Energy minimization via moment hierarchies

David de Laat (TU Delft)

ESI Workshop on Optimal Point Configurations and Applications
16 October 2014

Energy minimization

- What is the minimal potential energy E when we put N particles with pair potential h in a container V?

Energy minimization

- What is the minimal potential energy E when we put N particles with pair potential h in a container V?
- Example: For the Thomson problem we take

$$V = S^2 \quad \text{and} \quad h(\{x, y\}) = \frac{1}{\|x - y\|}$$

Energy minimization

- What is the minimal potential energy E when we put N particles with pair potential h in a container V?
- Example: For the Thomson problem we take

$$V = S^2 \quad \text{and} \quad h(\{x, y\}) = \frac{1}{\|x - y\|}$$

- As an optimization problem:

$$E = \min_{S \in \binom{V}{N}} \sum_{P \in \binom{S}{2}} h(P)$$

Approach

- Configurations provide upper bounds on the optimal energy E

Approach

- Configurations provide upper bounds on the optimal energy E
- To prove a configuration is good (or optimal) we need good lower bounds for E

Approach

- Configurations provide upper bounds on the optimal energy E
- To prove a configuration is good (or optimal) we need good lower bounds for E

Some systematic approaches for obtaining bounds:

- Linear programming bounds using the pair correlation function [Delsarte 1973, Delsarte-Goethals-Seidel 1977, Yudin 1992]

Approach

- Configurations provide upper bounds on the optimal energy E
- To prove a configuration is good (or optimal) we need good lower bounds for E

Some systematic approaches for obtaining bounds:

- Linear programming bounds using the pair correlation function [Delsarte 1973, Delsarte-Goethals-Seidel 1977, Yudin 1992]
- 3-point bounds using 3-point correlation functions and constraints arising from the stabilizer subgroup of 1 point [Schrijver 2005, Bachoc-Vallentin 2008, Cohn-Woo 2012]

Approach

- Configurations provide upper bounds on the optimal energy E
- To prove a configuration is good (or optimal) we need good lower bounds for E

Some systematic approaches for obtaining bounds:

- Linear programming bounds using the pair correlation function [Delsarte 1973, Delsarte-Goethals-Seidel 1977, Yudin 1992]
- 3-point bounds using 3-point correlation functions and constraints arising from the stabilizer subgroup of 1 point [Schrijver 2005, Bachoc-Vallentin 2008, Cohn-Woo 2012]
- k-point bounds using stabilizer subgroup of $k-2$ points [Musin 2007]

Approach

- Configurations provide upper bounds on the optimal energy E
- To prove a configuration is good (or optimal) we need good lower bounds for E

Some systematic approaches for obtaining bounds:

- Linear programming bounds using the pair correlation function [Delsarte 1973, Delsarte-Goethals-Seidel 1977, Yudin 1992]
- 3-point bounds using 3-point correlation functions and constraints arising from the stabilizer subgroup of 1 point [Schrijver 2005, Bachoc-Vallentin 2008, Cohn-Woo 2012]
- k-point bounds using stabilizer subgroup of $k-2$ points [Musin 2007]
- Hierarchy for packing problems [L.-Vallentin 2014]

This talk

- Hierarchy obtained by generalizing Lasserre's hierarchy from combinatorial optimization to the continuous setting

This talk

- Hierarchy obtained by generalizing Lasserre's hierarchy from combinatorial optimization to the continuous setting
- Finite convergence to the optimal energy

This talk

- Hierarchy obtained by generalizing Lasserre's hierarchy from combinatorial optimization to the continuous setting
- Finite convergence to the optimal energy
- A duality theory

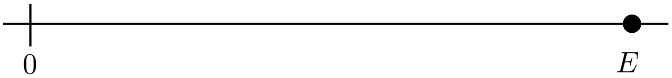
This talk

- Hierarchy obtained by generalizing Lasserre's hierarchy from combinatorial optimization to the continuous setting
- Finite convergence to the optimal energy
- A duality theory
- Reduction to a converging sequence of semidefinite programs

This talk

- Hierarchy obtained by generalizing Lasserre's hierarchy from combinatorial optimization to the continuous setting
- Finite convergence to the optimal energy
- A duality theory
- Reduction to a converging sequence of semidefinite programs
- Towards computations using several types of symmetry reduction

