The *Gelfand-Tsetlin basis* is a particularly convenient basis of V_λ [Mol06]. Here, the action is given by rational functions with *rational* coefficients. Moreover, the group action, Lie algebra action, and moment map can be computed in polynomial time when working in this basis [Bür00, BCMW17]. This makes it ideally suited as an input format for our algorithms. We caution that the Gelfand-Tsetlin basis is in general *not* orthonormal with respect to the $\mathrm{U}(n)$-invariant inner product that we use throughout the paper (e.g., to define the capacity). Thus the $\mathrm{U}(n)$-invariant norm is *not* the same as the Euclidean norm in this basis (see Lemma 7.2 below).

We now show how, given λ, to construct the representation in the Gelfand-Tsetlin basis. The basis elements of the target vector space V_λ will be indexed by *patterns*, which are arrays $\Lambda = (\lambda_{i,j})$

satisfying some additional properties. They are often depicted as follows:

$$\begin{tikzcd}
\lambda_{n,1} \arrow[rr, phantom, "\geqslant"] \arrow[dr, phantom, sloped, "\geqslant"] & & \lambda_{n,2} \arrow[rr, phantom, "\geqslant"] \arrow[dr, phantom, sloped, "\geqslant"] & & \cdots \arrow[rr, phantom, "\geqslant"] \arrow[dr, phantom, sloped, "\geqslant"] & & \lambda_{n,n} \\
& \lambda_{n-1,1} \arrow[ur, phantom, sloped, "\geqslant"] \arrow[rr, phantom, "\geqslant"] \arrow[dr, phantom, sloped, "\geqslant"] & & \cdots \arrow[rr, phantom, "\geqslant"] & & \lambda_{n-1,n-1} \arrow[ur, phantom, sloped, "\geqslant"] & \\
& & \cdots \arrow[dr, phantom, sloped, "\geqslant"] & & \cdots \arrow[ur, phantom, sloped, "\geqslant"] & & \\
& & & \lambda_{1,1} \arrow[ur, phantom, sloped, "\geqslant"] & & &
\end{tikzcd}$$

Formally, we say $\Lambda = (\lambda_{i,j})_{i\in[n],j\in[i]}$ is a *pattern* associated with λ if

- The upper row coincides with λ, i.e., $\lambda_{n,j} = \lambda_j$ for $j \in [n]$, and
- The following *betweenness conditions* hold for $2 \leqslant i \leqslant n$ and $1 \leqslant j \leqslant i-1$:

  $$\lambda_{i,j} \geqslant \lambda_{i-1,j} \geqslant \lambda_{i,j+1}$$

  That is, Λ weakly decreases along both southeast and northeast diagonals.

Let $\mathcal{P}_\lambda$ denote the set of all patterns associated with λ. Now, define $V_\lambda = \mathbb{C}^{\mathcal{P}_\lambda}$, the vector space with basis vectors ξ_Λ labeled by the pattern $\Lambda \in \mathcal{P}_\lambda$. The dimension $m_\lambda := \dim V_\lambda$ is then $|\mathcal{P}_\lambda|$. We frequently identify V_λ with $\mathbb{C}^{m_\lambda}$ by putting the patterns in decreasing lexicographic order.

We now describe how to define the *Lie algebra* representation $\Pi_\Lambda \colon \operatorname{Mat}(n) \to L(V_\lambda)$, where we recall that $\operatorname{Mat}(n)$ is the Lie algebra of $GL(n)$. Recall that a Lie algebra representation is a linear map that satisfies $\Pi([X,Y]) = [\Pi(X),\Pi(Y)]$ for all $X, Y \in \operatorname{Mat}(n)$. By linearity, it is enough to define Π_λ on the basis matrices $E_{i,j} \in \operatorname{Mat}(n)$ which are all zeroes apart from the i, j entry, which is one. In fact, we need only define Π on $E_{i,i}$ for $i \in [n]$ as well as on $E_{i,i+1}$ and $E_{i+1,i}$ for $i \in [n-1]$. This is because any other $E_{i,j}$ can be obtained as an $(|i-j|-1)$-fold commutator of $E_{i,i+1}$'s or $E_{i+1,i}$'s. Theorem 2.3 of [Mol06] asserts that the following defines a Lie algebra representation of $\operatorname{Mat}(n)$:

$$\begin{aligned}
\Pi(E_{i,i})\xi_\Lambda &:= \left(\sum_{j=1}^{i} \lambda_{i,j} - \sum_{j=1}^{i-1} \lambda_{i-1,j}\right) \xi_\Lambda, \\
\Pi(E_{i,i+1})\xi_\Lambda &:= -\sum_{j=1}^{i} \frac{(l_{i,j}-l_{i+1,1})\cdots(l_{i,j}-l_{i+1,i+1})}{(l_{i,j}-l_{i,1})\cdots\vee\cdots(l_{i,j}-l_{i,i})} \xi_{\Lambda+\delta_{i,j}}, \\
\Pi(E_{i+1,i})\xi_\Lambda &:= \sum_{j=1}^{i} \frac{(l_{i,j}-l_{i-1,1})\cdots(l_{i,j}-l_{i-1,i-1})}{(l_{i,j}-l_{i,1})\cdots\vee\cdots(l_{i,j}-l_{i,i})} \xi_{\Lambda-\delta_{i,j}},
\end{aligned} \tag{7.1}$$

where $l_{i,j} := \lambda_{i,j} - j + 1$. The arrays $\Lambda \pm \delta_{i,j}$ are obtained from Λ by replacing $\lambda_{i,j}$ by $\lambda_{i,j} \pm 1$. The symbol $\vee$ indicates that the zero factor in the denominator is skipped. We set $\xi_\Lambda := 0$ if the array Λ is not a pattern associated with λ.

The representation $\pi_\lambda \colon GL(n) \to GL(V_\lambda)$ of the group $GL(n)$ can be defined by exponentiation, i.e., $\pi_\lambda(e^X) := e^{\Pi(X)}$ for $X \in \operatorname{Mat}(n)$. The basis $\{\xi_\Lambda\}_{\Lambda\in\mathcal{P}_\lambda}$ of V_λ is called the *Gelfand-Tsetlin basis* for π_λ. Then, π_λ is an irreducible representation of $GL(n)$ with highest weight λ and highest weight vector $\xi := \xi_\Lambda$, associated with the pattern $\Lambda_{i,j} := \lambda_j$. As mentioned earlier, the matrix entries of $\pi_\lambda(g)$ in the Gelfand-Tsetlin basis are rational functions with rational coefficients in the matrix entries of g. When λ is a partition, the matrix entries are in fact *polynomials* with rational coefficients. In Theorem 7.7 we prove explicit bounds on the coefficients. These will be obtained by lifting the following bounds on the Lie algebra representation of the basis matrices:

Lemma 7.1. *Let λ be a partition of d with at most n parts. Let $\Pi\colon \mathrm{GL}(n) \to \mathrm{GL}(m_\lambda)$ be the Lie algebra representation in the Gelfand-Tsetlin basis, where we identify $V_\lambda \cong \mathbb{C}^{m_\lambda}$ using the lexicographic order. Then $\Pi(E_{i,i})$ is diagonal for all $i \in [n]$. Moreover, there exist positive integers $\beta \leqslant R$, where $R = e^{O(n^3 \log(\lambda_1+n))}$, such that the entries of $\beta\Pi(E_{i,j})$ are integers of absolute value at most R for all $i, j \in [n]$.*

Proof. We first focus on $\Pi(E_{i,i+1})$, which is defined in Eq. (7.1). Define $\beta_{i,i+1} := \prod_{1\leqslant j\neq k\leqslant i} |l_{i,j} - l_{i,k}|$. Then, $\beta_{i,i+1}$ as well as all entries of $\beta_{i,i+1}\Pi(E_{i,i+1})$ are integers bounded in absolute value by $(\lambda_1+n)^{n^2}$. Next consider $\Pi(E_{i,j})$ for $j > i+1$. Note that $E_{i,j} = [E_{i,j-1}, E_{j-1,j}]$, which implies that $\Pi(E_{i,j}) = [\Pi(E_{i,j-1}), \Pi(E_{j-1,j})]$ because Π is a Lie algebra representation. We can thus write $\Pi(E_{i,j})$ as an iterated commutator of $\Pi(E_{i,i+1})$, $\Pi(E_{i+1,i+2})$, $\dots$, $\Pi(E_{j-1,j})$. In particular, this shows that $\Pi(E_{i,j})$ is strictly upper-triangular. As an iterated commutator, $\Pi(E_{i,j})$ is a sum of at most 2^n terms of the form $\Pi(E_{\sigma(i),\sigma(i)+1})\cdots\Pi(E_{\sigma(j-1),\sigma(j-1)+1})$, where σ is a permutation of $\{i, i+1\dots, j-1\}$. Because the patterns appearing with nonzero coefficient in $\Pi(E_{k,k+1})\xi_\Lambda$ differ from ξ_Λ in the k^{th} row of Λ and nowhere else, the coefficient of each pattern appearing in

$$\Pi(E_{\sigma(i),\sigma(i)+1})\cdots\Pi(E_{\sigma(j-1),\sigma(j-1)+1})\xi_\Lambda$$

may be written as a product of entries of $\Pi(E_{k,k+1})$ for $k = i, \dots, j-1$. Define $\beta := \beta_{1,2}\cdots\beta_{n-1,n}$. Then, β is a common denominator of $\Pi(E_{i,j})$ for all $j > i$. Moreover, β and all entries of

$$\beta\Pi(E_{\sigma(i),\sigma(i)+1})\cdots\Pi(E_{\sigma(j-1),\sigma(j-1)+1})$$

are integers bounded in absolute value by $(\lambda_1+n)^{n^2(n-1)}$. It follows that, for all $j > i$, $\beta\Pi(E_{i,j})$ is an integer matrix with entries bounded in absolute value by $R := 2^n(\lambda_1+n)^{n^3} = e^{O(n^3\log(\lambda_1+n))}$. A completely analogous argument establishes the same bound when $i > j$.

Finally, consider $\Pi(E_{i,i})$. Using Eq. (7.1) and the betweenness conditions, each entry of $\Pi(E_{i,i})$ can be upper bounded by $\lambda_{i,1} \leqslant \lambda_1$. Thus, for all i, the entries of $\beta\Pi(E_{i,i})$ are certainly bounded by R. This concludes the proof. □

An important object in our setup is the $U(n)$-invariant inner product $\langle\cdot,\cdot\rangle$ on V_λ, which is unique up to a positive scalar. We will fix it by demanding that $\|\xi\| = 1$. Proposition 2.4 of [Mol06] computes this inner product explicitly:

$$\langle\xi_\Lambda, \xi_{\Lambda'}\rangle = \delta_{\Lambda,\Lambda'} \prod_{k=2}^{n} \prod_{1\leqslant i\leqslant j<k} \frac{(l_{k,i}-l_{k-1,j})!}{(l_{k-1,i}-l_{k-1,j})!} \prod_{1\leqslant i<j\leqslant k} \frac{(l_{k,i}-l_{k,j}-1)!}{(l_{k-1,i}-l_{k,j}-1)!}. \tag{7.2}$$

Thus, the Gelfand-Tsetlin basis is orthogonal, but *not* necessarily orthonormal. This means that the $U(n)$-invariant norm $\|\cdot\|$ (which underlies our basic setup and computational problems) need not be the same as the Euclidean norm $\|\cdot\|_2$ on $V_\lambda = \mathbb{C}^{\mathcal{P}_\lambda}$ (which is more directly related to the input size of our explicit algorithms in this section). The following lemma compares the two norms when λ is a partition.

Lemma 7.2. *Let λ be a partition of $d > 0$ with at most n parts, and let $v \in V_\lambda = \mathbb{C}^{\mathcal{P}_\lambda}$. Then,*

$$\|v\|_2 \leqslant \|v\| \leqslant e^{nd\log(nd)}\|v\|_2,$$

where $\|\cdot\|_2$ denotes the Euclidean norm and $\|v\|$ denotes the $U(n)$-invariant norm.

