A cycle-based formulation for the Distance Geometry Problem

Leo Liberti[1], **Gabriele Iommazzo**[1,2], Carlile Lavor[3], Nelson Maculan[4]

[1]DASCIM Team, LIX CNRS, Ecole Polytechnique, IPP, Palaiseau, France

[2]Operations Research Group, Dip. di Informatica, Università di Pisa, Italy

[3]IMECC, University of Campinas, Brazil

[4]COPPE, Federal University of Rio de Janeiro (UFRJ), Brazil

CTW 2020, September 14-16, 2020

Outline

1. Introduction
2. Our cycle-based formulation
3. Computational experiments

Distance Geometry Problem (DGP)

Definition (DGP)

Given: simple undirected weighted graph $G = (V, E, d)$ and $K \in \mathbb{N}_{>0}$, find *realization* of vertices $x : V \longrightarrow \mathbb{R}^K$ s.t.

$$\forall \{i, j\} \in E \quad \|x_i - x_j\| = d_{ij} \ , \tag{1}$$

Complexity [Beeker et al., 2013, Saxe, 1979]

- DGP_1: in NP + NP-complete;
- $\text{DGP}_{>1}$ (general graphs): may not be in NP;
- DGP on simple cycle graphs: weakly NP-hard;
- DGP with $d(\cdot) \in \{1, 2\}$ (general graphs): strongly NP-hard;

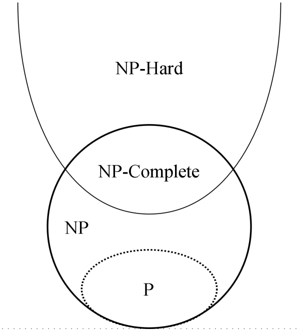


Figure: Complexity classes

Euclidean DGP (EDGP)

EDGP (parametrised over ℓ_2-norm): find $x : V \longrightarrow \mathbb{R}^K$ s.t.

$$\forall \{i,j\} \in E \quad \|x_i - x_j\|_2^2 = d_{ij}{}^2 \tag{2}$$

Applications

- clock synchronization ($K = 1$) [Singer, 2011];
- sensor network localization ($K = 2$) [Aspnes et al., 2006];
- molecular DGP ($K = 3$) [Lavor et al., 2009];
- ...

Solving EDGP: edge formulation

$$\min_x \sum_{\{i,j\}\in E} (\|x_i - x_j\|_2^2 - d_{ij}^2)^2 \quad (3)$$

- unconstrained, nonconvex, multimodal polynomial minimization problem;
- to overcome complexity: a) reformulations (ℓ_1-norm, constrained form, linearize, ...); b) local solvers; c) heuristic approaches
- optimal solution x^* is realization of given G $\iff$ value of x^* is zero (glob. min);

Solving EDGP: edge formulation

$$\min_x \sum_{\{i,j\} \in E} (\|x_i - x_j\|_2^2 - d_{ij}^2)^2 \tag{3}$$

- unconstrained, nonconvex, multimodal polynomial minimization problem;
- to overcome complexity: a) reformulations (ℓ_1-norm, constrained form, linearize, . . .); b) local solvers; c) heuristic approaches
- optimal solution x^* is realization of given G $\iff$ value of x^* is zero (glob. min);

Solving EDGP: edge formulation

$$\min_x \sum_{\{i,j\}\in E} (\|x_i - x_j\|_2^2 - d_{ij}^2)^2 \tag{3}$$

- unconstrained, nonconvex, multimodal polynomial minimization problem;
- to overcome complexity: a) reformulations (ℓ_1-norm, constrained form, linearize, ...); b) local solvers; c) heuristic approaches
- optimal solution x^* is realization of given G $\iff$ value of x^* is zero (glob. min);