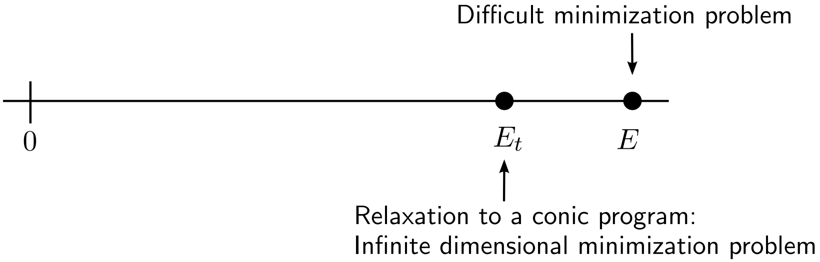
Approach



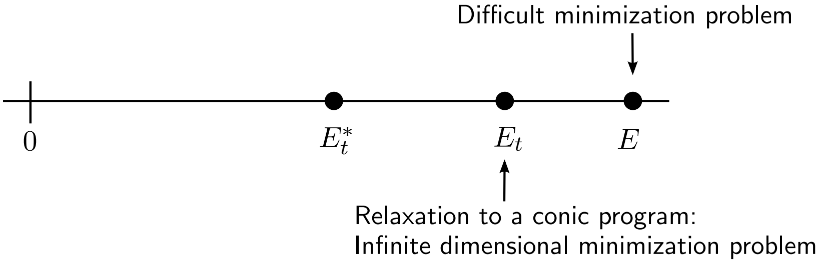
Approach

Approach

Approach

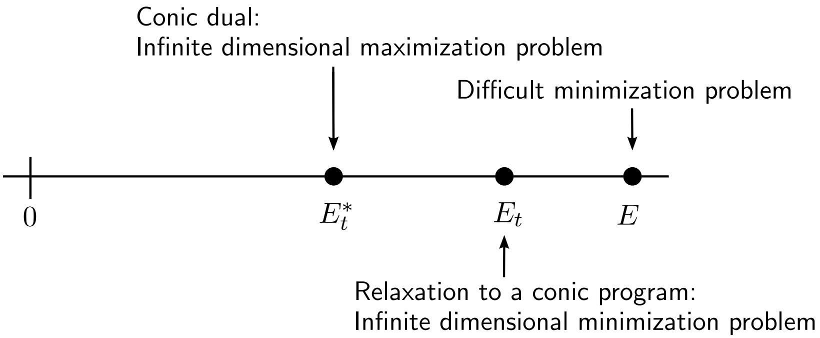

Difficult minimization problem
0
E_t
E
Relaxation to a conic program:
Infinite dimensional minimization problem

Approach



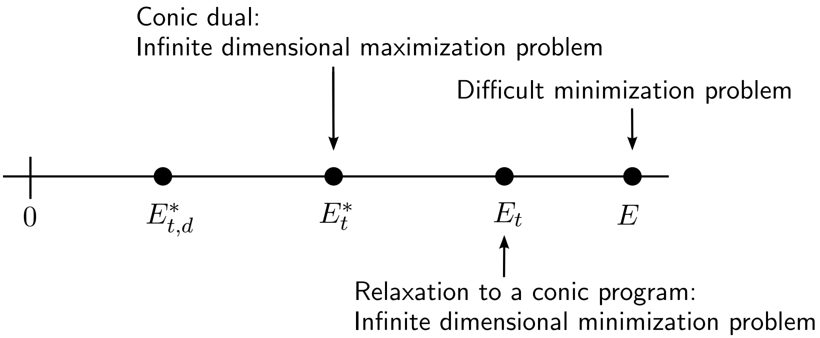
Difficult minimization problem
0
E_t^*
E_t
E
Relaxation to a conic program:
Infinite dimensional minimization problem

Approach

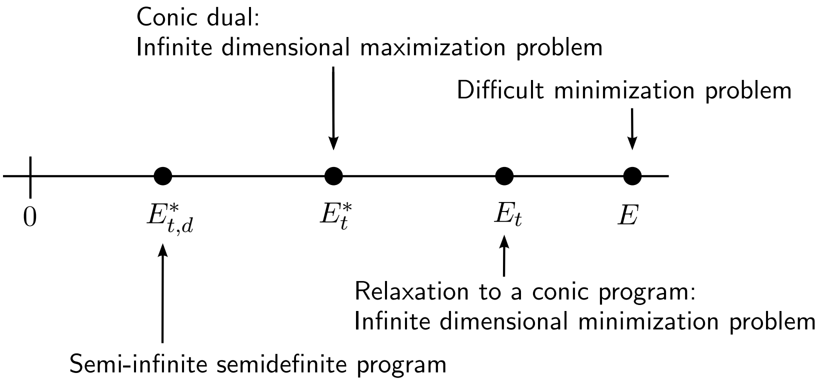


Conic dual:
Infinite dimensional maximization problem
Difficult minimization problem
0
E_t^*
E_t
E
Relaxation to a conic program:
Infinite dimensional minimization problem

Approach


Conic dual:
Infinite dimensional maximization problem
Difficult minimization problem
0
$E^*_{t,d}$
E^*_t
E_t
E
Relaxation to a conic program:
Infinite dimensional minimization problem

