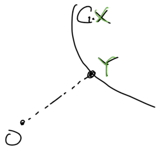
Introduction to geometric invariant theory II: Convexity, marginals & moment polytopes

Michael Walter (University of Amsterdam)

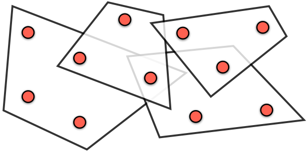
IAS, June 2018

University of Amsterdam QuSoft Research Center for Quantum Software

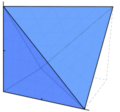
Plan for today



1. Convexity properties of $g \mapsto \|\pi(g)v\|^2$, which underlie optimization algorithms that we discuss this week.



2. Natural 'marginal' and 'scaling' problems, involving probability distributions and quantum states, related to the moment map.



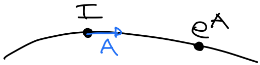
3. Moment polytopes that encode the answers to these problems, and their 'dual' optimization and invariant-theoretic characterization.

The geometry of invertible matrices

Any invertible matrix can be written as the exponential of an $n \times n$-matrix:

$$\mathrm{GL}_n = \{g = e^A \mid A \in \mathrm{Mat}_n\}$$

Since $e^{sA} = I + sA + O(s^2)$, can think of A as a tangent vector at I.



- If H, K Hermitian, then e^H positive definite, $u = e^{iK}$ unitary.

Polar decomposition: $g = u\, e^H$

Reminder: Moment map

Setup: A representation $\pi\colon \mathrm{GL}_n \to \mathrm{GL}(V)$ such that $\pi(U_n) \subseteq U(V)$.
Given a vector $v \in V$, consider squared norm function:

$$g \mapsto \|\pi(g)\, v\|^2$$

The moment map is its 'gradient':

$$\mu\colon V \to \mathrm{Herm}_n, \quad \mathrm{tr}\,[\mu(v)H] = \frac{1}{2}\partial_{s=0}\|\pi(e^{Hs})\, v\|^2 \quad (\forall H = H^\dagger)$$

Noncommutative duality from Ankit's talk: For $v \in V$,

$$0 \notin \overline{\pi(G)v} \quad \Leftrightarrow \quad \exists 0 \neq w \in \overline{\pi(G)v} : \mu(w) = 0.$$

Left-hand side: v not in null cone. Right-hand side: 'double stochastic'.

Reminder: Moment map

Setup: A representation $\pi\colon \mathrm{GL}_n \to \mathrm{GL}(V)$ such that $\pi(U_n) \subseteq U(V)$.
Given a vector $v \in V$, consider squared norm function:

$$g \mapsto \|\pi(g)\, v\|^2$$

The moment map is its 'gradient':

$$\mu\colon V \to \mathrm{Herm}_n, \quad \mathrm{tr}\left[\mu(v)H\right] = \frac{1}{2}\partial_{s=0}\|\pi(e^{Hs})\, v\|^2 \quad (\forall H = H^\dagger)$$

Noncommutative duality from Ankit's talk: For $v \in V$,

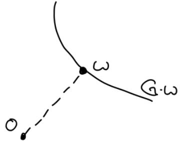
$$0 \notin \overline{\pi(G)v} \quad \Leftrightarrow \quad \exists 0 \neq w \in \overline{\pi(G)v} : \mu(w) = 0.$$

Left-hand side: v not in null cone. Right-hand side: 'double stochastic'.

Kempf-Ness theorem

It is implied by the following Kempf-Ness theorem:

$$\|\pi(g)w\|^2 \geq \|w\|^2 \quad (\forall g \in \mathrm{GL}_n) \quad \Leftrightarrow \quad \mu(w) = 0$$



($\Rightarrow$) since gradient vanishes at minimizers. Why ($\Leftarrow$)? Convexity!

Write $g = ue^H$. We only need to show that

$$f(s) := \|\pi(ue^{Hs})w\|^2$$

is convex, since then

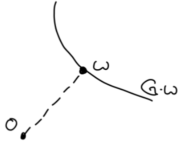
$$\|\pi(g)w\|^2 = f(1) \geq f(0) + f'(0) = \|w\|^2 + 2\underbrace{\mathrm{tr}\left[\mu(w)H\right]}_{=0} = \|w\|^2.$$

Conceptually, squared norm function is convex along geodesics (Thursday).

