# High precision computations for energy minimization

David de Laat (CWI Amsterdam)

Real algebraic geometry with a view toward moment problems and optimization, 6 March 2017, MFO

# Energy minimization

- Problem: Find the ground state energy of a system of $N$ particles in a compact metric space $(V, d)$ with pair potential $h$

# Energy minimization

- Problem: Find the ground state energy of a system of $N$ particles in a compact metric space $(V, d)$ with pair potential $h$
- Example: In the *Thomson problem* we minimize

$$\sum_{1 \leq i < j \leq N} \frac{1}{\|x_i - x_j\|_2}$$

over all sets $\{x_1, \ldots, x_N\}$ of $N$ distinct points in $S^2 \subseteq \mathbb{R}^3$

# Energy minimization

- Problem: Find the ground state energy of a system of $N$ particles in a compact metric space $(V, d)$ with pair potential $h$
- Example: In the *Thomson problem* we minimize

$$\sum_{1 \leq i < j \leq N} \frac{1}{\|x_i - x_j\|_2}$$

  over all sets $\{x_1, \ldots, x_N\}$ of $N$ distinct points in $S^2 \subseteq \mathbb{R}^3$
- Here $V = S^2$, $d(x, y) = \|x - y\|_2$, and $h(w) = 1/w$

# Energy minimization

- Problem: Find the ground state energy of a system of $N$ particles in a compact metric space $(V, d)$ with pair potential $h$
- Example: In the *Thomson problem* we minimize

$$\sum_{1 \leq i < j \leq N} \frac{1}{\|x_i - x_j\|_2}$$

  over all sets $\{x_1, \ldots, x_N\}$ of $N$ distinct points in $S^2 \subseteq \mathbb{R}^3$
- Here $V = S^2$, $d(x, y) = \|x - y\|_2$, and $h(w) = 1/w$
- Assume $h(w) \to \infty$ as $w \to 0$

# Energy minimization

- Problem: Find the ground state energy of a system of $N$ particles in a compact metric space $(V, d)$ with pair potential $h$
- Example: In the *Thomson problem* we minimize

  $$\sum_{1 \leq i < j \leq N} \frac{1}{\|x_i - x_j\|_2}$$

  over all sets $\{x_1, \ldots, x_N\}$ of $N$ distinct points in $S^2 \subseteq \mathbb{R}^3$
- Here $V = S^2$, $d(x, y) = \|x - y\|_2$, and $h(w) = 1/w$
- Assume $h(w) \to \infty$ as $w \to 0$
- Use moment techniques to find lower bounds (obstructions)

# Energy minimization

- Problem: Find the ground state energy of a system of $N$ particles in a compact metric space $(V, d)$ with pair potential $h$
- Example: In the *Thomson problem* we minimize

$$\sum_{1 \leq i < j \leq N} \frac{1}{\|x_i - x_j\|_2}$$

  over all sets $\{x_1, \ldots, x_N\}$ of $N$ distinct points in $S^2 \subseteq \mathbb{R}^3$
- Here $V = S^2$, $d(x, y) = \|x - y\|_2$, and $h(w) = 1/w$
- Assume $h(w) \to \infty$ as $w \to 0$
- Use moment techniques to find lower bounds (obstructions)
- Infinite dimensional moment techniques $\to$ computations

# Energy minimization

- Problem: Find the ground state energy of a system of $N$ particles in a compact metric space $(V, d)$ with pair potential $h$
- Example: In the *Thomson problem* we minimize

$$\sum_{1 \leq i < j \leq N} \frac{1}{\|x_i - x_j\|_2}$$

  over all sets $\{x_1, \ldots, x_N\}$ of $N$ distinct points in $S^2 \subseteq \mathbb{R}^3$
- Here $V = S^2$, $d(x, y) = \|x - y\|_2$, and $h(w) = 1/w$
- Assume $h(w) \to \infty$ as $w \to 0$
- Use moment techniques to find lower bounds (obstructions)
- Infinite dimensional moment techniques $\to$ computations (Compute sharp lower bound for the $N = 5$ case)

# Setup

- Let $B$ be an upper bound on the minimal energy

# Setup

- Let $B$ be an upper bound on the minimal energy
- Define a graph with vertex set $V$ where two distinct vertices $x$ and $y$ are adjacent if $h(d(x, y)) > B$

# Setup

- Let $B$ be an upper bound on the minimal energy
- Define a graph with vertex set $V$ where two distinct vertices $x$ and $y$ are adjacent if $h(d(x, y)) > B$
- Let $I_t$ be the set of independent sets with $\leq t$ elements

# Setup

- Let $B$ be an upper bound on the minimal energy
- Define a graph with vertex set $V$ where two distinct vertices $x$ and $y$ are adjacent if $h(d(x, y)) > B$
- Let $I_t$ be the set of independent sets with $\leq t$ elements
- Let $I_{=t}$ be the set of independent sets with $t$ elements

# Setup

- Let $B$ be an upper bound on the minimal energy
- Define a graph with vertex set $V$ where two distinct vertices $x$ and $y$ are adjacent if $h(d(x, y)) > B$
- Let $I_t$ be the set of independent sets with $\leq t$ elements
- Let $I_{=t}$ be the set of independent sets with $t$ elements
- These sets are compact metric spaces

# Setup

- Let $B$ be an upper bound on the minimal energy
- Define a graph with vertex set $V$ where two distinct vertices $x$ and $y$ are adjacent if $h(d(x,y)) > B$
- Let $I_t$ be the set of independent sets with $\leq t$ elements
- Let $I_{=t}$ be the set of independent sets with $t$ elements
- These sets are compact metric spaces
- Define $f \in \mathcal{C}(I_N)$ by

$$f(S) = \begin{cases} h(d(x,y)) & \text{if } S = \{x,y\} \text{ with } x \neq y, \\ 0 & \text{otherwise} \end{cases}$$

# Setup

- Let $B$ be an upper bound on the minimal energy
- Define a graph with vertex set $V$ where two distinct vertices $x$ and $y$ are adjacent if $h(d(x, y)) > B$
- Let $I_t$ be the set of independent sets with $\leq t$ elements
- Let $I_{=t}$ be the set of independent sets with $t$ elements
- These sets are compact metric spaces
- Define $f \in \mathcal{C}(I_N)$ by

$$f(S) = \begin{cases} h(d(x, y)) & \text{if } S = \{x, y\} \text{ with } x \neq y, \\ 0 & \text{otherwise} \end{cases}$$