Approach


Conic dual:
Infinite dimensional maximization problem
Difficult minimization problem
0
$E^*_{t,d}$
E^*_t
E_t
E
Relaxation to a conic program:
Infinite dimensional minimization problem
Semi-infinite semidefinite program

The minimization problem

- $I_{=t}$ (I_t) is the set of subsets of V which
  - have cardinality t $(\leq t)$
  - contain no points which are too close

The minimization problem

- $I_{=t}$ (I_t) is the set of subsets of V which
  - have cardinality t ($\leq t$)
  - contain no points which are too close
- Assuming $h(\{x, y\}) \to \infty$ when x and y converge, we have

$$E = \min_{S \in I_{=N}} \sum_{P \in \binom{S}{2}} h(P)$$

The minimization problem

- $I_{=t}$ (I_t) is the set of subsets of V which
  - have cardinality t ($\leq t$)
  - contain no points which are too close
- Assuming $h(\{x, y\}) \to \infty$ when x and y converge, we have

$$E = \min_{S \in I_{=N}} \sum_{P \in \binom{S}{2}} h(P)$$

- We will also assume that V is compact and h continuous

The minimization problem

- $I_{=t}$ (I_t) is the set of subsets of V which
  - have cardinality t ($\leq t$)
  - contain no points which are too close
- Assuming $h(\{x,y\}) \to \infty$ when x and y converge, we have

$$E = \min_{S \in I_{=N}} \sum_{P \in \binom{S}{2}} h(P)$$

- We will also assume that V is compact and h continuous
- $I_{=t}$ gets its topology as a subset of a quotient of V^t

Moment hierarchy of relaxations

- In the relaxation E_t we minimize over measures λ on the space I_s, where $s = \min\{2t, N\}$

Moment hierarchy of relaxations

- In the relaxation E_t we minimize over measures λ on the space I_s, where $s = \min\{2t, N\}$

Lemma

When $t = N$, the feasible measures λ are (generalized) convex combinations of measures

$$\chi_S = \sum_{R \subseteq S} \delta_R \quad \text{where} \quad S \in I_{=N}$$

Moment hierarchy of relaxations

- In the relaxation E_t we minimize over measures λ on the space I_s, where $s = \min\{2t, N\}$

Lemma

When $t = N$, the feasible measures λ are (generalized) convex combinations of measures

$$\chi_S = \sum_{R \subseteq S} \delta_R \quad \text{where} \quad S \in I_{=N}$$

- Objective function: $\lambda(h) = \int_{I_{=N}} h(S)\, d\lambda(S)$

Moment hierarchy of relaxations

- In the relaxation E_t we minimize over measures λ on the space I_s, where $s = \min\{2t, N\}$

Lemma

When $t = N$, the feasible measures λ are (generalized) convex combinations of measures

$$\chi_S = \sum_{R \subseteq S} \delta_R \quad \text{where} \quad S \in I_{=N}$$

- Objective function: $\lambda(h) = \int_{I_{=N}} h(S)\, d\lambda(S)$
- Moment constraints: $A_t^* \lambda \in \mathcal{M}(I_t \times I_t)_{\succeq 0}$

Moment hierarchy of relaxations

- In the relaxation E_t we minimize over measures λ on the space I_s, where $s = \min\{2t, N\}$

Lemma

When $t = N$, the feasible measures λ are (generalized) convex combinations of measures

$$\chi_S = \sum_{R \subseteq S} \delta_R \quad \text{where} \quad S \in I_{=N}$$

- Objective function: $\lambda(h) = \int_{I_{=N}} h(S)\, d\lambda(S)$
- Moment constraints: $A_t^* \lambda \in \mathcal{M}(I_t \times I_t)_{\succeq 0}$
- Here A_t^* is an operator $\mathcal{M}(I_s) \to \mathcal{M}(I_t \times I_t)$

Moment hierarchy of relaxations

- In the relaxation E_t we minimize over measures λ on the space I_s, where $s = \min\{2t, N\}$

Lemma

When $t = N$, the feasible measures λ are (generalized) convex combinations of measures

$$\chi_S = \sum_{R \subseteq S} \delta_R \quad \text{where} \quad S \in I_{=N}$$

- Objective function: $\lambda(h) = \int_{I_{=N}} h(S)\, d\lambda(S)$
- Moment constraints: $A_t^* \lambda \in \mathcal{M}(I_t \times I_t)_{\succeq 0}$
- Here A_t^* is an operator $\mathcal{M}(I_s) \to \mathcal{M}(I_t \times I_t)$
- $\mathcal{M}(I_t \times I_t)_{\succeq 0}$ is the cone dual to the cone $\mathcal{C}(I_t \times I_t)_{\succeq 0}$ of positive kernels

