Energy minimization via conic programming hierarchies

David de Laat (TU Delft)

IFORS
July 14, 2014, Barcelona

Energy minimization

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

Energy minimization

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

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

Energy minimization

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

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

- Optimization problem:

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

Approach

- Configurations provide upper bounds on the optimal energy E

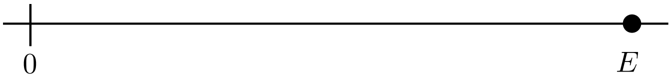
Approach

- 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
- For this we use infinite dimensional moment hierarchies and semidefinite programming

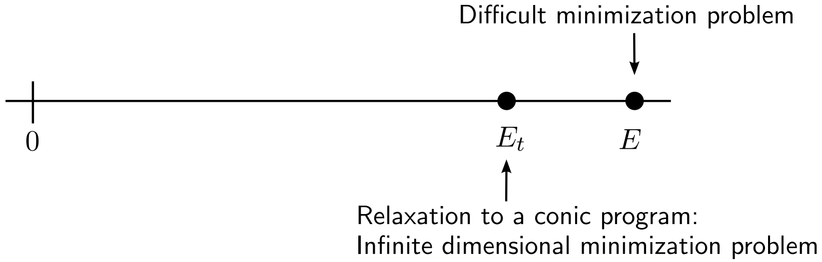
Approach



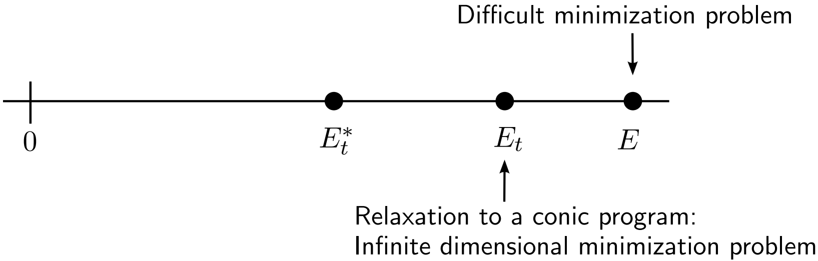
Approach

Approach

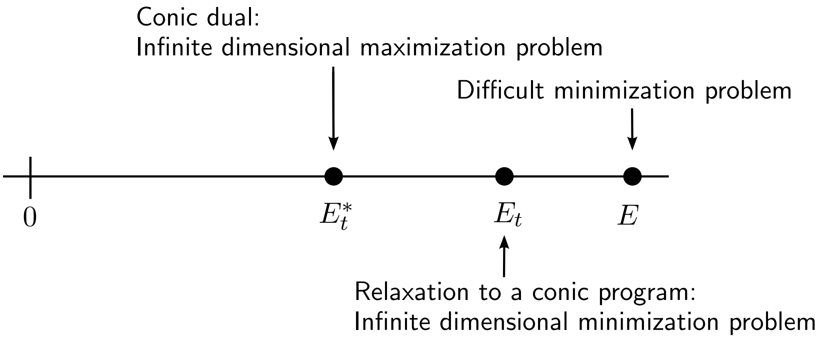
Approach


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

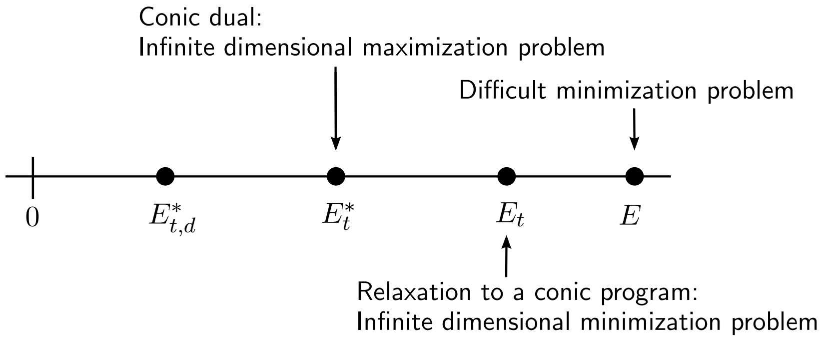
Approach


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

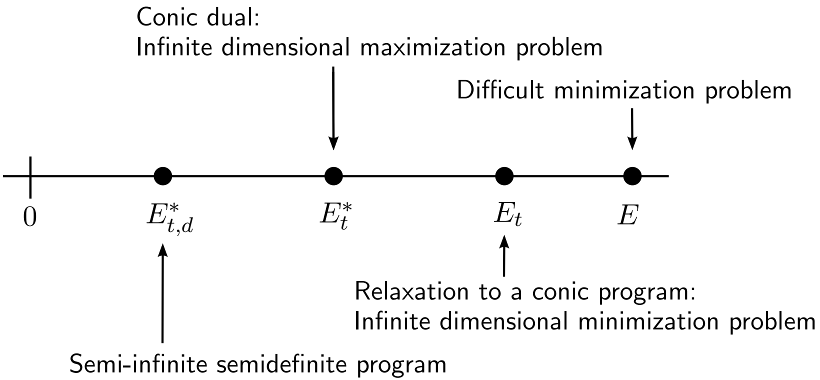
Approach


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

Finite container

- If $V = \{1, \ldots, n\}$ is a finite set, then E is a polynomial optimization problem:

$$E = \min \left\{ \sum_{\{i,j\} \in \binom{V}{2}} w(\{i,j\}) x_i x_j : x \in \{0,1\}^n, \sum_{i \in V} x_i = N \right\}$$

Finite container

- If $V = \{1, \ldots, n\}$ is a finite set, then E is a polynomial optimization problem:

$$E = \min \Big\{ \sum_{\{i,j\} \in \binom{V}{2}} w(\{i,j\}) x_i x_j : x \in \{0,1\}^n, \sum_{i \in V} x_i = N \Big\}$$

- The Lasserre hierarchy gives a chain $E_1 \leq E_2 \leq \cdots \leq E_n$ of lower bounds to the optimal energy E:

Finite container

- If $V = \{1, \ldots, n\}$ is a finite set, then E is a polynomial optimization problem:

$$E = \min \left\{ \sum_{\{i,j\} \in \binom{V}{2}} w(\{i,j\}) x_i x_j : x \in \{0,1\}^n, \sum_{i \in V} x_i = N \right\}$$