Cycle-based formulation: stemming idea

Lemma

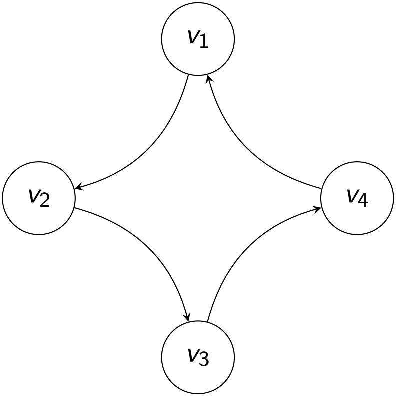
Given $K \in \mathbb{N}^+$, a simple undirected weighted graph $G = (V, E, d)$ and $x : V \longrightarrow \mathbb{R}^K$, then for each cycle C in G, each orientation of edges in C given by closed trail $W(C)$, and each $k \leq K$:

$$\sum_{(i,j) \in W(C)} (x_{ik} - x_{jk}) = 0 \tag{4}$$

Used in [Saxe, 1979] to prove weak NP hardness of DGP on simple cycle graphs

Cycle-based formulation: stemming idea (proof)

1) renumber vertices in cycle, according to trail order



Cycle-based formulation: stemming idea (proof)

2) for each $k \leq K$, start adding up:

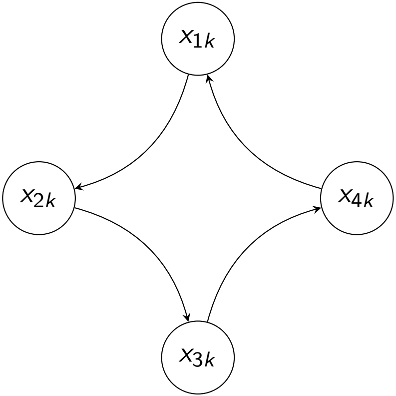
$(x_{2k} - x_{1k}) + (x_{3k} - x_{2k}) + \ldots + (x_{1k} - x_{4k})$

$=$

$-x_{1k} + (x_{2k} - x_{2k}) + \ldots + (x_{4k} - x_{4k}) + x_{1k}$

$=$

0



Cycle-based formulation: variables and constraints I

- new decision variables:

$$y_{ijk} = x_{ik} - x_{jk} \quad \forall k \leq K,\ \{i,j\} \in E \tag{5}$$

- bounds on the decision variables

$$\forall k \leq K,\ \{i,j\} \in E \quad -d_{ij} \leq y_{ijk} \leq d_{ij} \tag{6}$$

Cycle-based formulation: variables and constraints II

- EDGP Eq. (2) becomes the following set of constraints:

$$\forall \{i,j\} \in E \quad \sum_{k \le K} y_{ijk}^2 = d_{ij}^2 \tag{7}$$

- constraints on cycles

$$\forall k \le K, \ \boxed{C \subset G} \quad \Big(C \text{ is a cycle } \implies \sum_{\{i,j\} \in E} y_{ijk} = 0\Big) \tag{8}$$

Cycle-based formulation: main theorem

Theorem

$\exists$ a vector $y^* \in \mathbb{R}^{Km}$ which satisfies Eq. (7) and (8), parametrized on (K, G) $\iff$ (K, G) is a YES instance of the EDGP.

- *ergo*: find $y^* \longrightarrow$ any solution x^* of

$$\forall k \leq K, \{i, j\} \in E \quad x_{ik} - x_{jk} = y^*_{ijk} \tag{9}$$

  is valid realization of (K, G)

- issue: solving (7)–(8) is NP-hard: exponentially many constraints because of "$C \subset G$"

Cycle-based formulation: main theorem

Theorem

$\exists$ a vector $y^* \in \mathbb{R}^{Km}$ which satisfies Eq. (7) and (8), parametrized on $(K, G) \iff (K, G)$ is a YES instance of the EDGP.