Moment hierarchy of relaxations

- In the relaxation E_t we minimize over measures λ on the space I_s, where $s = \min\{2t, N\}$

Lemma

When $t = N$, the feasible measures λ are (generalized) convex combinations of measures

$$\chi_S = \sum_{R \subseteq S} \delta_R \quad \text{where} \quad S \in I_{=N}$$

- Objective function: $\lambda(h) = \int_{I_{=N}} h(S)\, d\lambda(S)$
- Moment constraints: $A_t^* \lambda \in \mathcal{M}(I_t \times I_t)_{\succeq 0}$
- Here A_t^* is an operator $\mathcal{M}(I_s) \to \mathcal{M}(I_t \times I_t)$
- $\mathcal{M}(I_t \times I_t)_{\succeq 0}$ is the cone dual to the cone $\mathcal{C}(I_t \times I_t)_{\succeq 0}$ of positive kernels: $\mu(K) \geq 0$ for all $K \succeq 0$

Moment hierarchy of relaxations

- In the relaxation E_t we minimize over measures λ on the space I_s, where $s = \min\{2t, N\}$

Lemma

When $t = N$, the feasible measures λ are (generalized) convex combinations of measures

$$\chi_S = \sum_{R \subseteq S} \delta_R \quad \text{where} \quad S \in I_{=N}$$

- Objective function: $\lambda(h) = \int_{I_{=N}} h(S)\, d\lambda(S)$
- Moment constraints: $A_t^* \lambda \in \mathcal{M}(I_t \times I_t)_{\succeq 0}$
- Here A_t^* is an operator $\mathcal{M}(I_s) \to \mathcal{M}(I_t \times I_t)$
- $\mathcal{M}(I_t \times I_t)_{\succeq 0}$ is the cone dual to the cone $\mathcal{C}(I_t \times I_t)_{\succeq 0}$ of positive kernels: $\mu(K) \geq 0$ for all $K \succeq 0$
- We have $\chi_S(h) = \sum_{P \in \binom{S}{2}} h(P)$

Moment hierarchy of relaxations

- In the relaxation E_t we minimize over measures λ on the space I_s, where $s = \min\{2t, N\}$

Lemma

When $t = N$, the feasible measures λ are (generalized) convex combinations of measures

$$\chi_S = \sum_{R \subseteq S} \delta_R \quad \text{where} \quad S \in I_{=N}$$

- Objective function: $\lambda(h) = \int_{I_{=N}} h(S)\, d\lambda(S)$
- Moment constraints: $A_t^* \lambda \in \mathcal{M}(I_t \times I_t)_{\succeq 0}$
- Here A_t^* is an operator $\mathcal{M}(I_s) \to \mathcal{M}(I_t \times I_t)$
- $\mathcal{M}(I_t \times I_t)_{\succeq 0}$ is the cone dual to the cone $\mathcal{C}(I_t \times I_t)_{\succeq 0}$ of positive kernels: $\mu(K) \geq 0$ for all $K \succeq 0$
- We have $\chi_S(h) = \sum_{P \in \binom{S}{2}} h(P)$

Theorem (Finite convergence)

We have $E_1 \leq \cdots \leq E_N = E$

Dual hierarchy

- E_t is a minimization problem, so we need an optimal solution to find a lower bound

Dual hierarchy

- E_t is a minimization problem, so we need an optimal solution to find a lower bound
- The conic dual E_t^* is a maximization problem where any feasible solution provides an upper bound

Dual hierarchy

- E_t is a minimization problem, so we need an optimal solution to find a lower bound
- The conic dual E_t^* is a maximization problem where any feasible solution provides an upper bound
- In E_t^* optimization is over scalars $a_i \in \mathbb{R}$ and positive definite kernels $K \in \mathcal{C}(I_t \times I_t)_{\succeq 0}$

Dual hierarchy

- E_t is a minimization problem, so we need an optimal solution to find a lower bound
- The conic dual E_t^* is a maximization problem where any feasible solution provides an upper bound
- In E_t^* optimization is over scalars $a_i \in \mathbb{R}$ and positive definite kernels $K \in \mathcal{C}(I_t \times I_t)_{\succeq 0}$
- The dual program:

$$E_t^* = \sup \Big\{ \sum_{i=0}^{s} \binom{N}{i} a_i : a_0, \ldots, a_s \in \mathbb{R},\ K \in \mathcal{C}(I_t \times I_t)_{\succeq 0}, \\ a_i + A_t K \le h \ \text{ on } \ I_{=i} \ \text{ for } \ i = 0, \ldots, s \Big\}$$