Proof. It is enough to show that $1 \leqslant \|\xi_\Lambda\| \leqslant e^{O(nd\log(nd)}$ for any $\Lambda \in \mathcal{P}$. The lower bound follows from Eq. (7.2) using the betweenness conditions. The upper bound can be obtained using Eq. (7.2) as follows. We first bound the left-hand product:

$$\begin{aligned}
&\prod_{k=2}^{n}\prod_{1\leqslant i\leqslant j<k}\frac{(l_{k,i}-l_{k-1,j})!}{(l_{k-1,i}-l_{k-1,j})!} \leqslant \prod_{k=2}^{n}\prod_{1\leqslant i\leqslant j<k}(l_{k,i}-l_{k-1,j})^{l_{k,i}-l_{k-1,i}} \\
&\leqslant \prod_{k=2}^{n}\prod_{1\leqslant i\leqslant j<k}(\lambda_1+n)^{l_{k,i}-l_{k-1,i}} = \prod_{1\leqslant i\leqslant j<n}(\lambda_1+n)^{l_{n,i}-l_{j,i}} \\
&= \prod_{1\leqslant i\leqslant j<n}(\lambda_1+n)^{\lambda_{n,i}-\lambda_{j,i}} \leqslant \prod_{1\leqslant i\leqslant j<n}(\lambda_1+n)^{\lambda_{n,i}} \\
&\leqslant \prod_{1\leqslant i<n}(\lambda_1+n)^{n\lambda_{n,i}} \leqslant (\lambda_1+n)^{nd} \leqslant e^{nd\log(nd)}.
\end{aligned}$$

We can similarly bound the right-hand product:

$$\begin{aligned}
&\prod_{k=2}^{n}\prod_{1\leqslant i<j\leqslant k}\frac{(l_{k,i}-l_{k,j}-1)!}{(l_{k-1,i}-l_{k,j}-1)!} \leqslant \prod_{k=2}^{n}\prod_{1\leqslant i<j\leqslant k}(l_{k,i}-l_{k,j}-1)^{l_{k,i}-l_{k-1,i}} \\
&\leqslant \prod_{k=2}^{n}\prod_{1\leqslant i<j\leqslant k}(\lambda_1+n)^{l_{k,i}-l_{k-1,i}} \leqslant e^{nd\log(nd)},
\end{aligned}$$

where the last inequality follows as above. Together, we obtain that $\|\xi_\Lambda\| \leqslant e^{nd\log(nd)}$. □

We will say that a representation $\pi\colon \mathrm{GL}(n) \to \mathrm{GL}(m)$ is *given in a Gelfand-Tsetlin basis* if π is a direct sum of irreducible representations each of which is given in a Gelfand-Tsetlin basis as defined above. That is, $\pi = \oplus_{i=1}^s \pi_{\lambda^i}$ and we identify $V_\lambda = \bigoplus_{i=1}^s V_{\lambda^i} \cong \mathbb{C}^m$ by using the Gelfand-Tsetlin basis in each irreducible summand. Up to isomorphism, any finite-dimensional representation of $\mathrm{GL}(n)$ is of this form. If we demand that the λ_i are partitions then π is polynomial, and any finite-dimensional polynomial representation of $\mathrm{GL}(n)$ is of this form. Any finite-dimensional representation of $\mathrm{SL}(n)$ can be obtained by restricting such a π. We will say that a representation of $\mathrm{SL}(n)$ is *given in a Gelfand-Tsetlin basis* if it is obtained in this way. It is clear that Lemmas 7.1 and 7.2 extend naturally to such representations. Our algorithms will take their input in a Gelfand-Tsetlin basis.

7.2 Coefficient upper bounds

Next, we prove upper bounds on the coefficients of the polynomials defining the matrix entries of a representation. This is a basis-dependent notion, so we assume that $V = \mathbb{C}^m$, corresponding to a choice of (not necessarily orthonormal) basis. We first introduce some notation.

Definition 7.3 (Degree, homogeneous, coefficient size)**.** *Let* $\pi\colon \mathrm{GL}(n) \to \mathrm{GL}(m)$ *be a* polynomial *representation, i.e., all matrix entries* $\pi_{k,l}(g) = \langle e_k, \pi(g)e_l\rangle$ *of the representation are polynomials in the matrix entries of the group element* $g \in \mathrm{GL}(n)$.

- *We define the* degree $d(\pi)$ *of* π *as the maximal degree of the polynomials* $\pi_{k,l}$.
- *We say that* π *is* homogeneous *if all nonzero* $\pi_{k,l}$ *are homogeneous of the same degree, namely* $d(\pi)$.
- *We define the* coefficient size *of* π *as the least positive integer* $R = R(\pi)$ *such that there is* $\alpha \in [R]$ *such that all* $\alpha\pi_{k,l}$ *have integer coefficients of absolute value at most* R *(that is,* R *bounds the numerators and common denominator of all coefficients for all entries). Set* $R(\pi) := \infty$ *if not all coefficients are rational.*

The following lemma is useful to bound coefficient sizes.

Lemma 7.4. *Let* $p_1, \ldots, p_\ell$ *be nonzero polynomials in variables* $X_1, \ldots, X_N$*. Suppose each* p_j *has degree* d_j *and integer coefficients bounded in absolute value by* R_j*. Then* $p := \prod_{i=1}^{\ell} p_i$ *is a polynomial of degree* $d = d_1 + \cdots + d_\ell$ *and has integer coefficients bounded in absolute value by*

$$R_1 \cdots R_\ell \min\left\{ \left(1 + \frac{d}{N}\right)^{N(\ell-1)} ,\ \ell^d \right\}.$$

Proof. It is clear that the degree of p is equal to $d_1 + \cdots + d_\ell$. To see the bound on its coefficients, write p and p_j as sums of monomials, say, $p = \sum_\omega p_\omega X^\omega$ and $p_j = \sum_\omega p_\omega^{(j)} X^\omega$. Then, using that the coefficients of p_j are bounded in absolute value by R_j,

$$\begin{aligned} |p_\omega| &= \left| \sum_{\omega^{(1)}+\ldots+\omega^{(\ell)}=\omega} p^{(1)}_{\omega^{(1)}} \cdots p^{(\ell)}_{\omega^{(\ell)}} \right| \leqslant R_1 \cdots R_\ell \left| \sum_{\omega^{(1)}+\ldots+\omega^{(\ell)}=\omega} 1 \right| \\ &= R_1 \cdots R_\ell \prod_{i=1}^{N} \binom{\omega_i + \ell - 1}{\omega_i} \leqslant R_1 \cdots R_\ell \prod_{i=1}^{N} \min\left\{ (\omega_i + 1)^{\ell-1}, \ell^{\omega_i} \right\}, \end{aligned}$$

where the last inequality follows from the bound $\binom{k+\ell-1}{k} \leqslant \min\{(k+1)^{\ell-1}, \ell^k\}$. To finish, note that $\prod_{i=1}^N \ell^{\omega_i} = \ell^{\sum_{i=1}^N \omega_i} = \ell^d$ and, by the AM-GM inequality, $\prod_{i=1}^N (\omega_i + 1) \leqslant (\sum_{i=1}^N (\omega_i + 1)/N)^N = (1 + d/N)^N$. Thus, the coefficients of p are bounded in absolute value as claimed. □

Next, we show how coefficient sizes behave under direct sums and tensor products.

Lemma 7.5 (Direct sums and tensor products)**.** *Let* $\rho_1\colon \mathrm{GL}(n) \to \mathrm{GL}(m_1)$*,* $\rho_2\colon \mathrm{GL}(n) \to \mathrm{GL}(m_2)$*, and* $\rho\colon \mathrm{GL}(n) \to \mathrm{GL}(m)$ *be polynomial representations. Then:*

1. *The direct sum representation* $\pi := \rho_1 \oplus \rho_2\colon \mathrm{GL}(n) \to \mathrm{GL}(m_1 + m_2)$ *has degree* $d(\pi) = d(\rho_1)d(\rho_2)$ *and coefficient size* $R(\pi) \leqslant R(\rho_1)R(\rho_2)$.
2. *The tensor product representation* $\pi := \rho_1 \otimes \rho_2\colon \mathrm{GL}(n) \to \mathrm{GL}(m_1 m_2)$ *has degree* $d(\pi) = d(\rho_1) + d(\rho_2)$ *and coefficient size*

$$R(\rho_1 \otimes \rho_2) \leqslant R(\rho_1)R(\rho_2) \min\left\{ \left(1 + \frac{d(\rho_1 \otimes \rho_2)}{n^2}\right)^{n^2} ,\ 2^{d(\rho_1 \otimes \rho_2)} \right\}.$$

3. *For $\ell \in \mathbb{N}$, the tensor power representation $\pi := \rho^{\otimes \ell} = \rho \otimes \cdots \otimes \rho \colon \mathrm{GL}(n) \to \mathrm{GL}(m^\ell)$ has degree $d(\pi) = \ell d(\rho)$ and coefficient size $R(\pi) \leqslant R(\rho)^\ell \ell^{\ell d(\rho)}$.*

Proof. The first statement is clear. The other two follow from Lemma 7.4 with $N = n^2$. □

We now aim to bound the coefficient size for representations in a Gelfand-Tsetlin basis. Our main tool is the following result which allows lifting coefficient bounds from Lie algebra to Lie group representations.

Proposition 7.6 (Lifting coefficient bounds)**.** *Let $\pi \colon \mathrm{GL}(n) \to \mathrm{GL}(m)$ be a homogeneous polynomial representation of degree $d > 0$. Let $\Pi \colon \mathrm{Mat}(n) \to \mathrm{Mat}(m)$ denote its Lie algebra representation and suppose $\Pi(E_{i,i})$ is diagonal for all $i \in [n]$. Let $\beta \leqslant R$ be positive integers such that the entries of $\beta\Pi(E_{i,j})$ are integers of absolute value at most R for all $i, j \in [n]$. Then, the entries of $\pi(g)$ are polynomials with rational coefficients in the entries of $g \in \mathrm{GL}(n)$, and $R(\pi) = R^{2m} e^{O(mn^3 d \log(mnd))}$.*

Proposition 7.6 is proved in Appendix A. As a consequence, we obtain the following fundamental bound on the coefficient size of representations in a Gelfand-Tsetlin basis.

Theorem 7.7 (Coefficient size in Gelfand-Tsetlin basis)**.** *Let $\pi \colon \mathrm{GL}(n) \to \mathrm{GL}(m)$ be a polynomial representation of degree $d > 0$ given in a Gelfand-Tsetlin basis. Then, the entries of $\pi(g)$ are polynomials with rational coefficients in the entries of $g \in \mathrm{GL}(n)$, and $R(\pi) = e^{O(mn^3 d \log(mnd))}$.*

Proof. We have, $\pi = \oplus_{i=1}^s \pi_{\lambda_i}$, where each π_{λ_i} is a homogeneous polynomial representation of degree at most d in a Gelfand-Tsetlin basis. Let $m_i := m_{\lambda_i}$ denote the dimension of π_{λ_i}. By Lemma 7.1 and Proposition 7.6, we have $R(\pi_{\lambda_i}) = e^{O(m_i n^3 d \log(m_i n d))}$. Thus, by Item 1 of Lemma 7.5, we have $R(\pi) \leqslant \prod_{i=1}^s R(\pi_{\lambda_i}) = e^{O(mn^3 d \log(mnd))}$, where we used that $m = \sum_{i=1}^s m_i$. □

7.3 Capacity lower bounds

We now prove capacity lower bounds for vectors of bounded bit complexity in a representation. Apart from the coefficient bounds from Section 7.2, we also need upper bounds on the degree and coefficients of invariant polynomials. As motivated in Sections 1.4 and 1.6, it is convenient to work with the subgroup $\mathrm{SL}(n)$ of $\mathrm{GL}(n)$. We shall say that a representation $\pi \colon \mathrm{SL}(n) \to \mathrm{GL}(V)$ is *polynomial* of degree $d(\pi)$ and coefficient size $R(\pi)$ if it is the restriction of a polynomial representation of $\mathrm{GL}(n)$ with this degree and coefficient size. We first borrow the following version of Derksen's general degree bound [Der01, Proposition 2.1].