Kempf-Ness theorem

It is implied by the following Kempf-Ness theorem:

$$\|\pi(g)w\|^2 \geq \|w\|^2 \quad (\forall g \in \mathrm{GL}_n) \quad \Leftrightarrow \quad \mu(w) = 0$$



($\Rightarrow$) since gradient vanishes at minimizers. Why ($\Leftarrow$)? Convexity!

Write $g = ue^H$. We only need to show that

$$f(s) := \|\pi(ue^{Hs})w\|^2$$

is convex, since then

$$\|\pi(g)w\|^2 = f(1) \geq f(0) + f'(0) = \|w\|^2 + 2\underbrace{\mathrm{tr}\,[\mu(w)H]}_{=0} = \|w\|^2.$$

Conceptually, squared norm function is convex along geodesics (Thursday).

Proof of convexity

$$f(s) = \|\pi(ue^{Hs})w\|^2 = \|\pi(e^{Hs})w\|^2 = \|e^{\widetilde{H}s}w\|^2$$

We calculate:

$$f(s) = \langle e^{\widetilde{H}s}w, e^{\widetilde{H}s}w\rangle\,,$$
$$f'(s) = 2\,\langle e^{\widetilde{H}s}w, \widetilde{H}e^{\widetilde{H}s}w\rangle\,,$$
$$f''(s) = 4\,\langle e^{\widetilde{H}s}w, \widetilde{H}^2e^{\widetilde{H}s}w\rangle = 4\|\widetilde{H}e^{\widetilde{H}s}w\|^2 \geq 0. \qquad \square$$

In fact, even $\log f(s)$ is convex!

Can interpret calculation in terms of moment (cumulant) generating function.
One more derivative yields 'second-order robustness' $|f'''(s)| \leq c_H f''(s)$.

Proof of convexity

$$f(s) = \|\pi(ue^{Hs})w\|^2 = \|\pi(e^{Hs})w\|^2 = \|e^{\widetilde{H}s}w\|^2$$

We calculate:

$$f(s) = \langle e^{\widetilde{H}s}w, e^{\widetilde{H}s}w\rangle\,,$$
$$f'(s) = 2\,\langle e^{\widetilde{H}s}w, \widetilde{H}e^{\widetilde{H}s}w\rangle\,,$$
$$f''(s) = 4\,\langle e^{\widetilde{H}s}w, \widetilde{H}^2e^{\widetilde{H}s}w\rangle = 4\|\widetilde{H}e^{\widetilde{H}s}w\|^2 \geq 0. \qquad \square$$

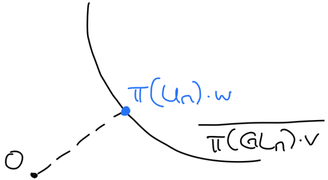
In fact, even $\log f(s)$ is convex!

Can interpret calculation in terms of moment (cumulant) generating function.
One more derivative yields 'second-order robustness' $|f'''(s)| \leq c_H f''(s)$.

Vectors of minimal norm

When is $\|\pi(g)w\|^2 = \|w\|^2$? Since $g = ue^H$,

$$f(1) = f(0) \;\Rightarrow\; f''(0) = 0 \;\Rightarrow\; \widetilde{H}w = 0 \;\Rightarrow\; \pi(g)w = \pi(u)w.$$



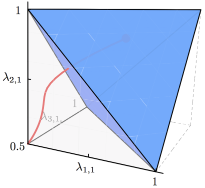
Theorem (Kempf-Ness)

In each GL_n-orbit closure, vectors of minimal norm form single U_n-orbit.

- can reduce orbit closure intersection problem $\overline{\pi(\mathrm{GL}_n)v} \cap \overline{\pi(\mathrm{GL}_n)v'} \neq \emptyset$ to orbit equality problem $\pi(U_n)w = \pi(U_n)w'$ for compact group

Algorithmic implications

Kirwan: Convexity ensures that gradient descent converges to global minimizer of $\|\pi(g)v\|^2$ (primal problem) and of $\frac{\|\mu(v)\|_F}{\|v\|^2}$ (dual problem).

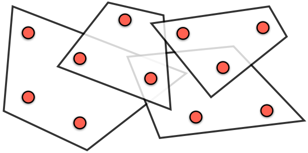


Suggests algorithmic solution by gradient methods:

- *continuous algorithms* such as continuous matrix scaling and operator scaling (Thursday)
- *discrete algorithms* can be understood as 'large step' variants: matrix, operator, tensor scaling (Avi, Rafael)

Also have general a priori bounds on primal and dual gaps (using invariant theory!).