- *ergo*: find $y^* \longrightarrow$ any solution x^* of

$$\forall k \leq K, \{i, j\} \in E \quad x_{ik} - x_{jk} = y^*_{ijk} \tag{9}$$

  is valid realization of (K, G)
- issue: solving (7)–(8) is NP-hard: exponentially many constraints because of "$C \subset G$"

Cycle-based formulation: main theorem (proof of $\Longleftarrow$)

G has realization x^*

- $\forall k \le K,\ \{i,j\} \in E,\ y_{ijk} = x^*_{ik} - x^*_{jk}$

$\Longrightarrow\ y^*$

- $\forall k \le K,\ \{i,j\} \in E,\ y^*_{ijk} = x^*_{ik} - x^*_{jk}$ (5)

Cycle-based formulation: main theorem (proof of $\Longleftarrow$)

G has realization x^*

- $\forall k \le K,\ \{i,j\} \in E,\ y_{ijk} = x^*_{ik} - x^*_{jk}$
- $\forall \{i,j\} \in E,\ \|x^*_i - x^*_j\|_2^2 = d_{ij}^2$ (2)

$\Longrightarrow\ y^*$

- $\forall k \le K,\ \{i,j\} \in E,\ y^*_{ijk} = x^*_{ik} - x^*_{jk}$ (5)
- $\forall \{i,j\} \in E,\ \sum_{k \le K} (y^*_{ijk})^2 = d_{ij}^2$ (7)

Cycle-based formulation: main theorem (proof of $\Longleftarrow$)

G has realization x^*

- $\forall k \le K,\ \{i,j\} \in E,\ y_{ijk} = x^*_{ik} - x^*_{jk}$
- $\forall \{i,j\} \in E,\ \|x^*_i - x^*_j\|_2^2 = d^2_{ij}$ (2)
- $\forall k \le K,\ C \subset G,\ \sum_{\{i,j\} \in E}(x^*_{ik} - x^*_{jk}) = 0$ (*Lemma*)

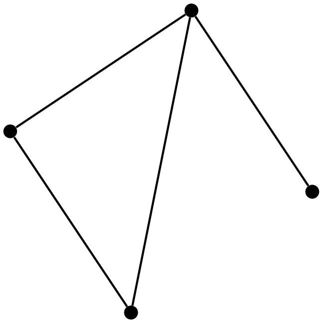
$\Longrightarrow y^*$

- $\forall k \le K,\ \{i,j\} \in E,\ y^*_{ijk} = x^*_{ik} - x^*_{jk}$ (5)
- $\forall \{i,j\} \in E,\ \sum_{k \le K} (y^*_{ijk})^2 = d^2_{ij}$ (7)
- $\forall k \le K,\ C \subset G,\ (C \text{ is a cycle } \Longrightarrow \sum_{\{i,j\} \in E} y^*_{ijk} = 0)$ (8)

Cycle-based formulation: main theorem (proof of $\implies$)

Use

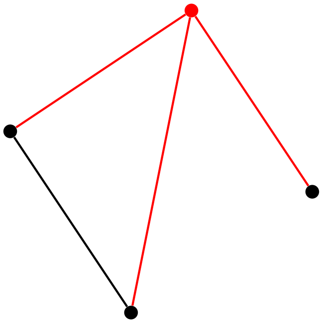
- tree: connected + no cycles
- biconnected graph: connected + no cut vertices



Cycle-based formulation: main theorem (proof of $\Longrightarrow$)

Use

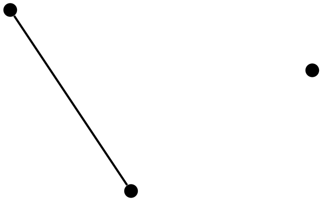
- tree: connected + no cycles
- biconnected graph: connected + no cut vertices



Cycle-based formulation: main theorem (proof of $\Longrightarrow$)

Use

- tree: connected + no cycles
- biconnected graph: connected + no cut vertices



Cycle-based formulation: main theorem (proof of $\Longrightarrow$)

1-DECOMPOSITION of $G = (V, E)$

Set of subgraphs $G_1, \ldots, G_r$, $r \in \mathbb{N}_{>0}$:

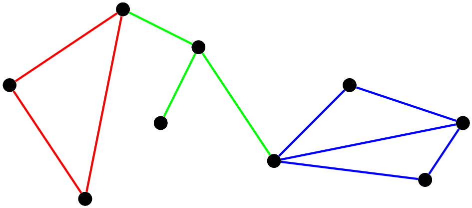
a. G_i biconnected or tree $i \leq r$
b. $\bigcup_{i \leq r} E(G_i) = E$
c. for any $i < j \leq r$, $V(G_i) \cap V(G_j)$ is $\emptyset$ or a cut vertex of G

Cycle-based formulation: main theorem (proof of $\Longrightarrow$)

1-DECOMPOSITION of $G = (V, E)$

Set of subgraphs $G_1, \ldots, G_r$, $r \in \mathbb{N}_{>0}$:

a. G_i biconnected or tree $i \leq r$
b. $\bigcup_{i \leq r} E(G_i) = E$
c. for any $i < j \leq r$, $V(G_i) \cap V(G_j)$ is $\emptyset$ or a cut vertex of G



$$\forall k \leq K,\ \{i, j\} \in E, \qquad y^*_{ijk} = x_{ik} - x_{jk}$$

Cycle-based formulation: enter cycle bases [Kavitha et al., 2009]

Problematic constraints

$$\forall k \leq K, \; \boxed{C \subset G} \quad \left(C \text{ is a cycle} \implies \sum_{\{i,j\} \in E} y_{ijk} = 0\right) \qquad (8)$$

BUT ...

... incidence vectors of cycles (in Euclidean space of dimension $|E|$) form vector space over field $\mathbb{F}$

SO ...

... use cycle bases $\implies \forall C$: weighted sum of cycles in basis $\mathcal{B}$

Cycle-based formulation: enter cycle bases

Easier constraints

$$\forall k \le K, \ \boxed{\forall B \in \mathcal{B}} \quad \ldots$$

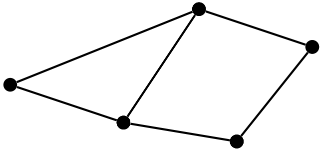
BUT ...

...incidence vectors of cycles (in Euclidean space of dimension $|E|$) form vector space over field $\mathbb{F}$

SO ...

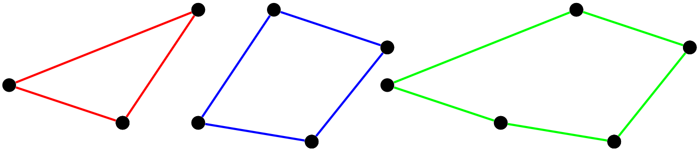
...use cycle bases $\Longrightarrow$ $\forall C$: weighted sum of cycles in basis $\mathcal{B}$

Cycle bases [Kavitha et al., 2009]



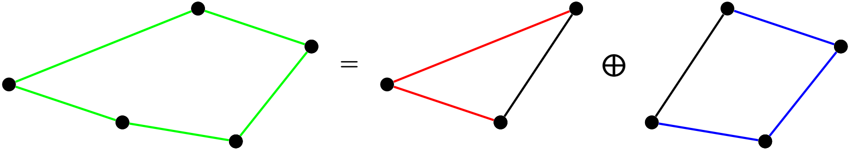
Cycle bases [Kavitha et al., 2009]

Each cycle encoded by a vector in $c \in \mathbb{Z}_2^m,\ m = |E|$



Cycle bases [Kavitha et al., 2009]

Sum operation



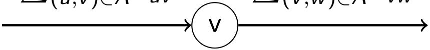
Characteristics:

- node degree is even
- vector space does not depend on orientation of G
- cases of interest: $\mathbb{Z}_2^m$, $\mathbb{Q}^m$

Cycle-based formulation: rewriting Eqs. (8)