- The Lasserre hierarchy gives a chain $E_1 \leq E_2 \leq \cdots \leq E_n$ of lower bounds to the optimal energy E:

$$E_t = \min \Bigg\{ \sum_{S \in \binom{V}{2}} w(S) y(S) : y \in \mathbb{R}^{\binom{V}{\leq 2t}},\ y(\emptyset) = 1,\ \big(y(A \cup B)\big)_{A,B \in \binom{V}{\leq t}} \succeq 0,$$
$$\sum_{x \in V} y(T \cup \{x\}) = N y(T) \text{ for } T \in \binom{V}{\leq 2t-1} \Bigg\}$$

Infinite container

- Assume V is a compact Hausdorff space and w continuous

Infinite container

- Assume V is a compact Hausdorff space and w continuous
- $\binom{V}{\leq t} \setminus \{\emptyset\}$ gets its topology as a quotient of V^t

Infinite container

- Assume V is a compact Hausdorff space and w continuous
- $\binom{V}{\leq t} \setminus \{\emptyset\}$ gets its topology as a quotient of V^t
- Generalization (here $s = \min\{2t, N\}$):

$$E_t = \min\Big\{\lambda(w) : \lambda \in \mathcal{M}(\tbinom{V}{\leq s})_{\geq 0},\ A_t^* \lambda \in \mathcal{M}(\tbinom{V}{\leq t} \times \tbinom{V}{\leq t})_{\succeq 0},$$
$$\lambda(\tbinom{V}{i}) = \tbinom{N}{i} \text{ for } i = 0, \ldots, s\Big\}$$

Infinite container

- Assume V is a compact Hausdorff space and w continuous
- $\binom{V}{\leq t} \setminus \{\emptyset\}$ gets its topology as a quotient of V^t
- Generalization (here $s = \min\{2t, N\}$):

$$E_t = \min \Big\{ \lambda(w) : \lambda \in \mathcal{M}(\tbinom{V}{\leq s})_{\geq 0},\ A_t^* \lambda \in \mathcal{M}(\tbinom{V}{\leq t} \times \tbinom{V}{\leq t})_{\succeq 0}, \\ \lambda(\tbinom{V}{i}) = \tbinom{N}{i} \text{ for } i = 0, \ldots, s \Big\}$$

- λ generalizes the moment vector y

Infinite container

- Assume V is a compact Hausdorff space and w continuous
- $\binom{V}{\leq t} \setminus \{\emptyset\}$ gets its topology as a quotient of V^t
- Generalization (here $s = \min\{2t, N\}$):

$$E_t = \min\left\{\lambda(w) : \lambda \in \mathcal{M}(\tbinom{V}{\leq s})_{\geq 0},\ A_t^* \lambda \in \mathcal{M}(\tbinom{V}{\leq t} \times \tbinom{V}{\leq t})_{\succeq 0},\right.$$
$$\left.\lambda(\tbinom{V}{i}) = \tbinom{N}{i} \text{ for } i = 0, \ldots, s\right\}$$