Proposition 7.8 (Degree bound, [Der01])**.** *Let $\pi \colon \mathrm{SL}(n) \to \mathrm{GL}(V)$ be a polynomial representation of degree d and let $v \in V$ such that $\mathrm{cap}(v) > 0$ (i.e., v is not in the null cone). Then there exists a nonconstant homogeneous invariant polynomial p on V of degree at most nd^{n^2-1} such that $p(v) \neq 0$.*

Next we recall the following coefficient bound for invariant polynomials [BGO+17, Theorem 7.10].

Proposition 7.9 (Coefficient bound for $\mathrm{SL}(n)$-invariants, [BGO+17])**.** *Let $\pi \colon \mathrm{SL}(n) \to \mathrm{GL}(m)$ be a homogeneous polynomial representation of degree $d > 0$ and coefficient size $R = R(\pi)$, and let $D > 0$. Then the space of invariant polynomials on V of degree D is spanned by invariant polynomials with integer coefficients bounded in absolute value by $e^{O(D \log(Rm) + dD \log(Ddn))}$.*

As a consequence, we obtain the following lower bound on the ℓ_2-norm of arbitrary scalings in terms of the coefficient size of the representation. The proof is very similar to the one of [BGO+17, Theorem 7.12], but our analysis gives a slightly better dependence on d.

Proposition 7.10 (ℓ_2-norm lower bound). *Let $\pi\colon \mathrm{SL}(n) \to \mathrm{GL}(m)$ be a homogeneous polynomial representation of degree $d > 0$ and coefficient size $R = R(\pi)$. If $v \in \mathbb{Z}[i]^m$ is a vector with $\operatorname{cap}(v) > 0$,*

$$-\inf_{g \in \mathrm{SL}(n)} \log \|\pi(g)v\|_2 = O\big(\log(Rm) + d\log(n) + n^2 d \log(d)\big).$$

Proof. According to Proposition 7.8, there exists a homogeneous invariant polynomial p on V of degree $1 \leqslant D \leqslant nd^{n^2-1}$. By Proposition 7.9, we may assume that p has integer coefficients bounded in absolute value by $L = e^{O(D\log(Rm)+dD\log(Ddn))}$. Then,

$$1 \leqslant |p(v)| = |p(\pi(g)v)| \leqslant m^D L \|\pi(g)v\|_\infty^D \leqslant m^D L \|\pi(g)v\|_2^D,$$

where we first used that $p(v) \in \mathbb{Z}[i]$, then the invariance of p, and finally that p is a sum of at most $\binom{D+m-1}{M-1} \leqslant m^D$ monomials with coefficients of bounded in absolute value by L. As a consequence,

$$\|\pi(g)v\|_2^{-1} \leqslant mL^{1/D} = e^{O(\log(Rm)+d\log(n)+n^2 d\log(d))},$$

which gives the desired bound. □

When using a Gelfand-Tsetlin basis, we can compare the ℓ^2-norm with the $U(n)$-invariant norm by using Lemma 7.2. Together with Proposition 7.10 and the bound on $R(\pi)$ from Theorem 7.7 we thus obtain the following fundamental capacity lower bound for $\mathrm{SL}(n)$-actions. As discussed in Section 1.6, the restriction to $\mathrm{SL}(n)$ is without loss of generality for homogeneous representations.

Corollary 7.11 (Capacity lower bound). *Let $\pi\colon \mathrm{SL}(n) \to \mathrm{GL}(m)$ be a homogeneous polynomial representation of degree $d > 0$ given in a Gelfand-Tsetlin basis. Let $v \in \mathbb{Z}[i]^m$ be a vector such that $\operatorname{cap}(v) > 0$. Then,*

$$-\log \operatorname{cap}(v) = O\big(mn^3 d \log(mnd)\big).$$

Proof. By Lemma 7.2, any lower bound on the ℓ_2-norm is also a lower bound on the $U(n)$-invariant norm. Thus, Proposition 7.10 and the bound on $R(\pi)$ from Theorem 7.7 show the desired bound. □

We now prove an analogous result for p-capacities associated with $\mathrm{GL}(n)$-actions. As discussed in Sections 1, 2.3 and 3.6, the moment polytope $\Delta(v)$ is naturally a subset of $C(n) = \{p \in \mathbb{R}^n : p_1 \geqslant \cdots \geqslant p_n\}$, see Table 2.1. In fact, when $\pi\colon \mathrm{GL}(n) \to \mathrm{GL}(V)$ is a homogeneous polynomial representation of degree d then $\Delta(v)$ is necessarily contained in

$$\Delta_d(n) := \left\{ p \in \mathbb{R}^n \;:\; p_1 \geqslant \cdots \geqslant p_n \geqslant 0, \sum_{i=1}^n p_i = d \right\} \subseteq C(n),$$

and so we only consider target points $p \in \Delta_d(n)$.

Theorem 7.12 (p-capacity lower bound). *Let $\pi\colon \mathrm{GL}(n) \to \mathrm{GL}(m)$ be a homogeneous polynomial representation of degree $d > 0$ given in a Gelfand-Tsetlin basis. Let $p \in \mathbb{Q}^n \cap \Delta_d(n)$ and let ℓ be a positive integer such that $\ell p \in \mathbb{Z}^n$. If $v \in \mathbb{Z}[i]^m$ is a vector such that $\operatorname{cap}_p(v) > 0$ then*

$$-\log \operatorname{cap}_p(v) = O\big(mn^3 d \log(\ell m n d)\big).$$

Proof. Set $\lambda := \ell p$. By Eq. (3.20), $\operatorname{cap}_p(v)^\ell = \operatorname{cap}(v^{\otimes \ell} \otimes v_{\lambda^*})$, where the right-hand side capacity is computed in the $\mathrm{GL}(n)$-representation $\mathrm{Sym}^\ell(\mathbb{C}^m) \otimes V_{\lambda^*}$. The latter representation has degree zero, so the capacity does not change when taken over $\mathrm{SL}(n)$ rather than $\mathrm{GL}(n)$. This in turn allows us to replace λ^* by $\mu := \lambda^* + (\lambda_1, \dots, \lambda_1)$, since shifting by multiples of the all-ones vector does not change the representation with respect to $\mathrm{SL}(n)$. Since μ is a partition, it corresponds to a homogeneous polynomial representation of degree $r := n\lambda_1 - \ell d$. Thus, $\operatorname{cap}_p(v)^\ell = \operatorname{cap}(v^{\otimes \ell} \otimes v_\mu)$, where the right-hand side capacity is computed in the $\mathrm{SL}(n)$-representation $\mathrm{Sym}^\ell(\mathbb{C}^m) \otimes V_\mu$.

We could now work in the Gelfand-Tsetlin basis of V_μ and apply Corollary 7.11, but the resulting capacity lower bounds scale poorly with ℓ. Instead, we shall realize π_μ using Weyl's construction (see Chapter 6 of [FH13]) as a subrepresentation of $\tau^{\otimes r}\colon \mathrm{SL}(n) \to \mathrm{GL}((\mathbb{C}^n)^{\otimes r})$, where $\tau\colon \mathrm{SL}(n) \to \mathrm{GL}(n)$ is the defining representation. We can use the Euclidean norm as the unitarily invariant norm $\|\cdot\|$ on $(\mathbb{C}^n)^{\otimes r}$. A unit norm highest weight vector of π_μ is then given by $v_\mu := w_\mu / \|w_\mu\|$, where

$$w_\mu := \bigotimes_{i=1}^{n} (e_1 \wedge \cdots \wedge e_i)^{\otimes(\mu_i - \mu_{i+1})} \in (\mathbb{C}^n)^{\otimes r}.$$

Here, $\wedge$ means to antisymmetrize *without* averaging, so that w_μ is a vector with integer coefficients in the standard tensor product basis. Note that $\|w_\mu\| \leqslant \sqrt{n^r}$ because w_μ is a linear combination of some subset of standard basis vectors with coefficients ± 1. We now have

$$\log \operatorname{cap}_p(v) = \frac{1}{\ell} \log \operatorname{cap}\left(v^{\otimes \ell} \otimes \frac{w_\mu}{\|w_\mu\|}\right) \geqslant \frac{1}{\ell} \log \operatorname{cap}(v^{\otimes \ell} \otimes w_\mu) - \frac{r}{2\ell} \log n. \tag{7.3}$$

Clearly, the capacity is unchanged when computed in the larger representation $\rho := \pi^{\otimes \ell} \otimes \tau^{\otimes r}$ on $(\mathbb{C}^m)^{\otimes \ell} \otimes (\mathbb{C}^n)^{\otimes r}$. To lower-bound it, note that $v^{\otimes \ell} \otimes w_\mu$ is a Gaussian integer vector in the standard tensor product basis of the latter space, which has dimension $m^\ell n^r$. Next, note that ρ has degree $d(\rho) = \ell d + r = n\lambda_1 \leqslant \ell n d$. Its coefficient size can be upper-bounded as follows:

$$R(\rho) \leqslant R(\pi^{\otimes \ell}) R(\tau^{\otimes r}) 2^{n\lambda_1} \leqslant R(\pi)^\ell \ell^{\ell d} R(\tau^{\otimes r}) 2^{n\lambda_1} = e^{O(\ell m n^3 d \log(\ell m n d))},$$

where we first used Item 2 and then Item 3 of Lemma 7.5; the last inequality follows from $R(\tau^{\otimes r}) = 1$ (which holds by direct inspection) and the bound from Theorem 7.7. Finally, by Lemma 7.2 and since we use the Euclidean norm as the unitarily invariant norm on $(\mathbb{C}^n)^{\otimes r}$, we obtain the following capacity lower bound from Proposition 7.10:

$$-\log \operatorname{cap}(v^{\otimes \ell} \otimes w_\mu) = O\big(\ell m n^3 d \log(\ell m n d) + \log(m^\ell n^r) + \ell n d \log(n) + n^2 \ell n d \log(\ell n d)\big)$$

Combining the above with Eq. (7.3) and using $r \leqslant \ell n d$ yields the desired bound. □

We discussed in Section 3.6 that $p \in \Delta(v)$ if and only if $\operatorname{cap}_p(\pi(g)v) > 0$ for generic $g \in G$, see Eq. (3.21). This motivates proving a p-capacity lower bound for random elements in the orbit of v. To start, we need an effective version of the equivalence statement. Such a result appeared in [BFG+18] for the tensor action. The proof extends to the more general setting by applying Derksen's degree bound in greater generality.

Proposition 7.13 (Effective version of Mumford's theorem). *Let $\pi\colon \mathrm{GL}(n) \to \mathrm{GL}(V)$ be a homogeneous polynomial representation of degree $d > 0$ and let $v \in V$. Let $p \in \mathbb{Q}^n \cap \Delta_d(n)$ and let ℓ be a positive integer such that $\ell p \in \mathbb{Z}^n$. Then the set of group elements $g \in \mathrm{GL}(n)$ such that $\operatorname{cap}_p(\pi(g)v) = 0$ is (as a subset of $\mathrm{GL}(n)$) defined by the zero set of polynomials of degree at most $(\ell n d)^{n^2}$ in the matrix entries $g_{i,j}$.*

Proof. Recall from the proof of Theorem 7.12 that $\operatorname{cap}_p(\nu)^\ell = \operatorname{cap}(\nu^{\otimes \ell} \otimes \nu_\mu)$, where the right-hand side capacity is computed in the $\mathrm{SL}(n)$-representation on $W := \mathrm{Sym}^\ell(\mathbb{C}^m) \otimes V_\mu$, with $\mu := \lambda^* + (\lambda_1, \ldots, \lambda_1)$. Note that the latter representation is polynomial of degree $\lambda_1 n \leqslant \ell n d$. Thus, it follows from Proposition 7.8 that there exist homogeneous $\mathrm{SL}(n)$-invariant polynomials p_k on W of degree at most $D := n(\ell n d)^{n^2-1}$ such that $\operatorname{cap}_p(\nu) = 0$ if and only if $p_k(\nu^{\otimes \ell} \otimes \nu_\mu) = 0$ for all k. Thus, $\operatorname{cap}_p(\pi(g)\nu) = 0$ if and only if $\tilde{p}_k(g) = 0$ for all k, where $\tilde{p}_k(g) := p_k((\pi(g)\nu)^{\otimes \ell} \otimes \nu_\mu)$ is a polynomial of degree at most $d\ell D = (\ell n d)^{n^2}$ in the matrix entries $g_{i,j}$. □

Lastly, we need to constrain the bit complexity of the vertices of the moment polytopes.