- direct edges in $E \longrightarrow$ directed simple graph $\bar{G}(V, A)$
- $c^{\bar{C}} \in \mathbb{R}^m$: incidence vector of directed cycle $\bar{C} \in \bar{G}$; $c^{\bar{C}} \in \{0, 1, -1\}$ if $\bar{C}$ is circuit;

$$\sum_{(u,v)\in A} c^{\bar{C}}_{uv} = \sum_{(v,w)\in A} c^{\bar{C}}_{vw}$$



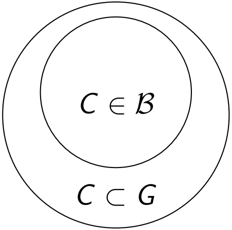
Proposition

Let $\mathcal{B}$ directed cycle basis of $\bar{G}$ over $\mathbb{Q}$. Eq. (8) holds $\iff$

$$\forall k \le K,\ \forall \boldsymbol{B} \in \boldsymbol{\mathcal{B}} \quad \sum_{(i,j)\in A(B)} c^B_{ij} y_{ijk} = 0 \tag{10}$$

Cycle-based formulation: rewriting Eq. (8) (proof)

$(\implies)$

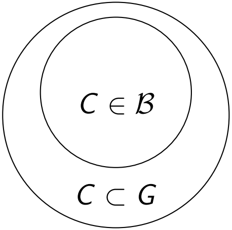


Cycle-based formulation: rewriting Eq. (8) (proof)

$(\Longleftarrow)$

$$\boxed{\bar{C} \subset \bar{G} : \quad c^{\bar{C}} = \sum_{B \in \mathcal{B}} \gamma_B c^B}$$

$(\Longrightarrow)$

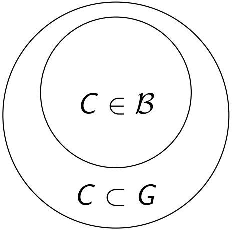


Cycle-based formulation: rewriting Eq. (8) (proof)

$(\Longleftarrow)$

$$\boxed{\bar{C} \subset \bar{G}: \quad c^{\bar{C}} = \sum_{B \in \mathcal{B}} \gamma_B c^B}$$

$(\Longrightarrow)$



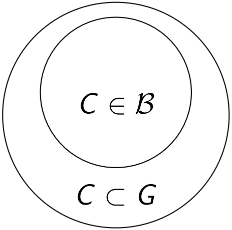
$$\forall k \le K: \quad \boxed{\sum_{B \in \mathcal{B}} \gamma_B \sum_{(i,j) \in A(B)} c_{ij}^B y_{ijk}}$$

Cycle-based formulation: rewriting Eq. (8) (proof)

$(\Longleftarrow)$

$$\boxed{\bar{C} \subset \bar{G}: \quad c^{\bar{C}} = \sum_{B \in \mathcal{B}} \gamma_B c^B}$$

$(\Longrightarrow)$



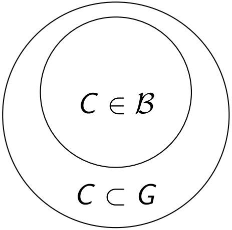
$$\boxed{\forall k \leq K: \quad 0 = \boxed{\sum_{B \in \mathcal{B}} \gamma_B \sum_{(i,j) \in A(B)} c_{ij}^B y_{ijk}}}$$

Cycle-based formulation: rewriting Eq. (8) (proof)

$$(\Longleftarrow)$$

$$\boxed{\bar{C} \subset \bar{G}: \quad c^{\bar{C}} = \sum_{B \in \mathcal{B}} \gamma_B c^B}$$

$$(\Longrightarrow)$$



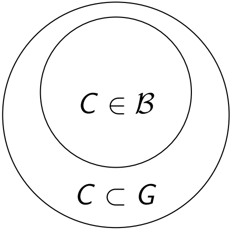
$$\boxed{\begin{aligned} \forall k \le K: \quad 0 &= \boxed{\sum_{B \in \mathcal{B}} \gamma_B \sum_{(i,j) \in A(B)} c_{ij}^B y_{ijk}} \\ &= \sum_{(i,j) \in A(\bar{C})} c_{ij}^{\bar{C}} y_{ijk} \end{aligned}}$$