- Ground state energy:

$$E = \min_{S \in I_{=N}} \sum_{P \subseteq S} f(P)$$

# Moment relaxations

- For $S \in I_{=N}$, define the measure $\chi_S = \sum_{R \subseteq S} \delta_R$ on $I_N$

# Moment relaxations

- For $S \in I_{=N}$, define the measure $\chi_S = \sum_{R \subseteq S} \delta_R$ on $I_N$
- We can use this measure to compute the energy of $S$

# Moment relaxations

- For $S \in I_{=N}$, define the measure $\chi_S = \sum_{R \subseteq S} \delta_R$ on $I_N$
- We can use this measure to compute the energy of $S$
- The energy of $S$ is given by

$$\chi_S(f) = \int f(P)\, d\chi_S(P) = \sum_{R \subseteq S} f(R) = \sum_{\{x,y\} \in I_{=2}} h(d(x,y))$$

# Moment relaxations

- For $S \in I_{=N}$, define the measure $\chi_S = \sum_{R \subseteq S} \delta_R$ on $I_N$
- We can use this measure to compute the energy of $S$
- The energy of $S$ is given by

$$\chi_S(f) = \int f(P)\, d\chi_S(P) = \sum_{R \subseteq S} f(R) = \sum_{\{x,y\} \in I_{=2}} h(d(x,y))$$

- This measure satisfies the following $3$ properties:

# Moment relaxations

- For $S \in I_{=N}$, define the measure $\chi_S = \sum_{R \subseteq S} \delta_R$ on $I_N$
- We can use this measure to compute the energy of $S$
- The energy of $S$ is given by

$$\chi_S(f) = \int f(P)\, d\chi_S(P) = \sum_{R \subseteq S} f(R) = \sum_{\{x,y\} \in I_{=2}} h(d(x,y))$$

- This measure satisfies the following $3$ properties:
  - $\chi_S$ is a positive measure

# Moment relaxations

- For $S \in I_{=N}$, define the measure $\chi_S = \sum_{R \subseteq S} \delta_R$ on $I_N$
- We can use this measure to compute the energy of $S$
- The energy of $S$ is given by

$$\chi_S(f) = \int f(P)\, d\chi_S(P) = \sum_{R \subseteq S} f(R) = \sum_{\{x,y\} \in I_{=2}} h(d(x,y))$$

- This measure satisfies the following $3$ properties:
  - $\chi_S$ is a positive measure
  - $\chi_S$ satisfies $\chi_S(I_{=i}) = \binom{N}{i}$ for all $i$

# Moment relaxations

- For $S \in I_{=N}$, define the measure $\chi_S = \sum_{R \subseteq S} \delta_R$ on $I_N$
- We can use this measure to compute the energy of $S$
- The energy of $S$ is given by

$$\chi_S(f) = \int f(P)\, d\chi_S(P) = \sum_{R \subseteq S} f(R) = \sum_{\{x,y\} \in I_{=2}} h(d(x,y))$$

- This measure satisfies the following $3$ properties:
  - $\chi_S$ is a positive measure
  - $\chi_S$ satisfies $\chi_S(I_{=i}) = \binom{N}{i}$ for all $i$
  - $\chi_S$ is a measure of *positive type* (see next slide)

# Moment relaxations

- For $S \in I_{=N}$, define the measure $\chi_S = \sum_{R \subseteq S} \delta_R$ on $I_N$
- We can use this measure to compute the energy of $S$
- The energy of $S$ is given by

$$\chi_S(f) = \int f(P)\, d\chi_S(P) = \sum_{R \subseteq S} f(R) = \sum_{\{x,y\} \in I_{=2}} h(d(x,y))$$

- This measure satisfies the following $3$ properties:
  - $\chi_S$ is a positive measure
  - $\chi_S$ satisfies $\chi_S(I_{=i}) = \binom{N}{i}$ for all $i$
  - $\chi_S$ is a measure of *positive type* (see next slide)
- Relaxations:

$$E_t = \min \Big\{ \lambda(f) : \lambda \in \mathcal{M}(I_{2t}) \text{ positive measure of positive type,} \\ \lambda(I_{=i}) = \binom{N}{i} \text{ for all } 0 \leq i \leq 2t \Big\}$$

# Moment relaxations

- For $S \in I_{=N}$, define the measure $\chi_S = \sum_{R \subseteq S} \delta_R$ on $I_N$
- We can use this measure to compute the energy of $S$
- The energy of $S$ is given by

$$\chi_S(f) = \int f(P)\, d\chi_S(P) = \sum_{R \subseteq S} f(R) = \sum_{\{x,y\} \in I_{=2}} h(d(x,y))$$

- This measure satisfies the following $3$ properties:
  - $\chi_S$ is a positive measure
  - $\chi_S$ satisfies $\chi_S(I_{=i}) = \binom{N}{i}$ for all $i$
  - $\chi_S$ is a measure of *positive type* (see next slide)
- Relaxations:

$$E_t = \min \Big\{ \lambda(f) : \lambda \in \mathcal{M}(I_{2t}) \text{ positive measure of positive type,} \\ \lambda(I_{=i}) = \binom{N}{i} \text{ for all } 0 \leq i \leq 2t \Big\}$$

- $E_t$ is a $\min\{2t, N\}$-point bound

# Moment relaxations

- For $S \in I_{=N}$, define the measure $\chi_S = \sum_{R \subseteq S} \delta_R$ on $I_N$
- We can use this measure to compute the energy of $S$
- The energy of $S$ is given by

$$\chi_S(f) = \int f(P)\, d\chi_S(P) = \sum_{R \subseteq S} f(R) = \sum_{\{x,y\} \in I_{=2}} h(d(x,y))$$

- This measure satisfies the following $3$ properties:
  - $\chi_S$ is a positive measure
  - $\chi_S$ satisfies $\chi_S(I_{=i}) = \binom{N}{i}$ for all $i$
  - $\chi_S$ is a measure of *positive type* (see next slide)
- Relaxations:

$$E_t = \min \Big\{ \lambda(f) : \lambda \in \mathcal{M}(I_{2t}) \text{ positive measure of positive type}, \\ \lambda(I_{=i}) = \binom{N}{i} \text{ for all } 0 \leq i \leq 2t \Big\}$$