Proposition 7.14. *Let* $\pi\colon \mathrm{GL}(n) \to \mathrm{GL}(V)$ *be a homogeneous polynomial representation of degree* $d > 0$ *and let* $v \in V$. *Then, every vertex* q *of* $\Delta(v)$ *is rational and* $\ell q \in \mathbb{Z}^n$ *for some integer* $1 \leqslant \ell \leqslant n^{3n/2} d^{n^2-n}$.

Proof. The moment polytope $\Delta(\nu)$ can be written as the intersection of the positive Weyl chamber $C(n)$ (see Table 2.1) and finitely many polytopes Δ_i with vertices in $\Omega(\pi)$ [Fra02]. Note that any $x \in \Omega(\pi)$ is integral and satisfies $\|x\|_2 \leqslant d$ by Proposition 3.10 and Lemma 6.1.

We claim that $\Delta' := \bigcap_i \Delta_i$ can be defined as the intersection of halfspaces bounded by hyperplanes passing through n affinely independent points $x \in \mathbb{Z}^n$ with $\|x\|_2 \leqslant d$. Clearly, it is enough to show that each Δ_i has this property. Indeed, if Δ_i has maximal dimension, the hyperplanes spanning its faces suffice. Otherwise, if Δ_i is of dimension $k < n$, there is a set S of at most $n-k$ points from $\{e_1, \ldots, e_n\}$ so that $\Delta_i' := \operatorname{conv}(\Delta_i \cup S)$ is of full dimension. Then Δ_i is a face of Δ_i', and so a subset of hyperplanes defining Δ_i' defines Δ_i.

Next, we claim that $\Delta(\nu) = C(n) \cap \Delta'$ can be defined as the intersection of halfspaces bounded by affine hyperplanes of the form $\{x \in \mathbb{R}^n : x \cdot y = b\}$, where $b \in \mathbb{Z}$, $y \in \mathbb{Z}^n$, and $\|y\|_\infty \leqslant M := n d^{n-1}$. Clearly, this holds for the halfspaces defining $C(n)$, so it remains to show the same for Δ'. Thus, it suffices to prove that any hyperplane passing through n affinely independent points $x \in \mathbb{Z}^n$ with $\|x\|_2 \leqslant d$ can be written in the form above. The following standard argument found in [Sch86, §17.1] shows that this is indeed the case. Let A denote the matrix whose rows are the n points spanning the hyperplane. If 0 is not in the hyperplane, A is invertible. By Cramer's rule, the unique solution to the equation $Ay = \det(A)\mathbf{1}$ is given by $y_i = \det(A_i)$, where A_i is the matrix with the i^{th} column replaced by $\mathbf{1}$. The vector y is the desired hyperplane normal and $b = \det(A)$. Expanding down the i^{th} column of A_i, using Hadamard's bound for each of the n minors, we find $|y_i| \leqslant M$. If 0 is in the hyperplane, then $\det(A) = 0$, so we can take y to be any nonzero column of the adjugate matrix, which obeys the same bound. This proves the second claim.

We now apply a similar argument to bound the complexity of the vertices. Every vertex q of $\Delta(\nu)$ is the intersection of some n of these hyperplanes $\{x \in \mathbb{R}^n : x \cdot y_i = b_i\}$ with linearly independent normal vectors $y_1, \ldots, y_n$ that satisfy $\|y_i\|_2 \leqslant n^{1/2}\|y_i\|_\infty \leqslant n^{1/2}M$. We may apply the argument of [Sch86] again to see that $\ell := |\det(y_1, \ldots, y_n)|$ satisfies $\ell q \in \mathbb{Z}^n$ and

$$1 \leqslant \ell \leqslant \|y_1\|_2 \cdots \|y_n\|_2 \leqslant (n^{1/2}M)^n = n^{3n/2} d^{n^2-n}.$$ □

Finally, we obtain our capacity lower bound for a random element in the orbit of ν.

Theorem 7.15 (Randomized p-capacity lower bound)**.** *Let* $\pi\colon \mathrm{GL}(n) \to \mathrm{GL}(m)$ *be a homogeneous polynomial representation of degree* $d > 0$ *given in a Gelfand-Tsetlin basis,* $\nu \in \mathbb{Z}[i]^m$ *a vector, and* $p \in \Delta(\nu)$

a point in its moment polytope. Set $S := 4n^{3n^3+1}d^{n^4}$*. If* $g \in \mathrm{Mat}(n)$ *is an integer matrix with entries drawn i.i.d. uniformly at random from* $[S]$*, then with probability at least* $1/2$ *we have* $g \in \mathrm{GL}(n)$ *and*

$$-\log \mathrm{cap}_p(\pi(g)v) = O\big(mn^5 d \log(mnd)\big).$$

Proof. Recall that $\Delta(v) \subseteq \Delta_d(n)$ is a convex polytope of dimension at most $n-1$. Thus, by Caratheodory's theorem, we know that p is contained in the convex hull of some set Q of at most n vertices of $\Delta(v)$. For every $q \in Q$, Proposition 7.14 shows that there exists a positive integer $\ell_q \leqslant n^{3n/2}d^{n^2-n}$ such that $\ell_q q \in \mathbb{Z}^n$. Applying Proposition 7.13 with $\ell = \ell_q$, and using that $q \in \Delta(v)$, there is a polynomial f_q of degree d_q at most $(\ell_q nd)^{n^2} \leqslant n^{3n^3}d^{n^4}$ that is not identically zero but that vanishes on all $g \in \mathrm{GL}(n)$ such that $\mathrm{cap}_q(\pi(g)v) = 0$.

Now observe that our choice of S ensures that $S \geqslant \max(4nd_q, 4n)$ for all $q \in Q$. If $g \in \mathrm{Mat}(n)$ is an integer matrix with entries drawn i.i.d. uniformly at random from $[S]$, then, by the Schwarz-Zippel lemma, $\Pr(f_q(g) = 0) \leqslant d_q/S \leqslant 1/4n$ for each $q \in Q$. Furthermore, $\Pr(\det(g) = 0) \leqslant 1/4$. By the union bound, it follows that, with probabilty at least $1/2$, we have $\det(g) \neq 0$ (i.e., $g \in \mathrm{GL}(n)$) and $f_q(g) \neq 0$ for all $q \in Q$. The latter implies that $\mathrm{cap}_q(\pi(g)v) \neq 0$. Since v has Gaussian integer entries and g is an integer matrix, there exists a positive integer $\alpha \leqslant R(\pi)$ such that $\alpha\pi(g)v$ has Gaussian integer entries. Applying Theorems 7.7 and 7.12, we obtain

$$\begin{aligned}-\log \mathrm{cap}_p(\pi(g)v) &= -\log \mathrm{cap}_p(\alpha\pi(g)v) + \log \alpha \\ &= O\big(mn^3 d \log(\ell_q mnd) + mn^3 d \log(mnd)\big) \\ &= O\big(mn^5 d \log(mnd)\big)\end{aligned}$$

for all $q \in Q$. To complete the proof, observe that by concavity of $p \mapsto \log \mathrm{cap}_p(\pi(g)v)$ (Proposition 3.32), the quantity $\log \mathrm{cap}_p(\pi(g)v)$ is at least $\min_{q \in Q} \log \mathrm{cap}_q(\pi(g)v)$. □

7.4 Explicit running time bounds

In this section, we apply the capacity lower bound of Section 7.3 to bound the running time of our algorithms on inputs specified as in Section 1.6. Firstly, let us discuss the issue of precision.

Remark 7.16 (Precision). *Each step of Algorithm 4.2 can be applied in time polynomial in* n*,* m*, and the desired number of bits of precision by [BCMW17]. It is not hard to verify that if* T *is the desired number of iterations then there is a* $\mathrm{poly}(n, m, T)$ *number of bits of precision such that Algorithm 4.2 with each step calculated to that precision still satisfies Theorem 4.2; see the discussion on bit-complexity in, e.g., [BFG+18, Roe18]. The same holds for Algorithm 5.1 and Theorem 5.6.*

We now restate and prove our explicit bound on the running time of our first order algorithm for the scaling problem for representations of $\mathrm{SL}(n)$.

Theorem 1.23 (First order algorithm for scaling in terms of input size)**.** *Let* (π, v, ε) *be an instance of the scaling problem for* $\mathrm{SL}(n)$ *such that* $0 \in \Delta(v)$ *and every entry of* v *is bounded in absolute value by* M*. Let* d *denote the degree and* m *the dimension of* π*. Then, Algorithm 4.2 with a number of iterations at most*

$$T = O\left(\frac{d^3}{\varepsilon^2} mn^3 \log(Mmnd)\right)$$

returns a group element $g \in \mathrm{SL}(n)$ *such that* $\|\mu(\pi(g)v)\|_F \leqslant \varepsilon$*. In particular, there is a* $\mathrm{poly}(\langle\pi\rangle, \langle v\rangle, \varepsilon^{-1})$ *time algorithm to solve the scaling problem (Problem 1.9) for* $\mathrm{SL}(n)$*.*

Proof. The bound on the maximal number of iterations follows by combining Theorem 4.2 with the upper bound on the weight norm from Lemma 6.1, the capacity lower bound from Corollary 7.11, and Lemma 7.2, which shows that $\|v\| \leqslant e^{nd\log(nd)}m^{1/2}M$. To see that this implies a $\mathrm{poly}(\langle\pi\rangle, \langle v\rangle, \varepsilon^{-1})$-time algorithm for the scaling problem, recall from Remark 1.22 that we can always preprocess so that that $d \leqslant m$ and use Remark 7.16. □

Next, we bound the running time of our second order algorithm.

Theorem 1.24 (Second order algorithm for norm minimization in terms of input size)**.** *Let* (π, v, ε) *be an instance of the scaling problem for* $\mathrm{SL}(n)$ *such that* $0 \in \Delta(v)$ *and every entry of* v *is bounded in absolute value by* M. *Let* d *denote the degree,* m *the dimension, and* γ *the weight margin of* π. *Then, Algorithm 5.1 applied to a suitably regularized objective function and a number of iterations at most*

$$T = O\left(\frac{d\sqrt{n}}{\gamma}\left(mn^3 d\log(Mmnd) + \log\frac{1}{\varepsilon}\right)\log\left(\frac{mnd\log M}{\varepsilon}\right)\right)$$

returns a group element $g \in \mathrm{SL}(n)$ *such that* $\log\|\pi(g)v\| \leqslant \log\mathrm{cap}(v) + \varepsilon$. *In particular, there is an algorithm to solve the norm minimization problem (Problem 1.10) for* $\mathrm{SL}(n)$ *in time* $\mathrm{poly}(\langle\pi\rangle, \langle v\rangle, \gamma^{-1}, \log(\varepsilon^{-1}))$, *which is at most* $\mathrm{poly}(\langle\pi\rangle, \langle v\rangle^n, \log(\varepsilon^{-1}))$.

Proof. Similarly to the proof of Theorem 1.23, the bound on the maximal number of iterations follows by combining Theorem 5.6 with Lemmas 6.1 and 7.2 and Corollary 7.11. To see that this implies a $\mathrm{poly}(\langle\pi\rangle, \langle v\rangle, \gamma^{-1}, \log(\varepsilon^{-1}))$-time algorithm for the scaling problem, recall from Remark 1.22 that we can always preprocess so that that $d \leqslant m$ and use Remark 7.16. The final claim follows using the bound $\gamma \geqslant d^{-n}n^{-3/2}$ from Table 1.1. □

Corollary 1.25 (Algorithm for null cone membership problem in terms of input size)**.** *There is an algorithm to solve the null cone membership problem (Problem 1.8) for* $\mathrm{SL}(n)$ *in time* $\mathrm{poly}(\langle\pi\rangle, \langle v\rangle, \gamma^{-1})$, *which is at most* $\mathrm{poly}(\langle\pi\rangle, \langle v\rangle^n)$.