Cycle-based formulation: rewriting Eq. (8) (proof)

$$(\Longleftarrow)$$

$$\boxed{\bar{C} \subset \bar{G}: \quad c^{\bar{C}} = \sum_{B \in \mathcal{B}} \gamma_B c^B}$$

$$(\Longrightarrow)$$



$$\boxed{\begin{aligned}\forall k \le K: \quad 0 &= \boxed{\sum_{B \in \mathcal{B}} \gamma_B \sum_{(i,j) \in A(B)} c_{ij}^B y_{ijk}} \\ &= \sum_{(i,j) \in A(\bar{C})} c_{ij}^{\bar{C}} y_{ijk}\end{aligned}}$$

$$\boxed{\forall k \le K: \quad \sum_{\{i,j\} \in E(C)} y_{ijk} = 0}$$

Cycle-based formulation: final make-over

Valid formulation for EDGP:

$$\left.\begin{array}{rl} \min\limits_{s\geq 0,\ y} & \sum\limits_{\{i,j\}\in E} (s_{ij}^{+} + s_{ij}^{-}) \\ \forall (i,j)\in A(\bar{G}) & \sum\limits_{k\leq K} y_{ijk}^2 - d_{ij}^2 = s_{ij}^{+} - s_{ij}^{-} \\ \forall k\leq K,\ B\in\mathcal{B} & \sum\limits_{(i,j)\in A(B)} c_{ij}^{B} y_{ijk} = 0 \end{array}\right\} \qquad (11)$$

Bonus: how to find directed cycle basis in $\mathbb{Q}$

- can be obtained from undirected cycle basis of G: use directions in $\bar{G}$ [Kavitha et al., 2009];
- algorithm used to find *Fundamental Cycle Basis* [Paton, 1969]:
  1. find spanning tree;
  2. pick $m - n + 1$ circuits that each nontree (*chord*) edge defines with tree

Results

cycle-based formulation

$$\left.\begin{array}{rl} \min\limits_{s\geq 0,\ y} & \sum\limits_{\{i,j\}\in E} (s^+_{ij} + s^-_{ij}) \\ \forall (i,j)\in A(\bar{G}) & \sum\limits_{k\leq K} y^2_{ijk} - d^2_{ij} = s^+_{ij} - s^-_{ij} \\ \forall k\leq K,\ B\in\mathcal{B} & \sum\limits_{(i,j)\in A(B)} c^B_{ij} y_{ijk} = 0 \end{array}\right\}$$

Results

cycle-based formulation

$$\left.\begin{array}{rl} \min\limits_{s\geq 0,\ y} & \sum\limits_{\{i,j\}\in E} (s_{ij}^{+} + s_{ij}^{-}) \\ \forall (i,j) \in A(\bar{G}) & \sum\limits_{k\leq K} y_{ijk}^2 - d_{ij}^2 = s_{ij}^{+} - s_{ij}^{-} \\ \forall k \leq K,\ B \in \mathcal{B} & \sum\limits_{(i,j)\in A(B)} c_{ij}^{B} y_{ijk} = 0 \end{array}\right\}$$

edge-based formulation

$$\left.\begin{array}{rl} \min\limits_{x} & \sum_{\{i,j\}\in E} (\|x_i - x_j\|_2^2 - d_{ij}^2)^2 \\ \forall k \leq K & \sum\limits_{i\leq n} x_{ik} = 0 \end{array}\right\}$$

Results

- 3-iteration multi-start: at each iteration a) call local NLP `IpOpt` solver from random starting point; b) update best incumbent solution
- formulations/heuristics implemented in AMPL
- Computed measures:
  - $\mathrm{MDE}(x, G) = \frac{1}{|E|} \big| \|x_i - x_j\|_2 - d_{ij} \big|$ (average error)
  - $\mathrm{LDE}(x, G) = \max\limits_{\{i,j\} \in E} \big| \|x_i - x_j\|_2 - d_{ij} \big|$ (max error)
  - CPU time