- $E_t$ is a $\min\{2t, N\}$-point bound

$$E_1 \leq E_2 \leq \cdots \leq E_N = E$$

# Measures of positive type [L–Vallentin 2015]

- Operator:

$$A_t \colon \mathcal{C}(I_t \times I_t)_{\mathrm{sym}} \to \mathcal{C}(I_{2t}),\; A_t K(S) = \sum_{J,J' \in I_t : J \cup J' = S} K(J,J')$$

# Measures of positive type [L–Vallentin 2015]

- Operator:

$$A_t \colon \mathcal{C}(I_t \times I_t)_{\mathrm{sym}} \to \mathcal{C}(I_{2t}),\; A_t K(S) = \sum_{J,J' \in I_t : J \cup J' = S} K(J, J')$$

- This is an infinite dimensional version of the adjoint of the opererator $y \mapsto M(y)$ that maps a moment sequence to a moment matrix

# Measures of positive type [L–Vallentin 2015]

- Operator:

$$A_t \colon \mathcal{C}(I_t \times I_t)_{\mathrm{sym}} \to \mathcal{C}(I_{2t}),\; A_t K(S) = \sum_{J,J' \in I_t : J \cup J' = S} K(J, J')$$

- This is an infinite dimensional version of the adjoint of the opererator $y \mapsto M(y)$ that maps a moment sequence to a moment matrix
- Dual operator

$$A_t^* \colon \mathcal{M}(I_{2t}) \to \mathcal{M}(I_t \times I_t)_{\mathrm{sym}}$$

# Measures of positive type [L–Vallentin 2015]

- Operator:

$$A_t \colon \mathcal{C}(I_t \times I_t)_{\mathrm{sym}} \to \mathcal{C}(I_{2t}),\; A_t K(S) = \sum_{J,J' \in I_t : J \cup J' = S} K(J, J')$$

- This is an infinite dimensional version of the adjoint of the opererator $y \mapsto M(y)$ that maps a moment sequence to a moment matrix
- Dual operator

$$A_t^* \colon \mathcal{M}(I_{2t}) \to \mathcal{M}(I_t \times I_t)_{\mathrm{sym}}$$

- Cone of positive definite kernels: $\mathcal{C}(I_t \times I_t)_{\succeq 0}$

# Measures of positive type [L–Vallentin 2015]

- Operator:

$$A_t \colon \mathcal{C}(I_t \times I_t)_{\mathrm{sym}} \to \mathcal{C}(I_{2t}),\ A_t K(S) = \sum_{J,J' \in I_t : J \cup J' = S} K(J, J')$$

- This is an infinite dimensional version of the adjoint of the opererator $y \mapsto M(y)$ that maps a moment sequence to a moment matrix
- Dual operator

$$A_t^* \colon \mathcal{M}(I_{2t}) \to \mathcal{M}(I_t \times I_t)_{\mathrm{sym}}$$

- Cone of positive definite kernels: $\mathcal{C}(I_t \times I_t)_{\succeq 0}$
- Dual cone:

$$\mathcal{M}(I_t \times I_t)_{\succeq 0} = \{\mu \in \mathcal{M}(I_t \times I_t)_{\mathrm{sym}} : \mu(K) \geq 0 \text{ for all } K \in \mathcal{C}(I_t \times I_t)_{\succeq 0}\}$$

# Measures of positive type [L–Vallentin 2015]

- Operator:

$$A_t \colon \mathcal{C}(I_t \times I_t)_{\mathrm{sym}} \to \mathcal{C}(I_{2t}),\ A_t K(S) = \sum_{J,J' \in I_t : J \cup J' = S} K(J,J')$$

- This is an infinite dimensional version of the adjoint of the opererator $y \mapsto M(y)$ that maps a moment sequence to a moment matrix
- Dual operator

$$A_t^* \colon \mathcal{M}(I_{2t}) \to \mathcal{M}(I_t \times I_t)_{\mathrm{sym}}$$

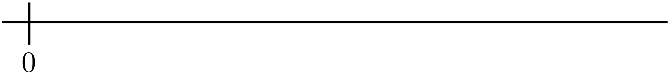
- Cone of positive definite kernels: $\mathcal{C}(I_t \times I_t)_{\succeq 0}$
- Dual cone:

$$\mathcal{M}(I_t \times I_t)_{\succeq 0} = \{\mu \in \mathcal{M}(I_t \times I_t)_{\mathrm{sym}} : \mu(K) \geq 0 \text{ for all } K \in \mathcal{C}(I_t \times I_t)_{\succeq 0}\}$$

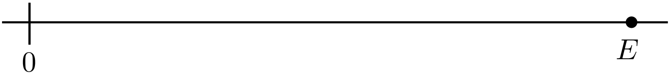
- A measure $\lambda \in \mathcal{M}(I_{2t})$ is of *positive type* if

$$A_t^* \lambda \in \mathcal{M}(I_t \times I_t)_{\succeq 0}$$

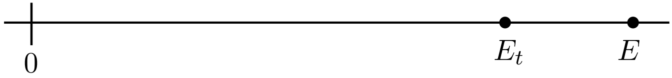
# The dual hierarchy



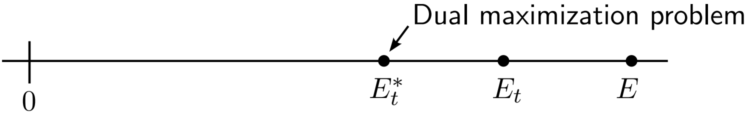
# The dual hierarchy



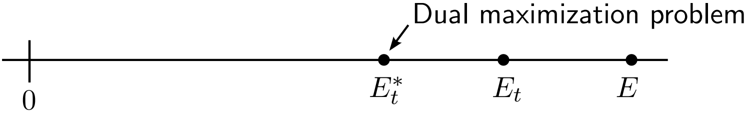
# The dual hierarchy



# The dual hierarchy

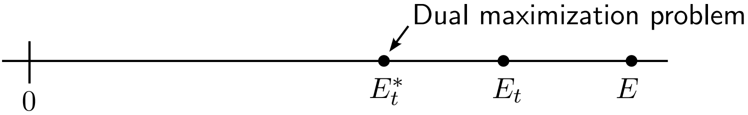


# The dual hierarchy




Strong duality holds: $E_t = E_t^*$

# The dual hierarchy

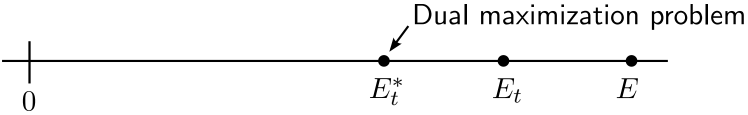



Strong duality holds: $E_t = E_t^*$

- In $E_t^*$ we optimize over kernels $K \in \mathcal{C}(I_t \times I_t)_{\succeq 0}$:

$$E_t^* = \sup \Big\{ \sum_{i=0}^{2t} \binom{N}{i} a_i : a \in \mathbb{R}^{\{0,\ldots,2t\}},\ K \in \mathcal{C}(I_t \times I_t)_{\succeq 0}, \\ a_i + A_t K(S) \leq f(S) \\ \text{for } S \in I_{=i} \text{ and } i = 0, \ldots, 2t \Big\},$$

# The dual hierarchy



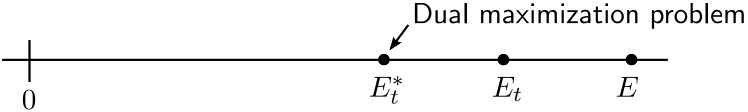

Strong duality holds: $E_t = E_t^*$

- In $E_t^*$ we optimize over kernels $K \in \mathcal{C}(I_t \times I_t)_{\succeq 0}$:

$$E_t^* = \sup \Big\{ \sum_{i=0}^{2t} \binom{N}{i} a_i : a \in \mathbb{R}^{\{0,\ldots,2t\}},\, K \in \mathcal{C}(I_t \times I_t)_{\succeq 0}, \\ a_i + A_t K(S) \le f(S) \\ \text{for } S \in I_{=i} \text{ and } i = 0, \ldots, 2t \Big\},$$

- Reduce to finite dimensional variable space:

# The dual hierarchy



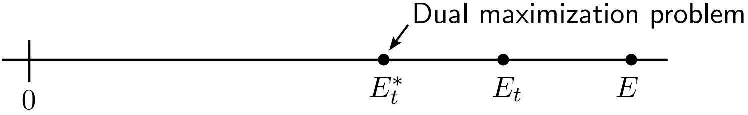

Strong duality holds: $E_t = E_t^*$

- In $E_t^*$ we optimize over kernels $K \in \mathcal{C}(I_t \times I_t)_{\succeq 0}$:

$$E_t^* = \sup \Big\{ \sum_{i=0}^{2t} \binom{N}{i} a_i : a \in \mathbb{R}^{\{0,\dots,2t\}},\, K \in \mathcal{C}(I_t \times I_t)_{\succeq 0}, \\ a_i + A_t K(S) \le f(S) \\ \text{for } S \in I_{=i} \text{ and } i = 0, \dots, 2t \Big\},$$

- Reduce to finite dimensional variable space:
  1. Express $K$ in terms of its Fourier coefficients

# The dual hierarchy



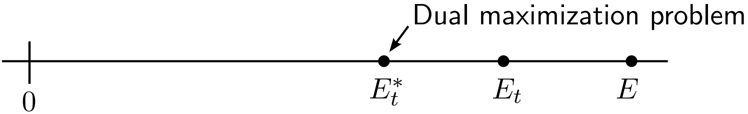

Strong duality holds: $E_t = E_t^*$

- In $E_t^*$ we optimize over kernels $K \in \mathcal{C}(I_t \times I_t)_{\succeq 0}$:

$$E_t^* = \sup\Big\{ \sum_{i=0}^{2t} \binom{N}{i} a_i : a \in \mathbb{R}^{\{0,\ldots,2t\}},\ K \in \mathcal{C}(I_t \times I_t)_{\succeq 0}, \\ a_i + A_t K(S) \leq f(S) \\ \text{for } S \in I_{=i} \text{ and } i = 0,\ldots,2t \Big\},$$

- Reduce to finite dimensional variable space:
  1. Express $K$ in terms of its Fourier coefficients
  2. Set all but finitely many of these coefficients to $0$

# The dual hierarchy




Strong duality holds: $E_t = E_t^*$

- In $E_t^*$ we optimize over kernels $K \in \mathcal{C}(I_t \times I_t)_{\succeq 0}$:

$$E_t^* = \sup \Big\{ \sum_{i=0}^{2t} \binom{N}{i} a_i : a \in \mathbb{R}^{\{0,\ldots,2t\}},\ K \in \mathcal{C}(I_t \times I_t)_{\succeq 0}, \\ a_i + A_t K(S) \leq f(S) \\ \text{for } S \in I_{=i} \text{ and } i = 0, \ldots, 2t \Big\},$$

- Reduce to finite dimensional variable space:
  1. Express $K$ in terms of its Fourier coefficients
  2. Set all but finitely many of these coefficients to $0$
  3. Optimize over the remaining coefficients

# Harmonic analysis on subset spaces

- Let $\Gamma$ be compact group with an action on $V$

# Harmonic analysis on subset spaces

- Let $\Gamma$ be compact group with an action on $V$
- Example: $\Gamma = O(3)$ and $V = S^2 \subseteq \mathbb{R}^3$

# Harmonic analysis on subset spaces

- Let $\Gamma$ be compact group with an action on $V$
- Example: $\Gamma = O(3)$ and $V = S^2 \subseteq \mathbb{R}^3$
- Assume the metric is $\Gamma$-invariant:
  $d(\gamma x, \gamma y) = d(x, y)$ for all $x, y \in V$ and $\gamma \in \Gamma$

# Harmonic analysis on subset spaces

- Let $\Gamma$ be compact group with an action on $V$
- Example: $\Gamma = O(3)$ and $V = S^2 \subseteq \mathbb{R}^3$
- Assume the metric is $\Gamma$-invariant:
  $d(\gamma x, \gamma y) = d(x, y)$ for all $x, y \in V$ and $\gamma \in \Gamma$
- Then the action extends to an action on $I_t$ by
  $\gamma\emptyset = \emptyset$ and $\gamma\{x_1, \ldots, x_t\} = \{\gamma x_1, \ldots, \gamma x_t\}$