Dual hierarchy

- E_t is a minimization problem, so we need an optimal solution to find a lower bound
- The conic dual E_t^* is a maximization problem where any feasible solution provides an upper bound
- In E_t^* optimization is over scalars $a_i \in \mathbb{R}$ and positive definite kernels $K \in \mathcal{C}(I_t \times I_t)_{\succeq 0}$
- The dual program:

$$E_t^* = \sup \Big\{ \sum_{i=0}^{s} \binom{N}{i} a_i : a_0, \ldots, a_s \in \mathbb{R},\, K \in \mathcal{C}(I_t \times I_t)_{\succeq 0}, \\ a_i + A_t K \leq h \text{ on } I_{=i} \text{ for } i = 0, \ldots, s \Big\}$$

- Here A_t is the linear operator $\mathcal{C}(I_t \times I_t) \to \mathcal{C}(I_t)$ given by $A_t K(S) = \sum_{J, J' \in I_t : J \cup J' = S} K(J, J')$

Dual hierarchy

- E_t is a minimization problem, so we need an optimal solution to find a lower bound
- The conic dual E_t^* is a maximization problem where any feasible solution provides an upper bound
- In E_t^* optimization is over scalars $a_i \in \mathbb{R}$ and positive definite kernels $K \in \mathcal{C}(I_t \times I_t)_{\succeq 0}$
- The dual program:

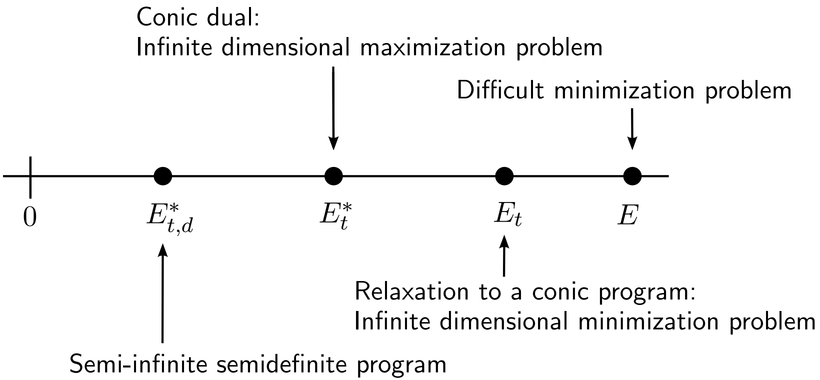
$$E_t^* = \sup \Big\{ \sum_{i=0}^{s} \binom{N}{i} a_i : a_0, \ldots, a_s \in \mathbb{R},\ K \in \mathcal{C}(I_t \times I_t)_{\succeq 0}, \\ a_i + A_t K \leq h \ \text{ on } \ I_{=i} \ \text{ for } \ i = 0, \ldots, s \Big\}$$

- Here A_t is the linear operator $\mathcal{C}(I_t \times I_t) \to \mathcal{C}(I_t)$ given by $A_t K(S) = \sum_{J, J' \in I_t : J \cup J' = S} K(J, J')$

Theorem

Strong duality holds: $E_t = E_t^*$ for each t

Closing the gaps



Finite dimensional approximations to E_t^*

- Define $E_{t,d}^*$ by replacing the cone $\mathcal{C}(I_t \times I_t)_{\succeq 0}$ in E_t^* by a finite dimensional inner approximating cone C_d

Finite dimensional approximations to E_t^*

- Define $E_{t,d}^*$ by replacing the cone $\mathcal{C}(I_t \times I_t)_{\succeq 0}$ in E_t^* by a finite dimensional inner approximating cone C_d
- Let $e_1, e_2, \ldots$ be a dense sequence in $\mathcal{C}(I_t)$ and define

$$C_d = \Big\{ \sum_{i,j=1}^{d} F_{i,j} e_i \otimes e_j : F \in \mathbb{R}^{d \times d} \text{ positive semidefinite} \Big\}$$

Finite dimensional approximations to E_t^*

- Define $E_{t,d}^*$ by replacing the cone $\mathcal{C}(I_t \times I_t)_{\succeq 0}$ in E_t^* by a finite dimensional inner approximating cone C_d
- Let $e_1, e_2, \ldots$ be a dense sequence in $\mathcal{C}(I_t)$ and define

$$C_d = \Big\{ \sum_{i,j=1}^{d} F_{i,j} e_i \otimes e_j : F \in \mathbb{R}^{d\times d} \text{ positive semidefinite} \Big\}$$

Lemma

Suppose X is a compact metric space. Then the extreme rays of the cone $\mathcal{C}(X \times X)_{\succeq 0}$ are precisely the kernels $f \otimes f$ with $f \in \mathcal{C}(X)$