Results: cycle formulation vs. edge formulation on proteins

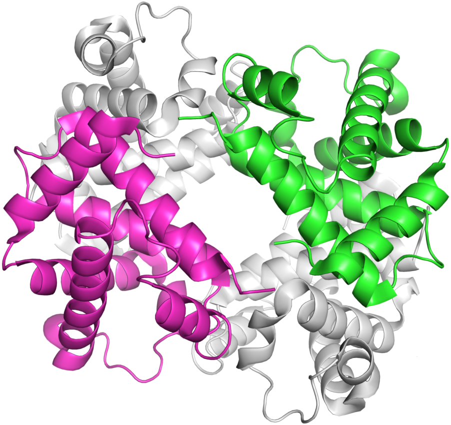


Figure: Deoxyhemoglobin (red blood cells)

Results: cycle formulation vs. edge formulation on proteins

Instance	*m*	*n*	mdeC	mdeE	ldeC	ldeE	cpuC	cpuE
1guu	955	150	**0.057**	0.061	1.913	**1.884**	**18.18**	37.14
1guu-1	959	150	**0.035**	0.038	2.025	**1.824**	24.27	**5.48**
1guu-4000	968	150	0.061	**0.060**	2.324	**2.121**	24.24	**6.97**
pept	999	107	**0.104**	0.161	3.367	**2.963**	34.67	**10.89**
2kxa	2711	177	**0.053**	0.155	**3.613**	3.936	169.95	**35.44**
res_2kxa	2627	177	0.131	**0.045**	**3.197**	3.442	153.00	**32.40**
C0030pkl	3247	198	**0.009**	0.059	**2.761**	3.965	156.09	**76.58**
cassioli-130731	4871	281	**0.005**	0.060	**3.447**	3.963	376.33	**143.31**
100d	5741	488	**0.146**	0.246	4.295	**4.090**	3024.67	**253.56**
helix_amber	6265	392	**0.038**	0.059	**3.528**	4.578	1573.10	**212.68**
water	11939	648	**0.222**	0.422	4.557	**4.322**	9384.08	**3836.23**
3al1	17417	678	**0.084**	0.124	4.165	**4.087**	4785.91	**1467.74**
1hpv	18512	1629	**0.334**	0.338	**4.256**	4.619	53848.33	**6620.70**
il2	45251	2084	1.481	**0.248**	9.510	**4.415**	**2323.90**	24321.25

Table: Performance on protein graphs ($K = 3$)

Results: cycle formulation vs. edge formulation

Instance	*m*	*n*	mdeC	mdeE	ldeC	ldeE	cpuC	cpuE
almostreg-3-100	298	100	**0**	**0**	0.048	**0.041**	0.88	**0.23**
almostreg-3-150	448	150	**0**	**0**	0.330	**0.282**	1.29	**0.30**
almostreg-3-200	598	200	**0**	**0**	0.030	**0.020**	2.15	**0.44**
almostreg-3-50	146	50	**0**	**0**	**0**	**0**	0.31	**0.11**
almostreg-6-100	591	100	**0.077**	0.093	0.740	**0.410**	6.85	**0.35**
almostreg-6-150	893	150	**0.085**	0.099	1.030	**0.485**	16.52	**0.68**
almostreg-6-200	1192	200	**0.076**	0.098	0.729	**0.501**	34.07	**1.35**
almostreg-6-50	292	50	**0.082**	0.099	0.648	**0.471**	1.80	**0.13**
almostreg-8-100	777	100	**0.105**	0.131	0.846	**0.577**	8.89	**0.42**
almostreg-8-150	1189	150	**0.104**	0.121	0.805	**0.528**	34.84	**0.83**
almostreg-8-200	1581	200	**0.104**	0.125	0.974	**0.654**	48.10	**1.79**
almostreg-8-50	387	50	**0.104**	0.113	0.670	**0.520**	2.46	**0.13**
bipartite-100-03	3044	200	**0.206**	0.218	0.931	**0.790**	209.15	**7.86**
bipartite-100-06	6024	200	**0.225**	0.234	0.978	**0.753**	439.74	**8.00**
bipartite-150-03	6708	300	**0.220**	0.232	0.951	**0.724**	582.71	**14.37**
bipartite-150-06	13466	300	**0.231**	0.240	0.852	**0.808**	1904.18	**30.79**
bipartite-200-03	11906	400	**0.223**	0.235	0.936	**0.812**	3183.43	**33.06**
bipartite-200-06	23963	400	**0.235**	0.244	0.888	**0.741**	4885.52	**64.03**
bipartite-50-03	744	100	**0.166**	0.185	0.936	**0.787**	29.27	**1.11**
bipartite-50-06	1468	100	**0.201**	0.217	1.011	**0.754**	80.80	**1.38**