Marginal problems and moment polytopes



Marginal problems

Visualize a joint probability distribution $p_{XY}(x, y)$ as matrix:

$$\begin{pmatrix} p_{XY}(1,1) & p_{XY}(1,2) & \dots \\ p_{XY}(2,1) & \ddots & \\ \vdots & & \end{pmatrix}$$

Then row & column sums are the marginal probability distributions:

$$p_X(x) = \sum_y p_{XY}(x, y), \quad p_Y(y) = \sum_x p_{XY}(x, y)$$

Any pair of marginals p_X, p_Y is compatible with a joint distribution.

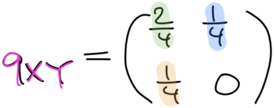
- Just choose $p_{XY}(x, y) = p_X(x) p_Y(y)$.

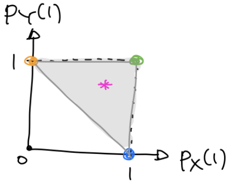
Which marginals can be obtained as scaling of some q_{XY}?

Matrix scaling as a marginal problem

Scalings are joint distributions $p_{XY}(x, y) = a(x)q_{XY}(x, y)b(y)$. Want:

$$\Delta(q_{XY}) := \{(p_X, p_Y) \mid p_{XY} \text{ is (asymptotic) scaling of } q_{XY}\}$$





Solution:

$$\Delta(q_{XY}) = \operatorname{conv}\{(\delta_x, \delta_y) \mid q_{XY}(x, y) \neq 0\}$$

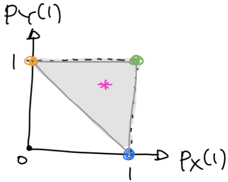
($\subseteq$) immediate from invariance of *support*. ($\supseteq$) not so obvious. . .

Matrix scaling as a marginal problem

Scalings are joint distributions $p_{XY}(x,y) = a(x) q_{XY}(x,y) b(y)$. Want:

$$\Delta(q_{XY}) := \{(p_X, p_Y) \mid p_{XY} \text{ is (asymptotic) scaling of } q_{XY}\}$$

$$q_{XY} = \begin{pmatrix} \frac{2}{4} & \frac{1}{4} \\ \frac{1}{4} & 0 \end{pmatrix}$$



Solution:

$$\Delta(q_{XY}) = \operatorname{conv}\{(\delta_x, \delta_y) \mid q_{XY}(x,y) \neq 0\}$$

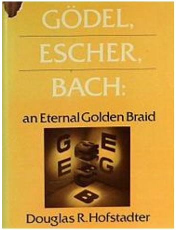
($\subseteq$) immediate from invariance of *support*. ($\supseteq$) not so obvious...

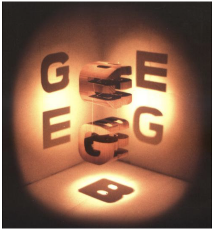
Further marginal problems

Given p_{XY} and p_{YZ}, are they compatible?

- Yes iff same p_Y.

But what if we want to obtain p_{XYZ} as a scaling? And how about p_{XZ}?





Solution to compatibility and scaling problems are convex polytopes.

- Key fact: Can relate $p_{XY} \mapsto (p_X, p_Y)$ etc. to moment maps for suitable representations (Ankit's talk)!

Convexity theorem for torus representations

Ankit's talk: Any representation $\pi\colon T \to \mathrm{GL}(V)$ of a torus $T = (\mathbb{C}^*)^n$ is of form $V = \bigoplus_{\omega\in\Omega} V_\omega$ for *weights* $\Omega \subseteq \mathbb{Z}^n$. Moment map:

$$\mu\colon V \to \mathbb{R}^n, \quad v = \sum_\omega v_\omega \mapsto \sum_{\omega\in\Omega} \|v_\omega\|^2 \omega$$

We are interested in:

$$\Delta = \left\{ \frac{\mu(v)}{\|v\|^2} \mid v \in V \right\}, \quad \Delta(w) = \mathrm{cl}\left\{ \frac{\mu(v)}{\|v\|^2} \mid v \in \pi(T)w, v \neq 0 \right\}$$

First object corresponds to compatibility, second to scaling problem.