Finite dimensional approximations to E_t^*

- Define $E_{t,d}^*$ by replacing the cone $\mathcal{C}(I_t \times I_t)_{\succeq 0}$ in E_t^* by a finite dimensional inner approximating cone C_d
- Let $e_1, e_2, \ldots$ be a dense sequence in $\mathcal{C}(I_t)$ and define

$$C_d = \Big\{ \sum_{i,j=1}^{d} F_{i,j} e_i \otimes e_j : F \in \mathbb{R}^{d \times d} \text{ positive semidefinite} \Big\}$$

Lemma

Suppose X is a compact metric space. Then the extreme rays of the cone $\mathcal{C}(X \times X)_{\succeq 0}$ are precisely the kernels $f \otimes f$ with $f \in \mathcal{C}(X)$

- This implies $\cup_{d=0}^{\infty} C_d$ is uniformly dense in $\mathcal{C}(I_t \times I_t)_{\succeq 0}$

Finite dimensional approximations to E_t^*

- Define $E_{t,d}^*$ by replacing the cone $\mathcal{C}(I_t \times I_t)_{\succeq 0}$ in E_t^* by a finite dimensional inner approximating cone C_d
- Let $e_1, e_2, \ldots$ be a dense sequence in $\mathcal{C}(I_t)$ and define

$$C_d = \Big\{ \sum_{i,j=1}^{d} F_{i,j} e_i \otimes e_j : F \in \mathbb{R}^{d \times d} \text{ positive semidefinite} \Big\}$$

Lemma

Suppose X is a compact metric space. Then the extreme rays of the cone $\mathcal{C}(X \times X)_{\succeq 0}$ are precisely the kernels $f \otimes f$ with $f \in \mathcal{C}(X)$

- This implies $\cup_{d=0}^{\infty} C_d$ is uniformly dense in $\mathcal{C}(I_t \times I_t)_{\succeq 0}$

Theorem

If V is a compact metric space, then $E_{t,d}^* \to E_t^*$ as $d \to \infty$ for all t

Block diagonalization

- For computations use the symmetry of V and h, expressed by the action of a group Γ, and Bochner's theorem to block diagonalize the matrix F

Block diagonalization

- For computations use the symmetry of V and h, expressed by the action of a group Γ, and Bochner's theorem to block diagonalize the matrix F
- For this we need a symmetry adapted basis of $\mathcal{C}(I_t)$

Block diagonalization

- For computations use the symmetry of V and h, expressed by the action of a group Γ, and Bochner's theorem to block diagonalize the matrix F
- For this we need a symmetry adapted basis of $\mathcal{C}(I_t)$
- If $t = 1$ and $V = S^2$, then

$$\mathcal{C}(I_t) \simeq \mathbb{R} \oplus \mathcal{C}(S^2)$$

Block diagonalization

- For computations use the symmetry of V and h, expressed by the action of a group Γ, and Bochner's theorem to block diagonalize the matrix F
- For this we need a symmetry adapted basis of $\mathcal{C}(I_t)$
- If $t = 1$ and $V = S^2$, then

$$\mathcal{C}(I_t) \simeq \mathbb{R} \oplus \mathcal{C}(S^2) = \mathbb{R} \oplus \bigoplus_{k=0}^{\infty} H_k$$

Block diagonalization

- For computations use the symmetry of V and h, expressed by the action of a group Γ, and Bochner's theorem to block diagonalize the matrix F
- For this we need a symmetry adapted basis of $\mathcal{C}(I_t)$
- If $t = 1$ and $V = S^2$, then

$$\mathcal{C}(I_t) \simeq \mathbb{R} \oplus \mathcal{C}(S^2) = \mathbb{R} \oplus \bigoplus_{k=0}^{\infty} H_k$$

- This will block diagonalize to a diagonal matrix and we get (something close to) Yudin's LP bound

Block diagonalization

- For computations use the symmetry of V and h, expressed by the action of a group Γ, and Bochner's theorem to block diagonalize the matrix F
- For this we need a symmetry adapted basis of $\mathcal{C}(I_t)$
- If $t = 1$ and $V = S^2$, then

$$\mathcal{C}(I_t) \simeq \mathbb{R} \oplus \mathcal{C}(S^2) = \mathbb{R} \oplus \bigoplus_{k=0}^{\infty} H_k$$

- This will block diagonalize to a diagonal matrix and we get (something close to) Yudin's LP bound
- In general $\mathcal{C}(I_t)$ injects into $\mathcal{C}(V)^{\odot t}$

Block diagonalization

- For computations use the symmetry of V and h, expressed by the action of a group Γ, and Bochner's theorem to block diagonalize the matrix F
- For this we need a symmetry adapted basis of $\mathcal{C}(I_t)$
- If $t = 1$ and $V = S^2$, then

$$\mathcal{C}(I_t) \simeq \mathbb{R} \oplus \mathcal{C}(S^2) = \mathbb{R} \oplus \bigoplus_{k=0}^{\infty} H_k$$