# Harmonic analysis on subset spaces

- Let $\Gamma$ be compact group with an action on $V$
- Example: $\Gamma = O(3)$ and $V = S^2 \subseteq \mathbb{R}^3$
- Assume the metric is $\Gamma$-invariant:
  $d(\gamma x, \gamma y) = d(x, y)$ for all $x, y \in V$ and $\gamma \in \Gamma$
- Then the action extends to an action on $I_t$ by
  $\gamma\emptyset = \emptyset$ and $\gamma\{x_1, \ldots, x_t\} = \{\gamma x_1, \ldots, \gamma x_t\}$
- By an "averaging argument" we may assume
  $K \in \mathcal{C}(I_t \times I_t)_{\succeq 0}$ to be $\Gamma$-invariant:
  $K(\gamma J, \gamma J') = K(J, J')$ for all $\gamma \in \Gamma$ and $J, J' \in I_t$

# Harmonic analysis on subset spaces

- Fourier inversion formula:

$$K(J, J') = \sum_{\pi \in \hat{\Gamma}} \sum_{i,j=1}^{m_\pi} \hat{K}(\pi)_{i,j} Z_\pi(J, J')_{i,j}$$

# Harmonic analysis on subset spaces

- Fourier inversion formula:

$$K(J, J') = \sum_{\pi \in \hat{\Gamma}} \sum_{i,j=1}^{m_\pi} \hat{K}(\pi)_{i,j} Z_\pi(J, J')_{i,j}$$

- The Fourier coefficients $\hat{K}(\pi)$ are psd matrices

# Harmonic analysis on subset spaces

- Fourier inversion formula:

$$K(J, J') = \sum_{\pi \in \hat{\Gamma}} \sum_{i,j=1}^{m_\pi} \hat{K}(\pi)_{i,j} Z_\pi(J, J')_{i,j}$$

- The Fourier coefficients $\hat{K}(\pi)$ are psd matrices
- The $Z_\pi(\cdot, \cdot)$ are matrix functions that depend on $\Gamma$ and $I_t$

# Harmonic analysis on subset spaces

- Fourier inversion formula:

$$K(J, J') = \sum_{\pi \in \hat{\Gamma}} \sum_{i,j=1}^{m_\pi} \hat{K}(\pi)_{i,j} Z_\pi(J, J')_{i,j}$$

- The Fourier coefficients $\hat{K}(\pi)$ are psd matrices
- The $Z_\pi(\cdot, \cdot)$ are matrix functions that depend on $\Gamma$ and $I_t$
- The action of $\Gamma$ on $I_t$ gives a linear action of $\Gamma$ on $\mathcal{C}(I_t)$ by

$$\gamma f(S) = f(\gamma^{-1} S)$$

# Harmonic analysis on subset spaces

- Fourier inversion formula:

$$K(J,J') = \sum_{\pi\in\hat{\Gamma}} \sum_{i,j=1}^{m_\pi} \hat{K}(\pi)_{i,j} Z_\pi(J,J')_{i,j}$$

- The Fourier coefficients $\hat{K}(\pi)$ are psd matrices
- The $Z_\pi(\cdot,\cdot)$ are matrix functions that depend on $\Gamma$ and $I_t$
- The action of $\Gamma$ on $I_t$ gives a linear action of $\Gamma$ on $\mathcal{C}(I_t)$ by

$$\gamma f(S) = f(\gamma^{-1}S)$$

- To construct the $Z_\pi(\cdot,\cdot)$ we need to decompose $\mathcal{C}(I_t)$ as a direct sum of irreducible $\Gamma$-invariant subspaces

# Harmonic analysis on subset spaces

- Fourier inversion formula:

$$K(J, J') = \sum_{\pi \in \hat{\Gamma}} \sum_{i,j=1}^{m_\pi} \hat{K}(\pi)_{i,j} Z_\pi(J, J')_{i,j}$$

- The Fourier coefficients $\hat{K}(\pi)$ are psd matrices
- The $Z_\pi(\cdot, \cdot)$ are matrix functions that depend on $\Gamma$ and $I_t$
- The action of $\Gamma$ on $I_t$ gives a linear action of $\Gamma$ on $\mathcal{C}(I_t)$ by

$$\gamma f(S) = f(\gamma^{-1} S)$$

- To construct the $Z_\pi(\cdot, \cdot)$ we need to decompose $\mathcal{C}(I_t)$ as a direct sum of irreducible $\Gamma$-invariant subspaces
- We give procedure to do this using symmetric tensor powers

## Harmonic analysis on subset spaces

- Fourier inversion formula:

$$K(J, J') = \sum_{\pi \in \hat{\Gamma}} \sum_{i,j=1}^{m_\pi} \hat{K}(\pi)_{i,j} Z_\pi(J, J')_{i,j}$$

- The Fourier coefficients $\hat{K}(\pi)$ are psd matrices
- The $Z_\pi(\cdot, \cdot)$ are matrix functions that depend on $\Gamma$ and $I_t$
- The action of $\Gamma$ on $I_t$ gives a linear action of $\Gamma$ on $\mathcal{C}(I_t)$ by

$$\gamma f(S) = f(\gamma^{-1} S)$$

- To construct the $Z_\pi(\cdot, \cdot)$ we need to decompose $\mathcal{C}(I_t)$ as a direct sum of irreducible $\Gamma$-invariant subspaces
- We give procedure to do this using symmetric tensor powers
- We do this explicitly for $V = S^2$, $\Gamma = O(3)$, and $t = 2$ (by using Clebsch–Gordan coefficients)

# Harmonic analysis on subset spaces

- Fourier inversion formula:

$$K(J, J') = \sum_{\pi \in \hat{\Gamma}} \sum_{i,j=1}^{m_\pi} \hat{K}(\pi)_{i,j} Z_\pi(J, J')_{i,j}$$