Theorem (Atiyah)

Both are convex polytopes, known as moment polytopes:

$$\Delta = \mathrm{conv}\,\Omega, \quad \Delta(w) = \mathrm{conv}\,\mathrm{supp}(w)$$

where $\mathrm{supp}(v) = \{\omega : v_\omega \neq 0\}$!

Convexity theorem for torus representations

Ankit's talk: Any representation $\pi\colon T \to \mathrm{GL}(V)$ of a torus $T = (\mathbb{C}^*)^n$ is of form $V = \bigoplus_{\omega\in\Omega} V_\omega$ for *weights* $\Omega \subseteq \mathbb{Z}^n$. Moment map:

$$\mu\colon V \to \mathbb{R}^n, \quad v = \sum_{\omega} v_\omega \mapsto \sum_{\omega\in\Omega} \|v_\omega\|^2 \omega$$

We are interested in:

$$\Delta = \left\{ \frac{\mu(v)}{\|v\|^2} \mid v \in V \right\}, \quad \Delta(w) = \mathrm{cl}\left\{ \frac{\mu(v)}{\|v\|^2} \mid v \in \pi(T)w, v \neq 0 \right\}$$

First object corresponds to compatibility, second to scaling problem.

Theorem (Atiyah)

Both are convex polytopes, known as moment polytopes:

$$\Delta = \operatorname{conv}\Omega, \quad \Delta(w) = \operatorname{conv}\operatorname{supp}(w)$$

where $\operatorname{supp}(v) = \{\omega : v_\omega \neq 0\}$!

Moment polytopes and computation

marginal problems for probability distributions
$\subseteq$ moment polytopes for T-representations

Can be solved in polynomial time if given in form $v = \sum_\omega v_\omega$.
Simply compute support of v and solve an LP.

Natural questions:

- What if vector is only implicitly given?
- How about noncommutative groups?

Another example: Newton polytopes

Newton polytope of a homogeneous polynomial $P = \sum_\omega a_\omega x_1^{\omega_1} \dots x_n^{\omega_n}$:

$$\Delta(P) := \text{conv } \{\omega \mid a_\omega \neq 0\}$$

E.g., for $P = 5x_1x_2 + 3x_1^3 + 7x_2^2$: $\Delta(P) = \text{conv } \{(1,1),(3,0),(0,2)\}$.

- Newton polytopes are moment polytopes!

How difficult is it to determine Newton polytope when polynomial is given as 'black box' that allows us only to evaluate?

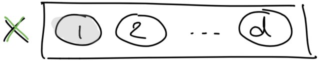
Efficient for class of 'hyperbolic' polynomials (Gurvits)!

What is a natural 'black box model' for general representations?

Quantum states and marginals

(Pure) quantum state of d particles is described by unit vector

$$X \in V = \mathbb{C}^{n_1} \otimes \ldots \otimes \mathbb{C}^{n_d}$$



Quantum marginals describe state of i-th particle: $n_i \times n_i$-matrices ρ_i^X

$$\mathrm{tr}[\rho_1^X A_1] = \langle X, (A_1 \otimes I \otimes \ldots \otimes I) X \rangle \quad \forall A_1$$

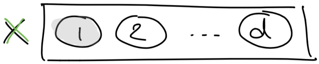
- $\rho_1^X = MM^\dagger$ if we ‘flatten’ X to $n_1 \times (n_2 \cdots n_d)$ matrix M (etc.)
- eigenvalues form probability distribution

We can similarly define ρ_S^X for any subset of particles $S \subseteq [d]$.

Quantum states and marginals

(Pure) quantum state of d particles is described by unit vector

$$X \in V = \mathbb{C}^{n_1} \otimes \ldots \otimes \mathbb{C}^{n_d}$$



Quantum marginals describe state of i-th particle: $n_i \times n_i$-matrices ρ_i^X

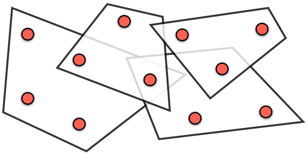
$$\mathrm{tr}[\rho_1^X A_1] = \langle X, (A_1 \otimes I \otimes \ldots \otimes I) X \rangle \quad \forall A_1$$

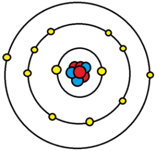
- $\rho_1^X = MM^\dagger$ if we 'flatten' X to $n_1 \times (n_2 \cdots n_d)$ matrix M (etc.)
- eigenvalues form probability distribution

We can similarly define ρ_S^X for any subset of particles $S \subseteq [d]$.