- This will block diagonalize to a diagonal matrix and we get (something close to) Yudin's LP bound
- In general $\mathcal{C}(I_t)$ injects into $\mathcal{C}(V)^{\odot t}$
- $\mathcal{C}(V)^{\odot t}$ can be written in terms of tensor products of the irreducible subspaces of $\mathcal{C}(V)$

Block diagonalization

- For computations use the symmetry of V and h, expressed by the action of a group Γ, and Bochner's theorem to block diagonalize the matrix F
- For this we need a symmetry adapted basis of $\mathcal{C}(I_t)$
- If $t = 1$ and $V = S^2$, then

$$\mathcal{C}(I_t) \simeq \mathbb{R} \oplus \mathcal{C}(S^2) = \mathbb{R} \oplus \bigoplus_{k=0}^{\infty} H_k$$

- This will block diagonalize to a diagonal matrix and we get (something close to) Yudin's LP bound
- In general $\mathcal{C}(I_t)$ injects into $\mathcal{C}(V)^{\odot t}$
- $\mathcal{C}(V)^{\odot t}$ can be written in terms of tensor products of the irreducible subspaces of $\mathcal{C}(V)$
- If we know how to decompose $\mathcal{C}(V)$ into irreducibles, and how to decompose tensor products of those irreducibles into irreducibles, then we have a symmerty adapted basis of V_t

The case $t = 2$ and $V = S^2$

- We know how to these decompositions from the quantum mechanics literature (angular momentum coupling): use Clebsch-Gordan coefficients

The case $t = 2$ and $V = S^2$

- We know how to these decompositions from the quantum mechanics literature (angular momentum coupling): use Clebsch-Gordan coefficients
- The affine constraints in $E_{t,d}^*$ are nonnegativity constraints of a polynomial $p \in \mathbb{R}[x_1, \ldots, x_4]$, where each x_i is a vector of 3 variables (the coefficients of these polynomials depend on the entries in the block diagonalization of F)

The case $t = 2$ and $V = S^2$

- We know how to these decompositions from the quantum mechanics literature (angular momentum coupling): use Clebsch-Gordan coefficients
- The affine constraints in $E_{t,d}^*$ are nonnegativity constraints of a polynomial $p \in \mathbb{R}[x_1, \ldots, x_4]$, where each x_i is a vector of 3 variables (the coefficients of these polynomials depend on the entries in the block diagonalization of F)
- We have $p(\gamma x_1, \ldots, \gamma x_4) = p(x_1, \ldots, x_4)$ for all $\gamma \in O(3)$

The case $t = 2$ and $V = S^2$

- We know how to these decompositions from the quantum mechanics literature (angular momentum coupling): use Clebsch-Gordan coefficients
- The affine constraints in $E^*_{t,d}$ are nonnegativity constraints of a polynomial $p \in \mathbb{R}[x_1, \ldots, x_4]$, where each x_i is a vector of 3 variables (the coefficients of these polynomials depend on the entries in the block diagonalization of F)
- We have $p(\gamma x_1, \ldots, \gamma x_4) = p(x_1, \ldots, x_4)$ for all $\gamma \in O(3)$
- Invariant theory: there is a polynomial q such that $p(x_1, \ldots, x_4) = q(x_1 \cdot x_2, \ldots, x_3 \cdot x_4)$

The case $t = 2$ and $V = S^2$

- We know how to these decompositions from the quantum mechanics literature (angular momentum coupling): use Clebsch-Gordan coefficients
- The affine constraints in $E_{t,d}^*$ are nonnegativity constraints of a polynomial $p \in \mathbb{R}[x_1, \ldots, x_4]$, where each x_i is a vector of 3 variables (the coefficients of these polynomials depend on the entries in the block diagonalization of F)
- We have $p(\gamma x_1, \ldots, \gamma x_4) = p(x_1, \ldots, x_4)$ for all $\gamma \in O(3)$
- Invariant theory: there is a polynomial q such that $p(x_1, \ldots, x_4) = q(x_1 \cdot x_2, \ldots, x_3 \cdot x_4)$
- Model nonnegativity constraints as sum of squares constraints using Putinar's theorem from real algebraic geometry