- λ generalizes the moment vector y
- $\mathcal{M}(\binom{V}{\leq t} \times \binom{V}{\leq t})_{\succeq 0}$ is dual to the cone $\mathcal{C}(\binom{V}{\leq t} \times \binom{V}{\leq t})_{\succeq 0}$ of positive definite kernels

Infinite container

- Assume V is a compact Hausdorff space and w continuous
- $\binom{V}{\leq t} \setminus \{\emptyset\}$ gets its topology as a quotient of V^t
- Generalization (here $s = \min\{2t, N\}$):

$$E_t = \min \Big\{ \lambda(w) : \lambda \in \mathcal{M}(\tbinom{V}{\leq s})_{\geq 0},\ A_t^* \lambda \in \mathcal{M}(\tbinom{V}{\leq t} \times \tbinom{V}{\leq t})_{\succeq 0}, \\ \lambda(\tbinom{V}{i}) = \tbinom{N}{i} \text{ for } i = 0, \ldots, s \Big\}$$

- λ generalizes the moment vector y
- $\mathcal{M}(\binom{V}{\leq t} \times \binom{V}{\leq t})_{\succeq 0}$ is dual to the cone $\mathcal{C}(\binom{V}{\leq t} \times \binom{V}{\leq t})_{\succeq 0}$ of positive definite kernels
- Relaxation: If S is an N subset of V, then

$$\chi_S = \sum_{R \in \binom{S}{\leq 2t}} \delta_R$$

is feasible for E_t

Infinite container

- Assume V is a compact Hausdorff space and w continuous
- $\binom{V}{\leq t} \setminus \{\emptyset\}$ gets its topology as a quotient of V^t
- Generalization (here $s = \min\{2t, N\}$):

$$E_t = \min\Big\{\lambda(w) : \lambda \in \mathcal{M}(\tbinom{V}{\leq s})_{\geq 0},\ A_t^*\lambda \in \mathcal{M}(\tbinom{V}{\leq t} \times \tbinom{V}{\leq t})_{\succeq 0},\ \lambda(\tbinom{V}{i}) = \tbinom{N}{i} \text{ for } i = 0, \ldots, s\Big\}$$

- λ generalizes the moment vector y
- $\mathcal{M}(\binom{V}{\leq t} \times \binom{V}{\leq t})_{\succeq 0}$ is dual to the cone $\mathcal{C}(\binom{V}{\leq t} \times \binom{V}{\leq t})_{\succeq 0}$ of positive definite kernels
- Relaxation: If S is an N subset of V, then

$$\chi_S = \sum_{R \in \binom{S}{\leq 2t}} \delta_R$$

  is feasible for E_t
- We have $E_N = E$

Infinite container

- Assume V is a compact Hausdorff space and w continuous
- $\binom{V}{\le t} \setminus \{\emptyset\}$ gets its topology as a quotient of V^t
- Generalization (here $s = \min\{2t, N\}$):

$$E_t = \min \Big\{ \lambda(w) : \lambda \in \mathcal{M}(\tbinom{V}{\le s})_{\ge 0},\ A_t^* \lambda \in \mathcal{M}(\tbinom{V}{\le t} \times \tbinom{V}{\le t})_{\succeq 0}, \\ \lambda(\tbinom{V}{i}) = \tbinom{N}{i} \text{ for } i = 0, \dots, s \Big\}$$

- λ generalizes the moment vector y
- $\mathcal{M}(\binom{V}{\le t} \times \binom{V}{\le t})_{\succeq 0}$ is dual to the cone $\mathcal{C}(\binom{V}{\le t} \times \binom{V}{\le t})_{\succeq 0}$ of positive definite kernels
- Relaxation: If S is an N subset of V, then

$$\chi_S = \sum_{R \in \binom{S}{\le 2t}} \delta_R$$

is feasible for E_t
- We have $E_N = E$
- Uses techniques from [de Laat-Vallentin 2013]: hierarchy for packing problems in discrete geometry

Dual hierarchy

- For lower bounds we need feasible solutions of the dual

Dual hierarchy

- For lower bounds we need feasible solutions of the dual
- In the dual hierarchy optimization is over scalars a_i and positive definite kernels $K \in \mathcal{C}(\binom{V}{\leq t} \times \binom{V}{\leq t})_{\succeq 0}$:

Dual hierarchy

- For lower bounds we need feasible solutions of the dual
- In the dual hierarchy optimization is over scalars a_i and positive definite kernels $K \in \mathcal{C}(\binom{V}{\leq t} \times \binom{V}{\leq t})_{\succeq 0}$:

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

Dual hierarchy

- For lower bounds we need feasible solutions of the dual
- In the dual hierarchy optimization is over scalars a_i and positive definite kernels $K \in \mathcal{C}(\binom{V}{\le t} \times \binom{V}{\le t})_{\succeq 0}$:

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

- Techniquality: we only put a linear constraint for $S \in \binom{V}{i}$ if the points in S are not too close

Dual hierarchy

- For lower bounds we need feasible solutions of the dual
- In the dual hierarchy optimization is over scalars a_i and positive definite kernels $K \in \mathcal{C}(\binom{V}{\leq t} \times \binom{V}{\leq t})_{\succeq 0}$:

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

- Techniquality: we only put a linear constraint for $S \in \binom{V}{i}$ if the points in S are not too close
- Strong duality holds: $E_t = E_t^*$

Dual hierarchy

- For lower bounds we need feasible solutions of the dual
- In the dual hierarchy optimization is over scalars a_i and positive definite kernels $K \in \mathcal{C}(\binom{V}{\le t} \times \binom{V}{\le t})_{\succeq 0}$:

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

- Techniquality: we only put a linear constraint for $S \in \binom{V}{i}$ if the points in S are not too close
- Strong duality holds: $E_t = E_t^*$
- If Γ acts on V and w is Γ-invariant, then we can restrict to Γ-invariant kernels: $K(\gamma J, \gamma J') = K(J, J')$ for all $J, J' \in \binom{V}{\le t}$ (Here $\gamma\{x_1, \ldots, x_t\} = \{\gamma x_1, \ldots, \gamma x_t\}$)

Inner approximiations to the cone $\mathcal{C}(\binom{V}{\leq t} \times \binom{V}{\leq t})^{\Gamma}_{\succeq 0}$

- Nested chain of inner approximations:

$$C_1 \subseteq C_2 \subseteq \cdots \subseteq \mathcal{C}(\binom{V}{\leq t} \times \binom{V}{\leq t})^{\Gamma}_{\succeq 0}$$

Inner approximiations to the cone $\mathcal{C}(\binom{V}{\leq t} \times \binom{V}{\leq t})_{\succeq 0}^{\Gamma}$

- Nested chain of inner approximations:

$$C_1 \subseteq C_2 \subseteq \cdots \subseteq \mathcal{C}(\binom{V}{\leq t} \times \binom{V}{\leq t})_{\succeq 0}^{\Gamma}$$

- Each cone C_i can be parametrized by a finite direct sum of positive semidefinite matrix cones

Inner approximiations to the cone $\mathcal{C}(\binom{V}{\leq t} \times \binom{V}{\leq t})^{\Gamma}_{\succeq 0}$

- Nested chain of inner approximations:

$$C_1 \subseteq C_2 \subseteq \cdots \subseteq \mathcal{C}(\binom{V}{\leq t} \times \binom{V}{\leq t})^{\Gamma}_{\succeq 0}$$

- Each cone C_i can be parametrized by a finite direct sum of positive semidefinite matrix cones
- Bochner: A kernel $K \in \mathcal{C}(\binom{V}{\leq t} \times \binom{V}{\leq t})^{\Gamma}_{\succeq 0}$ is of the form

$$K(J, J') = \sum_{k=0}^{\infty} \operatorname{trace}(F_k Z_k(J, J'))$$

  - F_k: (infinite) positive semidefinite matrices (the Fourier coefficients)
  - Z_k: zonal matrices corresponding to the action of Γ on $\binom{V}{\leq t}$ (generalizes $e^{2\pi i k x}$ in the Fourier transform on the circle)

Inner approximiations to the cone $\mathcal{C}(\binom{V}{\leq t} \times \binom{V}{\leq t})^{\Gamma}_{\succeq 0}$

- Nested chain of inner approximations:

$$C_1 \subseteq C_2 \subseteq \cdots \subseteq \mathcal{C}(\binom{V}{\leq t} \times \binom{V}{\leq t})^{\Gamma}_{\succeq 0}$$