- The Fourier coefficients $\hat{K}(\pi)$ are psd matrices
- The $Z_\pi(\cdot, \cdot)$ are matrix functions that depend on $\Gamma$ and $I_t$
- The action of $\Gamma$ on $I_t$ gives a linear action of $\Gamma$ on $\mathcal{C}(I_t)$ by

$$\gamma f(S) = f(\gamma^{-1} S)$$

- To construct the $Z_\pi(\cdot, \cdot)$ we need to decompose $\mathcal{C}(I_t)$ as a direct sum of irreducible $\Gamma$-invariant subspaces
- We give procedure to do this using symmetric tensor powers
- We do this explicitly for $V = S^2$, $\Gamma = O(3)$, and $t = 2$ (by using Clebsch–Gordan coefficients)
- In this way we lower bound $E_2^*$ by problems with finitely many variables and infinitely many constraints

# Invariant theory (for $V = S^2$)

- These constraints are of the form

$$p(x_1, \ldots, x_i) \geq 0 \quad \text{for} \quad \{x_1, \ldots, x_i\} \in I_{=i},$$

  where $p$ is a polynomial whose coefficients depend linearly on the entries of the matrix variables

# Invariant theory (for $V = S^2$)

- These constraints are of the form

$$p(x_1, \dots, x_i) \geq 0 \quad \text{for} \quad \{x_1, \dots, x_i\} \in I_{=i},$$

  where $p$ is a polynomial whose coefficients depend linearly on the entries of the matrix variables
- These polynomials satisfy

$$p(\gamma x_1, \dots, \gamma x_i) = p(x_1, \dots, x_i) \text{ for } x_1, \dots, x_i \in S^2 \text{ and } \gamma \in O(3)$$

# Invariant theory (for $V = S^2$)

- These constraints are of the form

$$p(x_1, \ldots, x_i) \geq 0 \quad \text{for} \quad \{x_1, \ldots, x_i\} \in I_{=i},$$

where $p$ is a polynomial whose coefficients depend linearly on the entries of the matrix variables

- These polynomials satisfy

$$p(\gamma x_1, \ldots, \gamma x_i) = p(x_1, \ldots, x_i) \text{ for } x_1, \ldots, x_i \in S^2 \text{ and } \gamma \in O(3)$$

- By a theorem of invariant theory we can write $p$ as a polynomial in the inner products:

$$p(x_1, \ldots, x_i) = q(x_1 \cdot x_1, x_1 \cdot x_2, \ldots, x_i \cdot x_i)$$

# Invariant theory (for $V = S^2$)

- These constraints are of the form

$$p(x_1, \ldots, x_i) \geq 0 \quad \text{for} \quad \{x_1, \ldots, x_i\} \in I_{=i},$$

where $p$ is a polynomial whose coefficients depend linearly on the entries of the matrix variables

- These polynomials satisfy

$$p(\gamma x_1, \ldots, \gamma x_i) = p(x_1, \ldots, x_i) \text{ for } x_1, \ldots, x_i \in S^2 \text{ and } \gamma \in O(3)$$

- By a theorem of invariant theory we can write $p$ as a polynomial in the inner products:

$$p(x_1, \ldots, x_i) = q(x_1 \cdot x_1, x_1 \cdot x_2, \ldots, x_i \cdot x_i)$$

- Now we have constraints of the form

$$q(u_1, \ldots, u_l) \geq 0 \quad \text{for} \quad (u_1, \ldots, u_l) \in \text{some semialgebraic set}$$

# Invariant theory

$$p(x_1, \ldots, x_i) = q(x_1 \cdot x_1, x_1 \cdot x_2, \ldots, x_i \cdot x_i), \qquad \deg(p) = 2d$$

- The theorem that gives the existence of $q$ is nonconstructive

# Invariant theory

$$p(x_1, \ldots, x_i) = q(x_1 \cdot x_1, x_1 \cdot x_2, \ldots, x_i \cdot x_i), \qquad \deg(p) = 2d$$

- The theorem that gives the existence of $q$ is nonconstructive
- Find $q$ by solving linear system $Ax = b$
  Rows indexed by monomials in $3i$ vars of degree $\leq 2d$
  Columns indexed by monomials in $\binom{i+1}{2}$ vars of degree $\leq d$

# Invariant theory

$$p(x_1, \ldots, x_i) = q(x_1 \cdot x_1, x_1 \cdot x_2, \ldots, x_i \cdot x_i), \qquad \deg(p) = 2d$$

- The theorem that gives the existence of $q$ is nonconstructive
- Find $q$ by solving linear system $Ax = b$
  Rows indexed by monomials in $3i$ vars of degree $\leq 2d$
  Columns indexed by monomials in $\binom{i+1}{2}$ vars of degree $\leq d$
- For $i = 4$, $d = 6$ we get over a million rows

# Invariant theory

$$p(x_1, \ldots, x_i) = q(x_1 \cdot x_1, x_1 \cdot x_2, \ldots, x_i \cdot x_i), \qquad \deg(p) = 2d$$

- The theorem that gives the existence of $q$ is nonconstructive
- Find $q$ by solving linear system $Ax = b$
  Rows indexed by monomials in $3i$ vars of degree $\leq 2d$
  Columns indexed by monomials in $\binom{i+1}{2}$ vars of degree $\leq d$
- For $i = 4$, $d = 6$ we get over a million rows
- Use custom pivoting, sparse, high precision, Cholesky factorization algorithm

# Invariant theory

$$p(x_1, \ldots, x_i) = q(x_1 \cdot x_1, x_1 \cdot x_2, \ldots, x_i \cdot x_i), \qquad \deg(p) = 2d$$

- The theorem that gives the existence of $q$ is nonconstructive
- Find $q$ by solving linear system $Ax = b$
  Rows indexed by monomials in $3i$ vars of degree $\leq 2d$
  Columns indexed by monomials in $\binom{i+1}{2}$ vars of degree $\leq d$
- For $i = 4$, $d = 6$ we get over a million rows
- Use custom pivoting, sparse, high precision, Cholesky factorization algorithm
- Computing the $q$ polynomials takes several days, but only needs to be done once for given $d$

# Sums of squares characterizations

- Putinar: Every positive polynomial on a compact set $S = \{x \in \mathbb{R}^n : g_1(x) \geq 0, \ldots, g_m(x) \geq 0\}$, where $\{g_1, \ldots, g_m\}$ has the Archimedean property, is of the form

$$f(x) = \sum_{i=0}^{m} g_i(x) s_i(x), \quad \text{where} \quad g_0 = 1 \text{ and } s_0, \ldots, s_m \text{ are SOS}$$