Proof. This follows from either Theorem 1.23 or Theorem 1.24. Choose ε as in Corollary 1.18, use the bounds $\gamma(\pi) \geqslant d^{-n}n^{-3/2}$ and $N(\pi) \leqslant d$ from Table 1.1, and preprocess such that $d \leqslant m$. □

We now consider the p-scaling problem for representations of $\mathrm{GL}(n)$. We start with a simple lemma that we will use in the theorem below.

Lemma 7.17. *Let* $\pi\colon \mathrm{GL}(n) \to \mathrm{GL}(m)$ *be a homogeneous polynomial representation of degree* d. *Then, for all* $g \in \mathrm{GL}(n)$,

$$\max_{k,l\in[m]} |\pi_{k,l}(g)| \leqslant n^{2d} R(\pi) \max_{i,j\in[n]} |g_{i,j}|^d$$

Proof. Each $\pi_{k,l}(g)$ is a homogeneous polynomial of degree d in the n^2 matrix entries $g_{i,j}$. There are at most $\binom{d+n^2-1}{d} \leqslant (n^2)^d = n^{2d}$ monomials in such a polynomial, and each coefficient is bounded by in absolute value by $R(\pi)$. Thus the claim follows. □

Finally, we state our algorithm for the p-scaling problem and bound its running time.

Input:

- A polynomial representation $\pi\colon \mathrm{GL}(n) \to \mathrm{GL}(m)$ given in a Gelfand-Tsetlin basis that is homogeneous of degree $d > 0$,
- a vector $v \in \mathbb{Z}[i]^m$,
- a target point $p \in \mathbb{Q}^n \cap \Delta_d(n)$,
- a number of iterations T.

Output: A group element g.

Algorithm:

1. Choose $g_0 \in \mathrm{Mat}(n)$ as an integer matrix with entries drawn i.i.d. uniformly at random from $[S]$, where $S := 4n^{3n^3+1}d^{n^4}$. If $\det(g_0) = 0$, **fail**.
2. Let g_1 be the output of Algorithm 4.3 applied to $\pi(g_0)v$, target point p, and number of iterations T.
3. **Return** $g_1 g_0$.

Algorithm 7.1: Randomized algorithm for the p-scaling problem for $\mathrm{GL}(n)$ (cf. Theorem 1.26).

Theorem 1.26 (First-order randomized algorithm for p-scaling in terms of input size)**.** *Let* (π, v, p, ε) *be an instance of the moment polytope problem for* $\mathrm{GL}(n)$ *such that* $p \in \Delta(v)$ *and every entry of* v *is bounded in absolute value by* M*. Let* d *denote the degree and* m *the dimension of* π*. Then, with probability at least* $1/2$*, Algorithm 7.1 with a number of iterations at most*

$$T = O\left(\frac{d^3}{\varepsilon^2} mn^5 \log(Mmnd)\right).$$

returns a group element $g \in \mathrm{GL}(n)$ *such that* $\|\mathrm{spec}(\mu(\pi(g)v) - p\|_2 \leqslant \varepsilon$*. In particular, there is a randomized algorithm to solve the* p*-scaling problem (Problem 1.12) for* $\mathrm{GL}(n)$ *in time* $\mathrm{poly}(\langle\pi\rangle, \langle v\rangle, \langle p\rangle, \varepsilon^{-1})$ *and using* $\mathrm{poly}(\langle\pi\rangle, \langle v\rangle)$ *bits of randomness.*

Proof. From Theorem 7.15, we know that $\det(g_0) \neq 0$ and

$$-\log \mathrm{cap}_p(\pi(g_0)v) = O\big(mn^5 d \log(mnd)\big)$$

with probability at least $1/2$. Note that $\|p\|_2 \leqslant N(\pi) \leqslant d$ by Lemmas 6.1 and 3.11, since $p \in \Delta(v)$. Thus, $N^2 := N(\pi)^2 + \|p\|_2 \leqslant d^2 + d$, and hence $N^2 = O(d^2)$. Furthermore, we find using Lemmas 7.2 and 7.17 and Theorem 7.7 that

$$\begin{aligned}\|\pi(g_0)v\| &\leqslant e^{nd\log(nd)}\|\pi(g_0)v\|_2 \leqslant e^{nd\log(nd)}\|\pi(g_0)\|_F\|v\|_2 \leqslant e^{nd\log(nd)} n \max_{k,l}|\pi_{k,l}(g_0)|\|v\|_2 \\ &\leqslant e^{nd\log(nd)} n^{2d+1} R(\pi) S^d \|v\|_2 = e^{O(mn^4 d\log(Mmnd))}\end{aligned}$$

The bound on the maximal number of iterations follows from these estimates by using Theorem 4.5. To see that this implies a $\mathrm{poly}(\langle\pi\rangle, \langle v\rangle, \langle p\rangle, \varepsilon^{-1})$-time algorithm for the scaling problem, recall from Remark 1.22 that we can always preprocess so that that $d \leqslant m$ and reason analogously to Remark 7.16. Since $\log(S) = O(n^4 \log(d))$, we only require $\mathrm{poly}(\langle\pi\rangle, \langle v\rangle)$ bits of randomness. □

Lemma 7.18. *Let $\pi\colon \mathrm{GL}(n) \to \mathrm{GL}(m)$ be a homogeneous polynomial representation of degree d. Let $p \in C(G)$ be rational with $\|p\|_2 \leqslant N(\pi)$ and let $\ell > 0$ be an integer such that ℓp is a highest weight. Set $\varepsilon := (2\ell)^{-n-1} d^{-n} n^{-1}$. Then, for all $v \in V \setminus \{0\}$, solving the p-scaling problem with input (π, v, p, ε) suffices to solve the moment polytope membership problem for (π, v, p).*

Proof. In view of Corollary 3.31 it suffices to verify that $\varepsilon \leqslant \gamma(\rho)/2\ell$, where $\gamma(\rho)$ refers to the weight margin of the representation $\rho := \pi^{\otimes \ell} \otimes \pi_{\lambda^*}$, with $\lambda := \ell p$. The weight norm of ρ can be estimated as $N(\rho) \leqslant \ell N(\pi) + \|\lambda\|_2 \leqslant 2\ell d$. Thus, the weight margin is at least $\gamma(\rho) \geqslant (2\ell d)^{-n} n^{-1}$ by Theorem 6.8. This confirms that $\varepsilon \leqslant \gamma(\rho)/2\ell$. □

If we choose ℓ as the product of the denominators of the entries of p, then $\ell \leqslant 2^{\langle p\rangle}$, hence

$$\log(\varepsilon^{-1}) = \log\big((2\ell)^{n+1} d^n n\big) = O(n\,\langle p\rangle) + n\log(d). \tag{7.4}$$

Thus, the bitsize of ε is polynomial in the input size provided that $d \leqslant m$ (which we can always be achieved by preprocessing).

Corollary 1.27 (Randomized algorithm for moment polytope membership in terms of input size)**.** *There is a randomized algorithm to solve the moment polytope membership problem (Problem 1.11) for $\mathrm{GL}(n)$ in time $\mathrm{poly}(\langle\pi\rangle, \langle v\rangle^n, 2^{n\langle p\rangle})$ and using $\mathrm{poly}(\langle\pi\rangle, \langle v\rangle)$ bits of randomness.*

Proof. We may assume that $\|p\|_2 \leqslant N(\pi)$, since otherwise $p \notin \Delta(v)$. By preprocessing, we may assume that $d \leqslant m$ (Remark 1.22). Then the corollary follows from Theorem 1.26 and Lemma 7.18 and the preceding discussion. □

8 Conclusion

This paper initiates a systematic development of a theory of *non-commutative* optimization which greatly extends ordinary (Euclidean) convex optimization. This setting captures natural geodesically convex optimization problems on Riemannian manifolds that arise from the symmetries of non-commutative groups. It unifies a diverse range of problems – many non-convex – in different areas of computer science, mathematics, and physics. Several of them were solved efficiently for the first time using non-commutative methods. The corresponding algorithms also lead to solutions of purely structural problems, and to many new connections between disparate fields. Our work points to intriguing open problems and suggests further research directions. We believe that extending this theory will be fruitful both from a mathematical and computational point of view. It provides a meeting place for ideas and techniques from several different areas of research, and promises better algorithms for existing and yet unforeseen applications. We mention a few concrete challenges:

1. Is the null cone membership problem for general group actions in P? A natural intermediate goal is to prove that they are in NP ∩ coNP. The quantitative duality theory developed in this paper makes such a result plausible. The same question may be asked about the moment polytope membership problem for general group actions [BCMW17].

2. Can we find more general classes of problems or group actions where our algorithms run in polynomial time? In view of the complexity parameters we have identified, it is of particular interest to understand when the *weight margin* is only inverse polynomially rather than exponentially small.

3. Interestingly, when restricted to the commutative case discussed in Section 1.4, our algorithms' guarantees do not match those of cut methods in the spirit of the ellipsoid algorithm. Can we extend non-commutative/geodesic optimization to include cut methods as well as interior point methods? The foundations we lay in extending first and second order methods to the non-commutative case makes one optimistic that similar extensions are possible of other methods in standard convex optimization.

4. Can geodesic optimization lead to new or different efficient algorithms in combinatorial optimization? We know that it captures algorithmic problems like bipartite matching (and more generally matroid intersection). How about perfect matching in general graphs – is the Edmonds polytope a moment polytope of a natural group action?

5. Can geodesic optimization lead to new or different efficient algorithms in algebraic complexity and derandomization? We know that it captures PIT (polynomial identity testing) in non-commuting variables. Is the classical PIT problem a null cone membership problem for some group action? If not, can we identify the required generalizations and extend our methods to solve it? Which algebraic varieties are *not* null cones of group actions?

A Lifting coefficient bounds

In this appendix we prove Proposition 7.6, which explains how bounds on the Lie algebra representation Π of the basis matrices $E_{i,j}$ can be used to bound $R(\pi)$. We first state a number of elementary lemmas.

Lemma A.1. *Let* p *be a polynomial of degree* d *in variables* $X_1,\dots,X_N$ *and with integer coefficients bounded in absolute value by* R. *Let* $\delta \in \{0,\pm1\}^N$. *Then,* $p(X+\delta)$ *has integer coefficients bounded in absolute value by* $R(d+1)^N 2^d$.

Proof. Write p as a sum of monomials, say, $p = \sum_\omega p_\omega X^\omega$. Then, $p(X+\delta) = \sum_\omega p_\omega (X_1 + \delta_1)^{\omega_1} \cdots (X_N + \delta_N)^{\omega_N} = \sum_\nu c_\nu X^\nu$, where

$$c_\nu = \sum_{\omega_1 \geqslant \nu_1,\dots,\omega_N \geqslant \nu_N} c_{\nu,\omega} p_\omega \binom{\omega_1}{\nu_1} \cdots \binom{\omega_N}{\nu_N}$$

and each $c_{\nu,\omega}$ is a product of entries of δ, hence in $\{0,\pm1\}$. Clearly,

$$|c_\nu| \leqslant R \sum_{\substack{\omega_1 \geqslant \nu_1,\dots,\omega_N \geqslant \nu_N \\ \omega_1+\cdots+\omega_N \leqslant d}} \binom{\omega_1}{\nu_1} \cdots \binom{\omega_N}{\nu_N} \leqslant R \sum_{\omega_1+\cdots+\omega_N \leqslant d} 2^d = R\binom{N+d}{d} 2^d \leqslant R(d+1)^N 2^d,$$

which concludes the proof. □

Lemma A.2. *Let* $X = (X_{i,j})_{i,j\in[n]}$ *be a symbolic square matrix. Then,* $\det(X+I)$ *is a multilinear (nonhomogeneous) polynomial with nonzero coefficients equal to* ±1.