- Each cone C_i can be parametrized by a finite direct sum of positive semidefinite matrix cones
- Bochner: A kernel $K \in \mathcal{C}(\binom{V}{\leq t} \times \binom{V}{\leq t})^{\Gamma}_{\succeq 0}$ is of the form

$$K(J, J') = \sum_{k=0}^{\infty} \mathrm{trace}(F_k Z_k(J, J'))$$

  - F_k: (infinite) positive semidefinite matrices (the Fourier coefficients)
  - Z_k: zonal matrices corresponding to the action of Γ on $\binom{V}{\leq t}$ (generalizes $e^{2\pi i k x}$ in the Fourier transform on the circle)
- Define C_d by truncating the above series

The semi-infinite semidefinite programs $E^*_{t,d}$

- Define $E^*_{t,d}$ by replacing the cone $\mathcal{C}(\binom{V}{\leq t} \times \binom{V}{\leq t})^{\Gamma}_{\succeq 0}$ in E^*_t by the cone C_d

The semi-infinite semidefinite programs $E_{t,d}^*$

- Define $E_{t,d}^*$ by replacing the cone $\mathcal{C}(\binom{V}{\leq t} \times \binom{V}{\leq t})_{\succeq 0}^{\Gamma}$ in E_t^* by the cone C_d
- This is an optimization problem with finitely many variables and infinitely many constraints

The semi-infinite semidefinite programs $E_{t,d}^*$

- Define $E_{t,d}^*$ by replacing the cone $\mathcal{C}(\binom{V}{\leq t} \times \binom{V}{\leq t})_{\succeq 0}^{\Gamma}$ in E_t^* by the cone C_d
- This is an optimization problem with finitely many variables and infinitely many constraints
- $E_{t,d}^* \to E_t^*$ as $d \to \infty$ follows from $\cup_{d=0}^{\infty} C_d$ being uniformly dense in $\mathcal{C}(\binom{V}{\leq t} \times \binom{V}{\leq t})_{\succeq 0}^{\Gamma}$

Example: $V = S^1$ with $O(2)$-invariant pair potential w

- The linear constraints in $E_{t,d}^*$ can be written as the nonnegativity of a trigonometric polynomial in $s-1$ variables

Example: $V = S^1$ with $O(2)$-invariant pair potential w

- The linear constraints in $E_{t,d}^*$ can be written as the nonnegativity of a trigonometric polynomial in $s-1$ variables
- Use trigonometric SOS characterizations [Dumitrescu 2006]

Example: $V = S^1$ with $O(2)$-invariant pair potential w

- The linear constraints in $E_{t,d}^*$ can be written as the nonnegativity of a trigonometric polynomial in $s-1$ variables
- Use trigonometric SOS characterizations [Dumitrescu 2006]
- For the Coulomb potential (or other completely monotonic potentials) the regular N-gon is the optimal configuration on the circle [Cohn-Kumar 2006]

Example: $V = S^1$ with $O(2)$-invariant pair potential w

- The linear constraints in $E_{t,d}^*$ can be written as the nonnegativity of a trigonometric polynomial in $s-1$ variables
- Use trigonometric SOS characterizations [Dumitrescu 2006]
- For the Coulomb potential (or other completely monotonic potentials) the regular N-gon is the optimal configuration on the circle [Cohn-Kumar 2006]
- Uses relaxation based on the 2-point correlation function [Yudin 1992] (This is similar to E_1)

Example: $V = S^1$ with $O(2)$-invariant pair potential w

- The linear constraints in $E_{t,d}^*$ can be written as the nonnegativity of a trigonometric polynomial in $s-1$ variables
- Use trigonometric SOS characterizations [Dumitrescu 2006]
- For the Coulomb potential (or other completely monotonic potentials) the regular N-gon is the optimal configuration on the circle [Cohn-Kumar 2006]
- Uses relaxation based on the 2-point correlation function [Yudin 1992] (This is similar to E_1)
- The bound E_2^* requires SOS characterizations in 3 variables

Example: $V = S^1$ with $O(2)$-invariant pair potential w

- The linear constraints in $E^*_{t,d}$ can be written as the nonnegativity of a trigonometric polynomial in $s-1$ variables
- Use trigonometric SOS characterizations [Dumitrescu 2006]
- For the Coulomb potential (or other completely monotonic potentials) the regular N-gon is the optimal configuration on the circle [Cohn-Kumar 2006]
- Uses relaxation based on the 2-point correlation function [Yudin 1992] (This is similar to E_1)
- The bound E^*_2 requires SOS characterizations in 3 variables
- Lennard-Jones potential: Based on a sampling implementation it appears that for e.g. $N = 3$ we have

$$E_1 < E_2 = E$$

Thank you!