# Sums of squares characterizations

- Putinar: Every positive polynomial on a compact set $S = \{x \in \mathbb{R}^n : g_1(x) \geq 0, \ldots, g_m(x) \geq 0\}$, where $\{g_1, \ldots, g_m\}$ has the Archimedean property, is of the form

$$f(x) = \sum_{i=0}^{m} g_i(x) s_i(x), \quad \text{where} \quad g_0 = 1 \text{ and } s_0, \ldots, s_m \text{ are SOS}$$

- The SOS polynomials $s_i$ can be modeled using psd matrices

# Sums of squares characterizations

- Putinar: Every positive polynomial on a compact set $S = \{x \in \mathbb{R}^n : g_1(x) \geq 0, \ldots, g_m(x) \geq 0\}$, where $\{g_1, \ldots, g_m\}$ has the Archimedean property, is of the form

$$f(x) = \sum_{i=0}^{m} g_i(x) s_i(x), \quad \text{where} \quad g_0 = 1 \text{ and } s_0, \ldots, s_m \text{ are SOS}$$

- The SOS polynomials $s_i$ can be modeled using psd matrices
- We use this to go from infinitely many linear constraints to finitely many semidefinite constraints

# Sums of squares characterizations

- Putinar: Every positive polynomial on a compact set $S = \{x \in \mathbb{R}^n : g_1(x) \geq 0, \ldots, g_m(x) \geq 0\}$, where $\{g_1, \ldots, g_m\}$ has the Archimedean property, is of the form

$$f(x) = \sum_{i=0}^{m} g_i(x) s_i(x), \quad \text{where} \quad g_0 = 1 \text{ and } s_0, \ldots, s_m \text{ are SOS}$$

- The SOS polynomials $s_i$ can be modeled using psd matrices
- We use this to go from infinitely many linear constraints to finitely many semidefinite constraints
- In energy minimization the particles are interchangeable

# Sums of squares characterizations

- Putinar: Every positive polynomial on a compact set $S = \{x \in \mathbb{R}^n : g_1(x) \geq 0, \ldots, g_m(x) \geq 0\}$, where $\{g_1, \ldots, g_m\}$ has the Archimedean property, is of the form

$$f(x) = \sum_{i=0}^{m} g_i(x) s_i(x), \quad \text{where} \quad g_0 = 1 \text{ and } s_0, \ldots, s_m \text{ are SOS}$$

- The SOS polynomials $s_i$ can be modeled using psd matrices
- We use this to go from infinitely many linear constraints to finitely many semidefinite constraints
- In energy minimization the particles are interchangeable
- This means

$$p(x_{\sigma(1)}, \ldots, x_{\sigma(i)}) = p(x_1, \ldots, x_i) \quad \text{for all} \quad \sigma \in S_i$$

# Sums of squares characterizations

- Putinar: Every positive polynomial on a compact set $S = \{x \in \mathbb{R}^n : g_1(x) \geq 0, \ldots, g_m(x) \geq 0\}$, where $\{g_1, \ldots, g_m\}$ has the Archimedean property, is of the form

$$f(x) = \sum_{i=0}^{m} g_i(x) s_i(x), \quad \text{where} \quad g_0 = 1 \text{ and } s_0, \ldots, s_m \text{ are SOS}$$

- The SOS polynomials $s_i$ can be modeled using psd matrices
- We use this to go from infinitely many linear constraints to finitely many semidefinite constraints
- In energy minimization the particles are interchangeable
- This means

$$p(x_{\sigma(1)}, \ldots, x_{\sigma(i)}) = p(x_1, \ldots, x_i) \quad \text{for all} \quad \sigma \in S_i$$

- Additional symmetries in the $q(u_1, \ldots, u_l)$ polynomials

# Sums of squares characterizations

- Symmetrization of Putinar's theorem to exploit the symmetry in the particles

# Sums of squares characterizations

- Symmetrization of Putinar's theorem to exploit the symmetry in the particles
- Assume the set $\{g_0, \ldots, g_m\}$ is $\Gamma$-invariant

# Sums of squares characterizations

- Symmetrization of Putinar's theorem to exploit the symmetry in the particles
- Assume the set $\{g_0, \ldots, g_m\}$ is $\Gamma$-invariant
- Denote by $\Gamma_{g_i}$ the stabilizer subgroup of $\Gamma$ with respect to $g_i$

# Sums of squares characterizations

- Symmetrization of Putinar's theorem to exploit the symmetry in the particles
- Assume the set $\{g_0, \ldots, g_m\}$ is $\Gamma$-invariant
- Denote by $\Gamma_{g_i}$ the stabilizer subgroup of $\Gamma$ with respect to $g_i$

A $\Gamma$-invariant polynomial that has a Putinar representation can be written as $p = \sum_{i=0}^m g_i s_i$, where $s_i$ is a $\Gamma_{g_i}$-invariant sum of squares polynomial

# Sums of squares characterizations

- Symmetrization of Putinar's theorem to exploit the symmetry in the particles
- Assume the set $\{g_0, \ldots, g_m\}$ is $\Gamma$-invariant
- Denote by $\Gamma_{g_i}$ the stabilizer subgroup of $\Gamma$ with respect to $g_i$

A $\Gamma$-invariant polynomial that has a Putinar representation can be written as $p = \sum_{i=0}^m g_i s_i$, where $s_i$ is a $\Gamma_{g_i}$-invariant sum of squares polynomial

- We can represent the $\Gamma_{g_i}$-invariant sum of squares polynomials $s_i$ using block diagonalized positive semidefinite matrices [Gatermann–Parillo 2004]

# Sums of squares characterizations

- Symmetrization of Putinar's theorem to exploit the symmetry in the particles
- Assume the set $\{g_0, \ldots, g_m\}$ is $\Gamma$-invariant
- Denote by $\Gamma_{g_i}$ the stabilizer subgroup of $\Gamma$ with respect to $g_i$

A $\Gamma$-invariant polynomial that has a Putinar representation can be written as $p = \sum_{i=0}^m g_i s_i$, where $s_i$ is a $\Gamma_{g_i}$-invariant sum of squares polynomial