Proof. Recall that $\det(X) = \sum_{\sigma \in S_n} (-1)^\sigma \prod_{i=1}^n X_{i,\sigma(i)}$. Consider one of its monomials and write

$$\prod_{i=1}^n X_{i,\sigma(i)} = \left(\prod_{i \in F(\sigma)} X_{i,i} \right) \left(\prod_{i \notin F(\sigma)} X_{i,\sigma(i)} \right),$$

where $F(\sigma)$ denotes the fixed point set of the permutation σ. When we substitute $X \mapsto X + I$, we obtain the polynomial

$$\left(\prod_{i \in F(\sigma)} (X_{i,i} + 1) \right) \left(\prod_{i \notin F(\sigma)} X_{i,\sigma(i)} \right) = \sum_{T \subseteq F(\sigma)} \left(\prod_{i \in T} X_{i,i} \right) \left(\prod_{i \notin F(\sigma)} X_{i,\sigma(i)} \right)$$

It remains to argue that any monomial in the right-hand side uniquely determines the permutation σ. But this holds since σ is determined by its action on the set of non-fixed points, which we can read off from the variables $X_{i,j}$ with $i \neq j$. This concludes the proof. □

Lemma A.3. *Let* f *and* $q_1, \dots, q_t$ *be polynomials in variables* $X_1, \dots, X_N$ *with integer coefficients bounded in absolute value by* R_f *and by* R_q*, respectively. Let* $d_f \geqslant 2$ *be an upper bound to the degree of* f *and suppose that the* q_i *have positive degree. Suppose, moreover, that* $\prod_{i=1}^t q_i$ *divides* f*,* $q_1(0) = \cdots = q_t(0) = 1$*, and that the quotient polynomial* $h = f / \prod_{i=1}^t q_i$ *has positive degree* d*. Then* h *has integer coefficients whose absolute values are bounded by* $R_f R_q^d (2 d_f d)^{4Nd}$.

Proof. Since $q_i(0) = 1$, we can write $q_i = 1 - \hat{q}_i$, where $\hat{q}_i(0) = 0$. That is, $\hat{q}_i$ has no constant term, which implies that, for every k, $\hat{q}_i^k$ contains no monomials of degree less than k. Now,

$$h = \frac{f}{\prod_{i=1}^t q_i} = \frac{f}{\prod_{i=1}^t (1 - \hat{q}_i)} = f \prod_{i=1}^t \left(1 + \hat{q}_i + \hat{q}_i^2 + \cdots + \hat{q}_i^d + G_i\right),$$

where $G_i = \sum_{k=d+1}^\infty \hat{q}_i^k$. The above holds on an open neighborhood of $X = 0$ and allows us to calculate the coefficients of h in terms of the right-hand side expansion. Since $h(x)$ is a polynomial of degree d and $G_i(x)$ does not have any terms of degree less than or equal to d, it follows that

$$h = \left[f \prod_{i=1}^t \left(1 + \hat{q}_i + \hat{q}_i^2 + \cdots + \hat{q}_i^d\right) \right]_{\leqslant d},$$

where we write $[p]_{\leqslant d}$ for the sum of homogeneous parts of degree less than or equal to d of a polynomial p. Thus, we are left with bounding the coefficients of the homogeneous parts of degree $\leqslant d$ of the right hand side above. Rewriting further,

$$h = \sum_{b_1 + \cdots + b_t \leqslant d} \left[f \prod_{i=1}^t \hat{q}_i^{b_i} \right]_{\leqslant d} = \sum_{\substack{b_1 + \cdots + b_t \leqslant d \\ c + c_1 + \cdots + c_t \leqslant d}} [f]_c \prod_{i=1}^t [\hat{q}_i^{b_i}]_{c_i}, \tag{A.1}$$

where we write $[p]_c$ for the homogeneous part of degree c of a polynomial p. Since q_i divides f, each $\hat{q}_i^{b_i}$ is a polynomial of degree at most $b_i d_f$, and its coefficients are integers bounded in absolute value by $R_q^{b_i}(1 + \frac{b_i d_f}{N})^{N(b_i - 1)}$ by Lemma 7.4. Clearly, $[\hat{q}_i^{b_i}]_{c_i}$ satisfies the same coefficient bound, but is homogeneous of degree c_i. Now consider the product $[f]_c \prod_{i=1}^t [\hat{q}_i^{b_i}]_{c_i}$. The right-hand side may be written as a product of at most d polynomials, since at most d of the c_i are nonzero and $[\hat{q}_i^{b_i}]_0 = \delta_{0,b_i}$. Thus, Lemma 7.4 shows that each summand of Eq. (A.1) is a polynomial of degree at most d and with integer coefficients bounded in absolute value by

$$R_f \left(\prod_{i=1}^t R_q^{b_i} \left(1 + \frac{d_f b_i}{N}\right)^{N(b_i - 1)} \right) d^d \leqslant R_f R_q^d \left(1 + \frac{d_f d}{N}\right)^{Nd} d^d \leqslant R_f R_q^d \,(d_f d)^{2Nd}$$

As h is a sum of $\binom{d+t}{d}\binom{d+t+1}{d} \leqslant (t+1)^d (t+2)^d$ of such polynomials, $h(x)$ has integer coefficients that can be bounded in absolute value by

$$(t+1)^d (t+2)^d R_f R_q^d \,(d_f d)^{2Nd} \leqslant R_f R_q^d \,(2 d_f d)^{4Nd}$$

where in the last equality we used the fact that $t \leqslant d_f$. □

We now prove Proposition 7.6, which we restate for convenience.

Proposition 7.6 (Lifting coefficient bounds). *Let $\pi\colon \mathrm{GL}(n) \to \mathrm{GL}(m)$ be a homogeneous polynomial representation of degree $d > 0$. Let $\Pi\colon \mathrm{Mat}(n) \to \mathrm{Mat}(m)$ denote its Lie algebra representation and suppose $\Pi(E_{i,i})$ is diagonal for all $i \in [n]$. Let $\beta \leqslant R$ be positive integers such that the entries of $\beta\Pi(E_{i,j})$ are integers of absolute value at most R for all $i, j \in [n]$. Then, the entries of $\pi(g)$ are polynomials with rational coefficients in the entries of $g \in \mathrm{GL}(n)$, and $R(\pi) = R^{2m} e^{O(mn^3 d \log(mnd))}$.*

Proof. We first prove that the coefficients are rational, and afterwards analyze the size of the numerators and the common denominator.

For the former, it suffices to prove that the entries of $\pi(g)$ are on a dense subset of $\mathrm{GL}(n)$ given by rational functions with rational coefficients. Indeed, by assumption the entries of $\pi(g)$ are polynomials, polynomials are uniquely determined by their values on a dense subset, and a ratio of polynomials with rational coefficients that is a polynomial must be a polynomial with rational coefficients. We now proceed with this plan and prove that the entries of $\pi(g)$ are rational functions with rational coefficients on the dense subset of $g \in \mathrm{GL}(n)$ such that all leading principal minors of g are nonzero. Indeed, in this case we can write $g = LDU$, where D is diagonal and L and U are lower and upper triangular, respectively, with ones on the diagonal. The entries of L, D, U are given by rational functions in the entries of g [Hou66]:

$$D_{i,i} = \frac{|g|_{[i],[i]}}{|g|_{[i-1],[i-1]}}, \quad L_{i,j} = \begin{cases} \frac{|g|_{[j-1]\cup\{i\},[j]}}{|g|_{[j],[j]}} & \text{if } i \geqslant j \\ 0 & \text{if } i < j \end{cases}, \quad U_{i,j} = \begin{cases} \frac{|g|_{[i],[i-1]\cup\{j\}}}{|g|_{[i],[i]}} & \text{if } i \leqslant j \\ 0 & \text{if } i > j, \end{cases} \tag{A.2}$$

where we write $|g|_{I,J}$ for the minor of g corresponding to rows $I \subseteq [n]$ and columns $J \subseteq [n]$. We now show that the entries of $\pi(L)$, $\pi(D)$, and $\pi(U)$ are given by polynomials with rational coefficients in the entries of L, D, and U, respectively. Then clearly the entries of $\pi(g)$ are on a dense subset of $\mathrm{GL}(n)$ given by rational function with rational coefficients. Note that our assumption on the $\Pi(E_{i,i})$ implies that $\pi(D)$ is a diagonal matrix with entries

$$\pi(D)_{k,k} = D_{1,1}^{\Pi_{k,k}(E_{1,1})} \cdots D_{d,d}^{\Pi_{k,k}(E_{d,d})}. \tag{A.3}$$

Since π is homogeneous and polynomial of degree d, the exponents are necessarily non-negative integers adding to d. Thus, the nonzero entries of $\pi(D)$ are monomials of degree d in the entries of D. Next, note that $L - I$ is nilpotent, so $(L-I)^n = 0$. Thus, $L = \exp(\log(L))$, where

$$\log(L) = \log(I - (I - L)) = -\sum_{i=1}^{n-1} \frac{(I-L)^i}{i}, \tag{A.4}$$

so the entries of $\log(L)$ are polynomials with rational coefficients in the entries of L. Since $\log(L)$ is strictly lower triangular, $\Pi(\log(L))$ is nilpotent by basic representation theory (cf. Eq. (7.1)). It follows that $\Pi(\log(L))^m = 0$ and, hence, each entry of

$$\pi(L) = \exp\bigl(\Pi(\log(L))\bigr) = \sum_{j=0}^{m-1} \frac{\Pi(\log(L))^j}{j!} \tag{A.5}$$

is a polynomial with rational coefficients in the entries of L. By symmetry this is true for $\pi(U)$ as well. By the argument mentioned above, we conclude that $\pi(g)$ is a polynomial with rational coefficients in the entries of $g \in GL(n)$.

We now analyze how the numerators and common denominator grow in each step of the proof. Using the formula for the entries $D_{i,i}$ in Eq. (A.2), we may write $D = M_0/p_0$ where the common denominator p_0 and every nonzero entry of the matrix M_0 is a product of n principal minors of g. Recall from Eq. (A.3) that each nonzero entry of $\pi(D)$ is a monomial of degree d in the variables $D_{i,i}$. It follows that $\pi(D) = M_1/p_1$, where p_1 and each nonzero entry of M_1 is a product of at most nd principal minors of g. By Lemma 7.4, these are then homogeneous polynomials of degree at most $n^2 d$ and with integer coefficients bounded in absolute value by $(nd)^{n^2 d}$.

We now turn our attention to $\pi(L)$. We first bound $\log(L)$ and then use Eq. (A.5). Using the formula for the entries $L_{i,j}$ in Eq. (A.2), and noting that L has ones on its diagonal, we may write $I - L = M_2/p_2$. Here, the common denominator $p_2 := \prod_{i=1}^{n} |g|_{[i],[i]}$ is the product of all leading principal minors of g, and each nonzero entry of M_2 is a product of n minors of g. For each $i \in [n]$, we may then write $(I-L)^i = M_{3,i}/p_3$ with common denominator $p_3 = p_2^n$. By the iterated matrix multiplication, the entries of $M_{3,i} = p_2^{n-i} M_2^i$ can be written as a sum of n^{i-1} many terms, each of which is a product of n^2 many minors. By Lemma 7.4, it follows that each nonzero entry of $M_{3,i}$ is a homogeneous polynomial of degree at most n^3 and with integer coefficients bounded in absolute value by n^{2n^3+n}. Using Eq. (A.4), we find that $\log(L) = M_4/p_4$ with common denominator $p_4 = n! p_3$ and numerator $M_4 = -\sum_{i=1}^{n-1} (n!/i) M_{3,i}$. Thus, the entries of M_4 are homogeneous polynomials of degree at most n^3 and have integer coefficients bounded in absolute value by n^{2n^3+2n+1}.