Quantum marginal problems

Given $\{\rho_S\}$, does there exist a compatible X ($\rho_S^X = \rho_S$ for given S)?





Fundamental problem: when can we patch together local data?

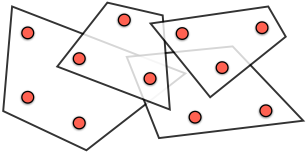
- Pauli principle: $\rho_i^X \leq I/d$ for electrons (X antisymmetric).

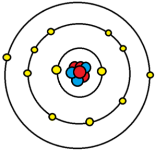
Physics is local: *energy*, *magnetization*, etc. depend only on few-particle marginals

- X exp large (in d), while marginals $\{\rho_S^X\}$ typically poly small.
- unfortunately, QMA-hard ('quantum NP'-hard) in general...
  (even if X need not be pure)

Quantum marginal problems

Given $\{\rho_S\}$, does there exist a compatible X ($\rho_S^X = \rho_S$ for given S)?





Fundamental problem: when can we patch together local data?

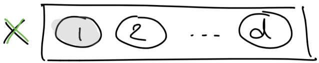
- Pauli principle: $\rho_i^X \leq I/d$ for electrons (X antisymmetric).

Physics is local: *energy*, *magnetization*, etc. depend only on few-particle marginals

- X exp large (in d), while marginals $\{\rho_S^X\}$ typically poly small.
- unfortunately, QMA-hard ('quantum NP'-hard) in general...
  (even if X need not be pure)

Single-particle quantum marginal problem

Given $(\rho_1, \ldots, \rho_d)$, are they compatible?

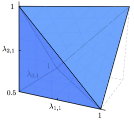


- $d = 2$: Yes iff ρ_1, ρ_2 have same nonzero eigenvalues.
- general answer only depends on eigenvalues:

$$X \mapsto (U_1 \otimes \ldots \otimes U_d) X \quad \rightsquigarrow \quad \rho_i^X \mapsto U_i \rho_i^X U_i^\dagger$$

Amazingly, answer is always given by convex polytope:

$$\Delta = \left\{ (\boldsymbol{p}_1^X, \ldots, \boldsymbol{p}_d^X) \mid \|X\| = 1 \right\}$$



where $\boldsymbol{p}_i^X$ ordered eigenvalues of quantum marginal ρ_i^X

$d = 3,\ n_i = 2$

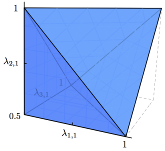
Tensor scaling as a marginal problem

Which quantum marginals can be obtained by scaling some Y? Recall a scaling is a quantum state of form $X = (g_1 \otimes \ldots \otimes g_d) Y$.

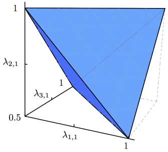
$$\Delta(Y) = \left\{ (\boldsymbol{p}_1^X, \ldots, \boldsymbol{p}_d^X) \mid X \text{ is (asymptotic) scaling of } Y \right\}$$

- $d = 2$: Only constraint is that rank cannot increase.

Again, $\Delta(Y)$ is *convex polytope*: the entanglement polytope of Y.
($\rightsquigarrow$ Matthias' talk)



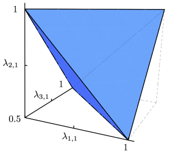
$Y = |000\rangle + |111\rangle$



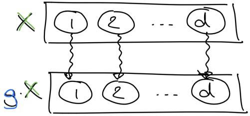
$W = |100\rangle + |010\rangle + |001\rangle$

Entanglement polytopes

$$\Delta(Y) = \left\{ (\boldsymbol{p}_1^X, \ldots, \boldsymbol{p}_d^X) \mid X \text{ is (asymptotic) scaling of } Y \right\}$$



Many applications:



- Quantum information & entanglement: tensors = quantum states, scalings = local transformations
- Algebraic complexity: tensors = computational problems, scalings = reductions
- Invariant theory and algebraic combinatorics: e.g., Kronecker coeffs
- *Operator scaling* and its many applications as 'special case' (Avi)

Why do we get convex polytopes?