The case $t = 2$ and $V = S^2$

- We know how to these decompositions from the quantum mechanics literature (angular momentum coupling): use Clebsch-Gordan coefficients
- The affine constraints in $E_{t,d}^*$ are nonnegativity constraints of a polynomial $p \in \mathbb{R}[x_1, \ldots, x_4]$, where each x_i is a vector of 3 variables (the coefficients of these polynomials depend on the entries in the block diagonalization of F)
- We have $p(\gamma x_1, \ldots, \gamma x_4) = p(x_1, \ldots, x_4)$ for all $\gamma \in O(3)$
- Invariant theory: there is a polynomial q such that $p(x_1, \ldots, x_4) = q(x_1 \cdot x_2, \ldots, x_3 \cdot x_4)$
- Model nonnegativity constraints as sum of squares constraints using Putinar's theorem from real algebraic geometry
- A sum of squares polynomial s can be written as $s(x) = v(x)^{\mathsf{T}} Q v(x)$, where Q is a positive semidefinite matrix and $v(x)$ a vector containing all monomials up to some degree

More symmetry

- More symmetry: $p(x_1, \ldots, x_4) = p(x_{\sigma(1)}, \ldots, x_{\sigma(4)})$ for all permutations $\sigma \in S_4$

More symmetry

- More symmetry: $p(x_1, \ldots, x_4) = p(x_{\sigma(1)}, \ldots, x_{\sigma(4)})$ for all permutations $\sigma \in S_4$
- This means that q is symmetric under a subgroup of S_6

More symmetry

- More symmetry: $p(x_1, \ldots, x_4) = p(x_{\sigma(1)}, \ldots, x_{\sigma(4)})$ for all permutations $\sigma \in S_4$
- This means that q is symmetric under a subgroup of S_6
- Use this to block diagonalize the positive semidefinite matrices showing up in the sums of squares characterizations

More symmetry

- More symmetry: $p(x_1, \ldots, x_4) = p(x_{\sigma(1)}, \ldots, x_{\sigma(4)})$ for all permutations $\sigma \in S_4$
- This means that q is symmetric under a subgroup of S_6
- Use this to block diagonalize the positive semidefinite matrices showing up in the sums of squares characterizations
- We give a symmetrized version of Putinar's theorem using the method of Gatermann and Parillo for symmetry reduction in sums of squares characterizations

More symmetry

- More symmetry: $p(x_1, \ldots, x_4) = p(x_{\sigma(1)}, \ldots, x_{\sigma(4)})$ for all permutations $\sigma \in S_4$
- This means that q is symmetric under a subgroup of S_6
- Use this to block diagonalize the positive semidefinite matrices showing up in the sums of squares characterizations
- We give a symmetrized version of Putinar's theorem using the method of Gatermann and Parillo for symmetry reduction in sums of squares characterizations
- Significant simplifications in the semidefinite programs

More symmetry

- More symmetry: $p(x_1, \ldots, x_4) = p(x_{\sigma(1)}, \ldots, x_{\sigma(4)})$ for all permutations $\sigma \in S_4$
- This means that q is symmetric under a subgroup of S_6
- Use this to block diagonalize the positive semidefinite matrices showing up in the sums of squares characterizations
- We give a symmetrized version of Putinar's theorem using the method of Gatermann and Parillo for symmetry reduction in sums of squares characterizations
- Significant simplifications in the semidefinite programs
- Not clear yet whether we can compute $E_{2,d}^*$ for large enough d (with current SDP solvers) to get improved bounds for S^2

More symmetry

- More symmetry: $p(x_1, \ldots, x_4) = p(x_{\sigma(1)}, \ldots, x_{\sigma(4)})$ for all permutations $\sigma \in S_4$
- This means that q is symmetric under a subgroup of S_6
- Use this to block diagonalize the positive semidefinite matrices showing up in the sums of squares characterizations
- We give a symmetrized version of Putinar's theorem using the method of Gatermann and Parillo for symmetry reduction in sums of squares characterizations
- Significant simplifications in the semidefinite programs
- Not clear yet whether we can compute $E_{2,d}^*$ for large enough d (with current SDP solvers) to get improved bounds for S^2
- Toy example: E_1 is not sharp for 3 points on S^1 with the Lennard-Jones potential

More symmetry

- More symmetry: $p(x_1, \ldots, x_4) = p(x_{\sigma(1)}, \ldots, x_{\sigma(4)})$ for all permutations $\sigma \in S_4$
- This means that q is symmetric under a subgroup of S_6
- Use this to block diagonalize the positive semidefinite matrices showing up in the sums of squares characterizations
- We give a symmetrized version of Putinar's theorem using the method of Gatermann and Parillo for symmetry reduction in sums of squares characterizations
- Significant simplifications in the semidefinite programs
- Not clear yet whether we can compute $E_{2,d}^*$ for large enough d (with current SDP solvers) to get improved bounds for S^2
- Toy example: E_1 is not sharp for 3 points on S^1 with the Lennard-Jones potential
- Using a reduction to 3 variables using trigonometric polynomials we compute that $E_2 = E$ (up to solver precision)

Thank you!