Now recall that the entries of $\beta\Pi(E_{i,j})$ are integers bounded in absolute value by $R \geqslant \beta$. Let us write $\Pi(\log(L)) = M_5/p_5$ with numerator $M_5 = \sum_{i,j} (M_4)_{i,j} \beta\Pi(E_{i,j})$ and common denominator $p_5 = \beta p_4 = \beta n! p_2^n$. The entries of M_5 are homogeneous of degree at most n^3 and have integer coefficients bounded in absolute value by Rn^{2n^3+2n+3}. By Lemma 7.4, the same holds for p_5 with plenty of slack. For $j \in \{0, \dots, m-1\}$, we may then write $\Pi(\log(L))^j = M_{6,j}/p_6$ with $M_{6,j} = p_5^{m-j} M_5^j$ and common denominator $p_6 = p_5^m$. By iterated matrix multiplication, each entry of $M_{6,j}$ can be written as a sum of at most m^m terms, each of which is a product of $m - j$ copies of p_5 and of j many entries of M_5. As before, we use Lemma 7.4 to find that the entries of $M_{6,j}$ are homogeneous polynomials of degree at most mn^3 and have integer coefficients bounded in absolute value by $m^m(Rn^{2n^3+2n+3})^m m^{mn^3} = R^m(mn)^{O(mn^3)}$. Finally, using Eq. (A.5), we write $\pi(L) = M_7/p_7$, with $M_7 = \sum_{j=0}^{m-1}(m!/j!)M_{6,j}$ and common denominator $p_7 = m!p_6$. Clearly, the entries of M_7 are again homogeneous polynomials of degree at most mn^3 with integer coefficients bounded in absolute value by $R^m(mn)^{O(mn^3)}$. By symmetry, the same bounds hold for $\pi(U)$ as well. The upshot of all the above is that we can write

$$\pi(g) = \pi(L)\pi(D)\pi(U) = \frac{M_8}{p_8}$$

with common denominator $p_8 = p_1 p_7^2$. Using Lemma 7.4, every entry of M_8 is a homogeneous polynomial of degree at most $n^2 d + 2mn^3$ with integer coefficients bounded in absolute value by $R_8 := R^{2m}e^{O(mn^3 d\log(mnd))}$. Since π is polynomial and $\pi(g) = M_8/p_8$ on a dense subset, it follows that each entry of M_8 must be a multiple of p_8.

To finish the proof, we need to bound the coefficients of $\pi(g)$ in terms of the numerator and denominator. This will be achieved by using Lemma A.3. To apply the lemma, recall that p_8 is proportional to a product of principal minors. Indeed, $p_8 = p_1 p_7^2 = \alpha\prod_{i=1}^t q_i$, where $t \leqslant nd + 2mn^2$, each q_i is a principal minor of g, and $\alpha = (m!)^2(\beta n!)^{2m} = R^{2m}e^{O(mn\log(mn))}$. In particular, $q_1(I) = \cdots = q_t(I) = 1$. Now let $h(g)$ be an arbitrary entry of $\alpha\pi(g)$. Let f denote the corresponding entry of M_8, so that $h = f/\prod_{i=1}^t q_i$. Define $\tilde{h}(X) := h(X+I)$, $\tilde{f}(X) := f(X+I)$, and $\tilde{q}_i(X) := q_i(X+I)$, so that $\tilde{q}_1(0) = \cdots = \tilde{q}_t(0) = 1$. Clearly, it still holds that $\tilde{h} = \tilde{f}/\prod_{i=1}^t \tilde{q}_i$ is a polynomial of degree d. Moreover, $\tilde{f}$ still has degree at most $d_{\tilde{f}} := n^2 d + 2mn^3$. By Lemma A.1, the coefficients of $\tilde{f}$ are bounded in absolute value by $R_{\tilde{f}} := R_8(d_{\tilde{f}}+1)^{n^2}2^{d_{\tilde{f}}}$, while the nonzero coefficients of $\tilde{q}_i$ remain ± 1 according to Lemma A.2. With these bounds, Lemma A.3 shows that $\tilde{h}$ is a polynomial with integer coefficients bounded in absolute value by $\tilde{R} := R_{\tilde{f}}(2d_{\tilde{f}}d)^{4n^2 d}$. Using Lemma A.1 one more time, it follows that h is a polynomial with integer coefficients bounded in absolute value by $\tilde{R}(d+1)^{n^2}2^d = R^{2m}e^{O(mn^3 d\log(mnd))}$. Since h is an arbitrary entry of $\alpha\pi$ and since $\alpha = R^{2m}e^{O(mn\log(mn))}$, we finally obtain the claim. □

References

[AMS09] P-A Absil, Robert Mahony, and Rodolphe Sepulchre. *Optimization algorithms on matrix manifolds*. Princeton University Press, 2009.

[Ati82] Michael F Atiyah. Convexity and commuting Hamiltonians. *Bulletin of the London Mathematical Society*, 14(1):1–15, 1982. doi:10.1112/blms/14.1.1.

[AZGL+18] Zeyuan Allen-Zhu, Ankit Garg, Yuanzhi Li, Rafael Oliveira, and Avi Wigderson. Operator scaling via geodesically convex optimization, invariant theory and polynomial identity testing. In *Proceedings of the Symposium on the Theory of Computing (STOC 2018)*, pages 172–181, 2018. arXiv:1804.01076.

[AZLOW17] Zeyuan Allen-Zhu, Yuanzhi Li, Rafael Oliveira, and Avi Wigderson. Much faster algorithms for matrix scaling. In *Proceedings of the Symposium on Foundations of Computer Science (FOCS 2017)*, pages 890–901. IEEE, 2017. arXiv:1704.02315.

[BCMW17] Peter Bürgisser, Matthias Christandl, Ketan D Mulmuley, and Michael Walter. Membership in moment polytopes is in NP and coNP. *SIAM Journal on Computing*, 46(3):972–991, 2017. arXiv:1511.03675, doi:10.1137/15M1048859.

[BFG$^+$18] Peter Bürgisser, Cole Franks, Ankit Garg, Rafael Oliveira, Michael Walter, and Avi Wigderson. Efficient algorithms for tensor scaling, quantum marginals, and moment polytopes. In *Proceedings of the Symposium on Foundations of Computer Science (FOCS 2018)*, pages 883–897. IEEE, 2018. arXiv:1804.04739.

[BGO$^+$17] Peter Bürgisser, Ankit Garg, Rafael Oliveira, Michael Walter, and Avi Wigderson. Alternating minimization, scaling algorithms, and the null-cone problem from invariant theory. In *Proceedings of Innovations in Theoretical Computer Science (ITCS 2018)*, 2017. arXiv:1711.08039, doi:10.4230/LIPIcs.ITCS.2018.24.

[BI13] Peter Bürgisser and Christian Ikenmeyer. Deciding positivity of Littlewood-Richardson coefficients. *SIAM Journal on Discrete Mathematics*, 27(4):1639–1681, 2013. arXiv:1204.2484, doi:10.1137/120892532.

[BK06] Prakash Belkale and Shrawan Kumar. Eigenvalue problem and a new product in cohomology of flag varieties. *Inventiones mathematicae*, 166(1):185–228, 2006. arXiv:math/0407034, doi:10.1007/s00222-006-0516-x.

[Bri87] Michel Brion. Sur l'image de l'application moment. In *Séminaire d'algebre Paul Dubreil et Marie-Paule Malliavin*, volume 1296 of *Lecture Notes in Mathematics*, pages 177–192. Springer, 1987.

[BS00] Arkady Berenstein and Reyer Sjamaar. Coadjoint orbits, moment polytopes, and the Hilbert–Mumford criterion. *Journal of the American Mathematical Society*, 13(2):433–466, 2000. arXiv:math/9810125, doi:10.1090/S0894-0347-00-00327-1.

[Bür00] Peter Bürgisser. The computational complexity to evaluate representations of general linear groups. *SIAM Journal on Computing*, 30(3):1010–1022, 2000.

[BV04] Stephen Boyd and Lieven Vandenberghe. *Convex optimization*. Cambridge University Press, 2004.

[BVW16] Nicole Berline, Michèle Vergne, and Michael Walter. The Horn inequalities from a geometric point of view. *L'Enseignement Mathématique*, 63:403–470, 2016. arXiv:1611.06917.

[BVW18a] Velleda Baldoni, Michèle Vergne, and Michael Walter. Computation of dilated Kronecker coefficients. *Journal of Symbolic Computation*, 84:113–146, 2018. arXiv:1601.04325, doi:10.1016/j.jsc.2017.03.005.

[BVW18b] Velleda Baldoni, Michèle Vergne, and Michael Walter. Horn inequalities and quivers. 2018. arXiv:1804.00431.

[BVW19] Velleda Baldoni, Michèle Vergne, and Michael Walter. Horn conditions for Schubert positions of general quiver subrepresentations. 2019. arXiv:1901.07194.

[CDKW14] Matthias Christandl, Brent Doran, Stavros Kousidis, and Michael Walter. Eigenvalue distributions of reduced density matrices. *Communications in Mathematical Physics*, 332(1):1–52, 2014. arXiv:1204.0741, doi:10.1007/s00220-014-2144-4.

[CDW12] Matthias Christandl, Brent Doran, and Michael Walter. Computing multiplicities of Lie group representations. In *Proceedings of the Symposium on Foundations of Computer Science (FOCS 2012)*, pages 639–648. IEEE, 2012. arXiv:1204.4379, doi:10.1109/FOCS.2012.43.

[CGT00] Andrew R. Conn, Nicholas I. M. Gould, and Philippe L. Toint. *Trust Region Methods*, volume 1 of *MPS-SIAM Series on Optimization*. Society for Industrial and Applied Mathematics, 2000.

[CHM07] Matthias Christandl, Aram W Harrow, and Graeme Mitchison. Nonzero Kronecker coefficients and what they tell us about spectra. *Communications in Mathematical Physics*, 270(3):575–585, 2007. arXiv:quant-ph/0511029, doi:10.1007/s00220-006-0157-3.

[CL55] Earl A Coddington and Norman Levinson. *Theory of ordinary differential equations*. Tata McGraw-Hill Education, 1955.

[CM06] Matthias Christandl and Graeme Mitchison. The spectra of quantum states and the Kronecker coefficients of the symmetric group. *Communications in Mathematical Physics*, 261(3):789–797, 2006. arXiv:quant-ph/0409016, doi:10.1007/s00220-005-1435-1.

[CMTV17] Michael B Cohen, Aleksander Madry, Dimitris Tsipras, and Adrian Vladu. Matrix scaling and balancing via box constrained Newton's method and interior point methods. In *Proceedings of the Symposium on Foundations of Computer Science (FOCS 2017)*, pages 902–913. IEEE, 2017. arXiv:1704.02310.

[Der01] Harm Derksen. Polynomial bounds for rings of invariants. *Proceedings of the American Mathematical Society*, 129(4):955–963, 2001. doi:10.1090/S0002-9939-00-05698-7.

[DH05] Sumit Daftuar and Patrick Hayden. Quantum state transformations and the Schubert calculus. *Annals of Physics*, 315(1):80–122, 2005. arXiv:quant-ph/0410052, doi:10.1016/j.aop.2004.09.012.

[DK15] Harm Derksen and Gregor Kemper. *Computational invariant theory*. Springer, 2015.

[DLM06] Jesús A. De Loera and Tyrrell B. McAllister. On the computation of Clebsch-Gordan coefficients and the dilation effect. *Experimental Mathematics*, 15(1):7–19, 2006. arXiv:math/0501446.

[DM17] Harm Derksen and Visu Makam. Polynomial degree bounds for matrix semi-invariants. *Advances in Mathematics*, 310:44–63, 2017. arXiv:1512.03393, doi:10.1016/j.aim.2017.01.018.

[DM18] Harm Derksen and Visu Makam. Algorithms for orbit closure separation for invariants and semi-invariants of matrices. 2018. arXiv:1801.02043.

[DW00] Harm Derksen and Jerzy Weyman. Semi-invariants of quivers and saturation for Littlewood-Richardson coefficients. *Journal of the American Mathematical Society*, 13(3):467–479, 2000. doi:10.1090/S0894-0347-00-00331-3.

[DW17] Harm Derksen and Jerzy Weyman. *An introduction to quiver representations*, volume 184. American Mathematical Society, 2017.

[FH13] William Fulton and Joe Harris. *Representation theory: a first course*, volume 129. Springer, 2013.

[FK94] Jacques Faraut and Adam Korányi. *Analysis on symmetric cones*. Oxford University Press, 1994.

[Fra02] Matthias Franz. Moment polytopes of projective G-varieties and tensor products of symmetric group representations. *Journal of Lie Theory*, 12(2):539–549, 2002.

[Fra18] Cole Franks. Operator scaling with specified marginals. In *Proceedings of the 50th Annual ACM SIGACT Symposium on Theory of Computing*, pages 190–203. ACM, 2018. arXiv:1801.01412.