- Key fact: The map $X \mapsto (\rho_1^X, \ldots, \rho_d^X)$ is a moment map (Ankit's talk)!

Convexity theorem for general actions

Setup: Representation $\pi\colon \mathrm{GL}_{n_1} \times \cdots \times \mathrm{GL}_{n_d} \to \mathrm{GL}(V)$ and moment map $\boldsymbol{\mu} = (\mu_1, \ldots, \mu_d)\colon V \to \mathrm{Herm}_{n_1} \oplus \ldots \oplus \mathrm{Herm}_{n_d}$. Compute:

$$v \longmapsto \underbrace{\frac{\boldsymbol{\mu}(v)}{\|v\|^2} = (\rho_1, \ldots, \rho_d)}_{\text{image of moment map}} \longmapsto \underbrace{\boldsymbol{p}(v) = (\boldsymbol{p}_1, \ldots, \boldsymbol{p}_d)}_{\text{ordered eigenvalues}} \in \mathbb{R}^{n_1 + \cdots + n_d}$$

We are interested in:

$$\Delta = \{\boldsymbol{p}(v) \mid v \in V\}, \quad \Delta(w) = \mathrm{cl}\,\{\boldsymbol{p}(v) \mid v \in \pi(G)w, v \neq 0\}$$

First object corresponds to compatibility, second to scaling problem.

Theorem (Kirwan, Mumford)

Both are convex polytopes, known as 'noncommutative' moment polytopes.

Can also study varieties that sit between orbit closure and entire space.

Moment polytopes and computation

marginal problems for quantum state
$\subseteq$ moment polytopes for G-representations

In contrast to the commutative case, polytopal nature not obvious and theorem does not give explicit description.

For the compatibility problem:

- Explicit inequalities known (Ressayre, ...), but quickly 'intractable'.
  In general, exponentially many facets!
- Membership problem is in NP $\cap$ coNP.

(a,b,c)	$(2,2,2)$ [43]	$(3,3,3)$ [45]	$(4,4,4)$
Inequalities	9 (3)	114 (25)	1749 (323)
Facets	6 (2)	45 (10)	270 (50)
Extreme Rays	5 (3)	33 (11)	328 (65)

Calls for algorithmic explanations!

Another example: Horn's problem

What are the possible eigenvalues $\boldsymbol{a}$, $\boldsymbol{b}$, $\boldsymbol{c}$ of Hermitian $n \times n$-matrices A, B, C such that $A + B = C$?

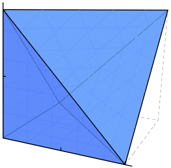
- Horn conjectured complete set of linear inequalities (e.g., $a_1 + b_1 \geq c_1$)
- proved by Knutson-Tao as consequence of saturation conjecture
- membership problem in polynomial time (Mulmuley)

Compatible eigenvalues characterized by moment polytope!

- $G = \mathrm{GL}_n^3$, $V = \mathrm{Mat}_n^2$, $\pi(g, h, k)(M, N) = (gMk^{-1}, hMk^{-1})$

Many further examples in physics (classical mechanics, geometric quantization, etc). Interestingly, not all quantum marginal problems fall into this framework!

Moment polytopes and noncommutative duality

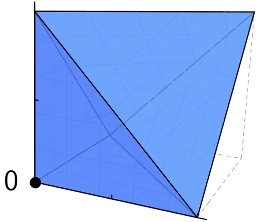


Reminder: Noncommutative duality

Can scale to *uniform marginals* iff not in null cone (Ankit), and null cone is defined by invariant polynomials (Harm). In our language:

$$0 \in \Delta(w) \;\Leftrightarrow\; \inf_{g \in G} \|\pi(g) w\|^2 > 0 \;\Leftrightarrow\; \exists P \in \mathbb{C}[V]^G : P(w) \neq P(0)$$

Uniform marginals correspond to *origin* of entanglement polytope:



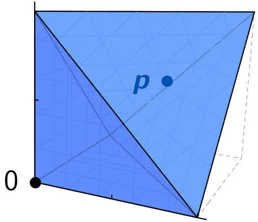
How about general marginals? When is $\boldsymbol{p} \in \Delta(w)$?