- We can represent the $\Gamma_{g_i}$-invariant sum of squares polynomials $s_i$ using block diagonalized positive semidefinite matrices [Gatermann–Parillo 2004]
- For energy minimization on the sphere this yields large reductions in solver time (Ex. 150 hours $\rightarrow$ 7 hours)

# Computations

- Mow we have an SDP given as high precision numbers whose optimal value lower bounds the ground state energy

# Computations

- Mow we have an SDP given as high precision numbers whose optimal value lower bounds the ground state energy
- Want to solve with high precision SDP solver

# Computations

- Mow we have an SDP given as high precision numbers whose optimal value lower bounds the ground state energy
- Want to solve with high precision SDP solver
- Problem 1: Free variables in the SDP $\rightarrow$ Dual SDP not strictly feasible $\rightarrow$ Cannot solve with high precision solver

# Computations

- Mow we have an SDP given as high precision numbers whose optimal value lower bounds the ground state energy
- Want to solve with high precision SDP solver
- Problem 1: Free variables in the SDP $\rightarrow$ Dual SDP not strictly feasible $\rightarrow$ Cannot solve with high precision solver
- Bound free variables with big $M$ constraints

# Computations

- Mow we have an SDP given as high precision numbers whose optimal value lower bounds the ground state energy
- Want to solve with high precision SDP solver
- Problem 1: Free variables in the SDP $\rightarrow$ Dual SDP not strictly feasible $\rightarrow$ Cannot solve with high precision solver
- Bound free variables with big $M$ constraints
- Problem 2: The additional symmetry exploitation leads to hard to predict linear dependencies in the constraints

# Computations

- Mow we have an SDP given as high precision numbers whose optimal value lower bounds the ground state energy
- Want to solve with high precision SDP solver
- Problem 1: Free variables in the SDP $\to$ Dual SDP not strictly feasible $\to$ Cannot solve with high precision solver
- Bound free variables with big $M$ constraints
- Problem 2: The additional symmetry exploitation leads to hard to predict linear dependencies in the constraints
- Use QR factorization of the constraint matrix to remove these

# Computations

- In the Thomson problem we take

$$V = S^2, \quad d(x,y) = \|x - y\|_2, \quad \text{and} \quad h(w) = \frac{1}{w}$$

# Computations

- In the Thomson problem we take

$$V = S^2, \quad d(x,y) = \|x-y\|_2, \quad \text{and} \quad h(w) = \frac{1}{w}$$

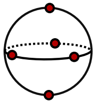
- $E_1^*$ is sharp for $2$, $3$, $4$, $6$, and $12$ particles (Yudin's LP bound)

# Computations

- In the Thomson problem we take

$$V = S^2, \quad d(x,y) = \|x-y\|_2, \quad \text{and} \quad h(w) = \frac{1}{w}$$

- $E_1^*$ is sharp for $2$, $3$, $4$, $6$, and $12$ particles (Yudin's LP bound)
- The triangular bipiramid is optimal for $N = 5$ (Schwartz 2010)

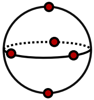


# Computations

- In the Thomson problem we take

$$V = S^2, \quad d(x,y) = \|x-y\|_2, \quad \text{and} \quad h(w) = \frac{1}{w}$$

- $E_1^*$ is sharp for $2$, $3$, $4$, $6$, and $12$ particles (Yudin's LP bound)
- The triangular bipiramid is optimal for $N = 5$ (Schwartz 2010)



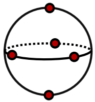
- High precision SDP solver gives the first $28$ decimal digits of a lower bound on $E_2$

# Computations

- In the Thomson problem we take

$$V = S^2, \quad d(x,y) = \|x-y\|_2, \quad \text{and} \quad h(w) = \frac{1}{w}$$

- $E_1^*$ is sharp for $2$, $3$, $4$, $6$, and $12$ particles (Yudin's LP bound)
- The triangular bipiramid is optimal for $N = 5$ (Schwartz 2010)



- High precision SDP solver gives the first $28$ decimal digits of a lower bound on $E_2$
- These all agree with the energy of the triangular bipiramid

# Computations

- We should be able to use this to construct an optimality certificate for the $N = 5$ case of the Thomson problem, but need to replace linear algebra by Gröbner bases

# Computations

- We should be able to use this to construct an optimality certificate for the $N = 5$ case of the Thomson problem, but need to replace linear algebra by Gröbner bases
- The system of $5$ particles on $S^2$ admits a phase transition

# Computations

- We should be able to use this to construct an optimality certificate for the $N = 5$ case of the Thomson problem, but need to replace linear algebra by Gröbner bases
- The system of $5$ particles on $S^2$ admits a phase transition
- Using SDP solver we see $E_2$ is also (numerically) sharp for many other pair potentials

# Computations

- We should be able to use this to construct an optimality certificate for the $N = 5$ case of the Thomson problem, but need to replace linear algebra by Gröbner bases
- The system of $5$ particles on $S^2$ admits a phase transition
- Using SDP solver we see $E_2$ is also (numerically) sharp for many other pair potentials
- Conjecture: $E_2$ is universally sharp for $5$ particles on $S^2$

# Computations

- We should be able to use this to construct an optimality certificate for the $N = 5$ case of the Thomson problem, but need to replace linear algebra by Gröbner bases
- The system of $5$ particles on $S^2$ admits a phase transition
- Using SDP solver we see $E_2$ is also (numerically) sharp for many other pair potentials
- Conjecture: $E_2$ is universally sharp for $5$ particles on $S^2$
- This is the first time a four $4$-bound has been computed for a continuous problem

# Computations

- We should be able to use this to construct an optimality certificate for the $N = 5$ case of the Thomson problem, but need to replace linear algebra by Gröbner bases
- The system of $5$ particles on $S^2$ admits a phase transition
- Using SDP solver we see $E_2$ is also (numerically) sharp for many other pair potentials
- Conjecture: $E_2$ is universally sharp for $5$ particles on $S^2$
- This is the first time a four $4$-bound has been computed for a continuous problem
- Future work: apply these techniques to packing problems

Thank you!

D. de Laat, *Moment methods in energy minimization: New bounds for Riesz minimal energy problems*, arXiv:1610.04905.