[FS13] Michael A. Forbes and Amir Shpilka. Explicit Noether normalization for simultaneous conjugation via polynomial identity testing. *Lecture Notes in Computer Science*, pages 527–542, 2013. arXiv:1303.0084, doi:10.1007/978-3-642-40328-6_37.

[Ful00] William Fulton. Eigenvalues, invariant factors, highest weights, and Schubert calculus. *Bulletin of the American Mathematical Society*, 37(3):209–249, 2000. arXiv:math/9908012.

[GGOW16] Ankit Garg, Leonid Gurvits, Rafael Oliveira, and Avi Wigderson. A deterministic polynomial time algorithm for non-commutative rational identity testing. In *Proceedings of the Symposium on Foundations of Computer Science (FOCS 2016)*, pages 109–117. IEEE, 2016. arXiv:1511.03730, doi:10.1109/FOCS.2016.95.

[GGOW17] Ankit Garg, Leonid Gurvits, Rafael Oliveira, and Avi Wigderson. Algorithmic and optimization aspects of Brascamp-Lieb inequalities, via operator scaling. In *Proceedings of the Symposium on the Theory of Computing (STOC 2017)*, pages 397–409. ACM, 2017. arXiv:1607.06711.

[GS82] V. Guillemin and S. Sternberg. Convexity properties of the moment mapping. *Inventiones mathematicae*, 67:491–513, 1982.

[Gur04a] Leonid Gurvits. Classical complexity and quantum entanglement. *Journal of Computer and System Sciences*, 69(3):448–484, 2004. arXiv:quant-ph/0303055, doi:10.1016/j.jcss.2004.06.003.

[Gur04b] Leonid Gurvits. Combinatorial and algorithmic aspects of hyperbolic polynomials. 2004. arXiv:math/0404474.

[Gur06] Leonid Gurvits. Hyperbolic polynomials approach to van der Waerden/Schrijver-Valiant like conjectures: sharper bounds, simpler proofs and algorithmic applications. In *Proceedings of the thirty-eighth annual ACM Symposium on Theory of Computing*, pages 417–426. ACM, 2006. arXiv:math/0510452.

[GY98] Leonid Gurvits and Peter N. Yianilos. The deflation-inflation method for certain semidefinite programming and maximum determinant completion problems. *Technical Report, NECI*, 1998.

[Hel79] Sigurdur Helgason. *Differential geometry, Lie groups, and symmetric spaces*, volume 80. Academic Press, 1979.

[Hil93] David Hilbert. Über die vollen Invariantensysteme. *Math. Ann.*, 42:313–370, 1893.

[HM18] Linus Hamilton and Ankur Moitra. The Paulsen problem made simple. In *Proceedings of Innovations in Theoretical Computer Science (ITCS 2019)*, 2018. arXiv:1809.04726, doi:10.4230/LIPIcs.ITCS.2019.41.

[Hou66] Alston S. Householder. *The theory of matrices in numerical analysis*. Dover, 1966.

[IMW17] Christian Ikenmeyer, Ketan D Mulmuley, and Michael Walter. On vanishing of Kronecker coefficients. *computational complexity*, 26(4):949–992, 2017. arXiv:1507.02955, doi:10.1007/s00037-017-0158-y.

[IQS17a] Gábor Ivanyos, Youming Qiao, and KV Subrahmanyam. Constructive non-commutative rank computation is in deterministic polynomial time. In *Proceedings of Innovations in Theoretical Computer Science (ITCS 2017)*, 2017. arXiv:1512.03531, doi:10.4230/LIPIcs.ITCS.2017.55.

[IQS17b] Gábor Ivanyos, Youming Qiao, and KV Subrahmanyam. Non-commutative Edmonds' problem and matrix semi-invariants. *Computational Complexity*, 26(3):717–763, 2017. arXiv:1508.00690.

[Kar84] Narendra Karmarkar. A new polynomial-time algorithm for linear programming. In *Proceedings of Symposium on the Theory of Computing (STOC 1984)*, pages 302–311. ACM, 1984.

[Kha79] Leonid G Khachiyan. A polynomial algorithm in linear programming. In *Doklady Academii Nauk SSSR*, volume 244, pages 1093–1096, 1979.

[KI04] Valentine Kabanets and Russell Impagliazzo. Derandomizing polynomial identity tests means proving circuit lower bounds. *Computational Complexity*, 13(1-2):1–46, 2004.

[Kin94] Alastair D King. Moduli of representations of finite dimensional algebras. *The Quarterly Journal of Mathematics*, 45(4):515–530, 1994. doi:10.1093/qmath/45.4.515.

[Kir84a] Frances Kirwan. Convexity properties of the moment mapping, III. *Inventiones mathematicae*, 77(3):547–552, 1984.

[Kir84b] Frances Clare Kirwan. *Cohomology of quotients in symplectic and algebraic geometry*, volume 31. Princeton University Press, 1984.

[KLLR18] Tsz Chiu Kwok, Lap Chi Lau, Yin Tat Lee, and Akshay Ramachandran. The Paulsen problem, continuous operator scaling, and smoothed analysis. In *Proceedings of the 50th Annual ACM SIGACT Symposium on Theory of Computing*, pages 182–189. ACM, 2018. arXiv:1710.02587.

[Kly98] Alexander A Klyachko. Stable bundles, representation theory and hermitian operators. *Selecta Mathematica, New Series*, 4(3):419–445, 1998. doi:10.1007/s000290050037.

[Kly04] Alexander Klyachko. Quantum marginal problem and representations of the symmetric group. 2004. arXiv:quant-ph/0409113.

[KN79] George Kempf and Linda Ness. The length of vectors in representation spaces. In *Algebraic geometry*, pages 233–243. Springer, 1979.

[Kos73] B. Kostant. On convexity, the Weyl group and the Iwasawa decomposition. *Ann. scient. E.N.S*, 6:413–455, 1973.

[Kra07] V. M. Kravtsov. Combinatorial properties of noninteger vertices of a polytope in a three-index axial assignment problem. *Cybernetics and Systems Analysis*, 43(1):25–33, 2007.

[KT99] A Knutson and T Tao. The honeycomb model of $\mathrm{GL}_n(\mathbb{C})$ tensor products I: Proof of the saturation conjecture. *Journal of the American Mathematical Society*, 12(4):1055–1090, 1999. arXiv:math/9807160.

[LSW98] Nati Linial, Alex Samorodnitsky, and Avi Wigderson. A deterministic strongly polynomial algorithm for matrix scaling and approximate permanents. In *Proceedings of the Symposium on the Theory of Computing (STOC 1998)*, pages 644–652, 1998.

[MNS12] Ketan D Mulmuley, Hariharan Narayanan, and Milind Sohoni. Geometric complexity theory III: on deciding nonvanishing of a Littlewood–Richardson coefficient. *Journal of Algebraic Combinatorics*, 36(1):103–110, 2012.

[Mol06] Alexander I. Molev. Gelfand-Tsetlin bases for classical Lie algebras. In *Handbook of Algebra*, volume 4, pages 109–170. Elsevier, 2006. arXiv:math/0211289, doi:10.1016/S1570-7954(06)80006-9.

[Mul12] Ketan D Mulmuley. Geometric complexity theory V: Equivalence between blackbox derandomization of polynomial identity testing and derandomization of Noether's normalization lemma. In *2012 IEEE 53rd Annual Symposium on Foundations of Computer Science*, pages 629–638. IEEE, 2012.

[Mul17] Ketan Mulmuley. Geometric complexity theory V: Efficient algorithms for Noether normalization. *Journal of the American Mathematical Society*, 30(1):225–309, 2017. arXiv:1209.5993.

[Mum65] David Mumford. *Geometric invariant theory*. Springer-Verlag, 1965.

[NM84] Linda Ness and David Mumford. A stratification of the null cone via the moment map. *American Journal of Mathematics*, 106(6):1281–1329, 1984. doi:10.2307/2374395.

[NN94] Yurii Nesterov and Arkadii Nemirovskii. *Interior-point polynomial algorithms in convex programming*, volume 13. SIAM, 1994.

[Res10] Nicolas Ressayre. Geometric invariant theory and the generalized eigenvalue problem. *Inventiones mathematicae*, 180(2):389–441, 2010. arXiv:0903.1187, doi:10.1007/s00222-010-0233-3.

[Res12] Nicolas Ressayre. Git-cones and quivers. *Mathematische Zeitschrift*, 270(1-2):263–275, 2012. arXiv:0903.1202, doi:10.1007/s00209-010-0796-0.

[Roe18] Philip Roeleveld. *A Tensor Scaling Algorithm with Truncation*. Bachelor's thesis, University of Amsterdam, 2018.

[RS05] Ran Raz and Amir Shpilka. Deterministic polynomial identity testing in non commutative models. *Computational Complexity*, 14:1–19, 2005.

[Sch86] Alexander Schrijver. *Theory of linear and integer programming*. Wiley-Interscience Series in Discrete Mathematics. John Wiley & Sons, 1986.

[SKM19] Hiroyuki Sato, Hiroyuki Kasai, and Bamdev Mishra. Riemannian stochastic variance reduced gradient algorithm with retraction and vector transport. *SIAM Journal on Optimization*, 29:1444–1472, 2019. arXiv:1702.05594.

[Stu08] Bernd Sturmfels. *Algorithms in invariant theory*. Springer, 2008.

[SV14] Mohit Singh and Nisheeth K Vishnoi. Entropy, optimization and counting. In *Proceedings of the Symposium on the Theory of Computing (STOC 2014)*, pages 50–59. ACM, 2014. arXiv:1304.8108.

[SV19] Damian Straszak and Nisheeth K Vishnoi. Computing maximum entropy distributions everywhere. In *Proceedings of Machine Learning Research, 32nd Annual Conference on Learning Theory*, 2019. arXiv:1711.02036.

[SVdB01] Aidan Schofield and Michel Van den Bergh. Semi-invariants of quivers for arbitrary dimension vectors. *Indagationes Mathematicae*, 12(1):125–138, 2001. arXiv:math/9907174, doi:10.1016/S0019-3577(01)80010-0.

[Udr94] Constantin Udriste. *Convex functions and optimization methods on Riemannian manifolds*, volume 297. Springer, 1994.

[VDDM03] Frank Verstraete, Jeroen Dehaene, and Bart De Moor. Normal forms and entanglement measures for multipartite quantum states. *Physical Review A*, 68(1):012103, 2003. arXiv:quant-ph/0105090, doi:10.1103/PhysRevA.68.012103.

[VW17] Michele Vergne and Michael Walter. Inequalities for moment cones of finite-dimensional representations. *Journal of Symplectic Geometry*, 15(4):1209–1250, 2017. arXiv:1410.8144, doi:10.4310/JSG.2017.v15.n4.a8.

[Wal14] Michael Walter. *Multipartite Quantum States and their Marginals*. PhD thesis, ETH Zurich, 2014. arXiv:1410.6820, doi:10.3929/ethz-a-010250985.

[Wal17] Norbert Wallach. *Geometric Invariant Theory Over the Real and Complex Numbers*. Springer, 2017.

[WDGC13] Michael Walter, Brent Doran, David Gross, and Matthias Christandl. Entanglement polytopes: multiparticle entanglement from single-particle information. *Science*, 340(6137):1205–1208, 2013. arXiv:1208.0365, doi:10.1126/science.1232957.

[Woo10] Christopher T Woodward. Moment maps and geometric invariant theory. *Les cours du CIRM*, 1:55–98, 2010. arXiv:0912.1132.

[ZRS16] Hongyi Zhang, Sashank J Reddi, and Suvrit Sra. Riemannian SVRG: Fast stochastic optimization on Riemannian manifolds. In *Advances in Neural Information Processing Systems*, pages 4592–4600, 2016. arXiv:1605.07147.

[ZS16] Hongyi Zhang and Suvrit Sra. First-order methods for geodesically convex optimization. In *Conference on Learning Theory*, pages 1617–1638, 2016. arXiv:1602.06053.

[ZS18] Hongyi Zhang and Suvrit Sra. Towards Riemannian accelerated gradient methods. 2018. arXiv:1806.02812.