Reminder: Noncommutative duality

Can scale to *uniform marginals* iff not in null cone (Ankit), and null cone is defined by invariant polynomials (Harm). In our language:

$$0 \in \Delta(w) \;\Leftrightarrow\; \inf_{g \in G} \|\pi(g) w\|^2 > 0 \;\Leftrightarrow\; \exists P \in \mathbb{C}[V]^G : P(w) \neq P(0)$$

Uniform marginals correspond to *origin* of entanglement polytope:



How about general marginals? When is $\boldsymbol{p} \in \Delta(w)$?

Moment polytopes and invariant theory

Invariant polynomials span *trivial* irreducible representations in $\mathbb{C}[V]$.
Recall (Peter): Irreducible representations $\leftrightarrow$ highest weight vector $P_{\boldsymbol{\lambda}}$

$$P_{\boldsymbol{\lambda}}(\pi(b)^{-1}v) = \chi_{\boldsymbol{\lambda}}(b)\, P_{\boldsymbol{\lambda}}(v) \quad (\forall b \in B_n), \qquad \chi_{\boldsymbol{\lambda}}(b) = \prod_{j=1}^{n} {b_{jj}}^{\lambda_i}$$

Theorem (Mumford)

$$\Delta(w) = \left\{ \boldsymbol{p} = \frac{\boldsymbol{\lambda}}{k} \;\middle|\; \exists \text{HWV } P_{\boldsymbol{\lambda}^*} \in \mathbb{C}[V]_k : P_{\boldsymbol{\lambda}^*}(\pi(g_0)w) \neq 0 \right\}$$

Two complications:

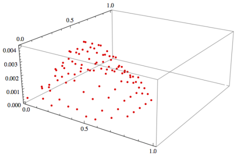
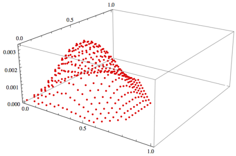
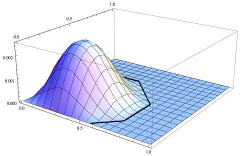
- Need to use 'dual' $\boldsymbol{\lambda}^* = (-\lambda_n, \ldots, -\lambda_1)$ (related to the $\pi(g)^{-1}$).
- Need to first apply generic $g_0 \in \mathrm{GL}_n$ (e.g., random unitary).

Moment polytopes and representation theory

Let $m_k(\boldsymbol{\lambda})$ denote *multiplicity* of $V_{\boldsymbol{\lambda}^*}$ in $\mathbb{C}[V]_k$. Then:

$$\Delta = \left\{ \boldsymbol{p} = \frac{\boldsymbol{\lambda}}{k} \mid \exists V_{\boldsymbol{\lambda}}^* \subseteq \mathbb{C}[V]_k \right\} = \left\{ \boldsymbol{p} = \frac{\boldsymbol{\lambda}}{k} \mid m_k(\boldsymbol{\lambda}) > 0 \right\}$$

e.g., Kronecker (quantum marginals) and Littlewood-Richardson cofficients (Horn)

Computational problems:

- *Counting*: $m_k(\boldsymbol{\lambda}) = ?$
- *Positivity*: $m_k(\boldsymbol{\lambda}) > 0$
- *Moment polytope*: $\frac{\boldsymbol{\lambda}}{k} \in \Delta$, i.e., $\exists s > 0 \colon m_{sk}(s\boldsymbol{\lambda}) > 0$

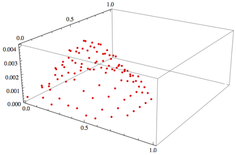
Generically, first #P-hard, second NP-hard, while third in NP ∩ coNP.

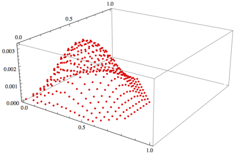
Moment polytopes and representation theory

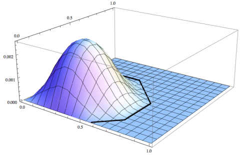
Let $m_k(\boldsymbol{\lambda})$ denote *multiplicity* of $V_{\boldsymbol{\lambda}^*}$ in $\mathbb{C}[V]_k$. Then:

$$\Delta = \left\{ \boldsymbol{p} = \frac{\boldsymbol{\lambda}}{k} \mid \exists V_{\boldsymbol{\lambda}}^* \subseteq \mathbb{C}[V]_k \right\} = \left\{ \boldsymbol{p} = \frac{\boldsymbol{\lambda}}{k} \mid m_k(\boldsymbol{\lambda}) > 0 \right\}$$

e.g., Kronecker (quantum marginals) and Littlewood-Richardson cofficients (Horn)







Computational problems:

- *Counting*: $m_k(\boldsymbol{\lambda}) = ?$
- *Positivity*: $m_k(\boldsymbol{\lambda}) > 0$
- *Moment polytope*: $\frac{\boldsymbol{\lambda}}{k} \in \Delta$, i.e., $\exists s > 0 \colon m_{sk}(s\boldsymbol{\lambda}) > 0$

Generically, first #P-hard, second NP-hard, while third in NP $\cap$ coNP.

Moment polytopes and optimization

We also have *noncommutative optimization duality* for general points in the moment polytope:

$$\boldsymbol{p} \in \Delta(w) \;\Leftrightarrow\; \inf_{b \in B_n} |\chi_{\boldsymbol{p}^*}(b)|^2 \, \|\pi(b)\pi(g_0)w\|^2 > 0$$

- scaling by upper-triangular matrices $b \in B_n$
- 'twisted' norm $=$ ordinary norm in larger space
- minimizers have desired marginals $\boldsymbol{p}$

For uniform marginals $\boldsymbol{p} = (1/n, \ldots, 1/n)$:

- $\chi_{\boldsymbol{p}^*}(b) b = \det(b)^{-1/n} b$ has determinant one!
- condition reduces to $\inf_{g \in \mathrm{SL}_n} \|\pi(g)w\|^2 > 0$ (Ankit's talk)

Moment polytopes and optimization

We also have *noncommutative optimization duality* for general points in the moment polytope:

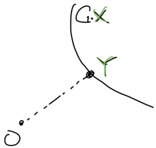
$$\boldsymbol{p} \in \Delta(w) \;\Leftrightarrow\; \inf_{b \in B_n} |\chi_{\boldsymbol{p}^*}(b)|^2 \, \|\pi(b)\pi(g_0)w\|^2 > 0$$

- scaling by upper-triangular matrices $b \in B_n$
- 'twisted' norm = ordinary norm in larger space
- minimizers have desired marginals $\boldsymbol{p}$

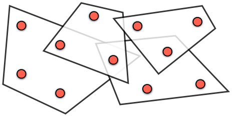
For uniform marginals $\boldsymbol{p} = (1/n, \ldots, 1/n)$:

- $\chi_{\boldsymbol{p}^*}(b) b = \det(b)^{-1/n} b$ has determinant one!
- condition reduces to $\inf_{g \in \mathrm{SL}_n} \|\pi(g)w\|^2 > 0$ (Ankit's talk)

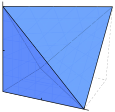
Summary



Convexity properties of $g \mapsto \|\pi(g)v\|^2$ underlying optimization algorithms that we will discuss this week.



The moment map (its 'gradient') is related to natural 'marginal' and 'scaling' problems involving probability distributions and quantum states.



Moment polytopes encode answers to these problems. 'Dual' optimization and invariant-theoretic characterizations. Often exponentially many facets, yet can admit efficient algorithms.

Many open questions: Poly-time algorithms? Quantum algorithms? $\mathbb{C} \rightsquigarrow \mathbb{F}$? Computational invariant theory without computing invariants?

Thank you for your attention!

Reductions to uniform marginals: shifting trick

Key idea: Modify representation so that $\frac{\boldsymbol{\lambda}}{k}$ becomes new origin.

Building blocks:

- $V \rightsquigarrow \mathrm{Sym}^k(V)$: $\quad \mu(v^{\otimes k}) = k\mu(v)$
- $V, W \rightsquigarrow V \otimes W$: $\quad \mu(v \otimes w) = \mu(v) + \mu(w)$
- $W = V_{\boldsymbol{\lambda}}$: $\quad \Delta(v_{\boldsymbol{\lambda}}) = \boldsymbol{\lambda}$

Shifting trick: $V' = \mathrm{Sym}^k(V) \otimes V_{\boldsymbol{\lambda}^*}$ and $v' := v^{\otimes k} \otimes g_0 v_{\boldsymbol{\lambda}^*}$. Then:

$$\frac{\boldsymbol{\lambda}}{k} \in \Delta(v) \quad \Leftrightarrow \quad 0 \in \Delta(v')$$

for generic g.

In special cases: 'elementary' reductions to uniform marginal that only involve change of parameters (e.g., $n \times n$ to $n' \times n'$ matrices).