Table: Performance on small sized graphs ($K = 2$)

Results: cycle formulation vs. edge formulation

Instance	m	n	mdeC	mdeE	ldeC	ldeE	cpuC	cpuE
cluster-120-4-05-01	1495	120	**0.191**	0.206	0.873	**0.838**	98.67	**1.69**
cluster-120-8-05-01	1149	120	**0.181**	0.196	0.892	**0.740**	62.29	**1.04**
cluster-150-2-05-01	3337	150	**0.218**	0.230	**0.901**	0.936	605.00	**3.66**
cluster-150-8-05-01	1750	150	**0.190**	0.205	0.886	**0.831**	70.66	**2.44**
cluster-200-2-05-01	5957	200	**0.231**	0.241	**0.931**	0.952	612.82	**8.01**
cluster-200-4-05-01	4155	200	**0.221**	0.233	0.924	**0.906**	397.45	**7.67**
cluster-200-8-05-01	3046	200	**0.206**	0.220	0.988	**0.851**	462.46	**5.61**
cluster-50-2-05-01	361	50	**0.159**	0.171	0.742	**0.679**	7.52	**0.20**
cluster-50-4-05-01	242	50	**0.145**	0.167	0.899	**0.588**	3.63	**0.18**
cluster-50-8-05-01	187	50	**0.113**	0.133	0.716	**0.500**	2.73	**0.16**
euclid-150-02	2341	150	**0**	**0**	**0**	**0**	286.09	**2.69**
euclid-150-05	5678	150	**0**	**0**	**0**	**0**	991.87	**2.86**
euclid-150-08	8915	150	**0**	**0**	**0**	**0**	1507.94	**3.88**
euclid-200-05	10037	200	**0**	**0**	**0**	**0**	1881.40	**5.47**
euclid-200-08	15877	200	**0**	**0**	**0**	**0**	3114.95	**7.96**
flowersnark120	720	480	**0**	**0**	0.151	**0.109**	**7.86**	8.21
flowersnark-150	900	600	**0**	**0**	0.101	**0.086**	36.53	**15.50**
flowersnark-200	1200	800	**0**	**0**	0.141	**0.123**	**18.02**	31.04
flowersnark40	240	160	**0**	**0**	0.016	**0.005**	1.92	**0.35**
flowersnark80	480	320	**0**	**0**	0.068	**0.059**	3.18	**1.08**
hypercube-10	5120	1024	**0.128**	0.152	1.004	**0.653**	4965.30	**133.93**
hypercube-5	80	32	**0.054**	0.058	0.401	**0.321**	0.95	**0.10**
hypercube-6	192	64	**0.075**	0.087	0.774	**0.426**	4.20	**0.20**
hypercube-8	1024	256	**0.104**	0.127	0.876	**0.631**	81.68	**2.59**

Table: Performance on small sized graphs ($K = 2$)

Results: cycle formulation vs. edge formulation

Instance	*m*	*n*	mdeC	mdeE	ldeC	ldeE	cpuC	cpuE
powerlaw-100-2-05	148	100	**0.024**	0.025	0.338	**0.309**	1.24	**0.38**
powerlaw-100-2-08	178	100	**0.042**	**0.042**	0.464	**0.398**	1.64	**0.59**
powerlaw-150-2-05	223	150	**0.034**	0.035	0.404	**0.360**	**1.37**	1.94
powerlaw-150-2-08	268	150	**0.047**	**0.047**	0.471	**0.404**	2.44	**1.73**
powerlaw-200-2-05	298	200	**0.025**	0.026	0.581	**0.443**	2.64	**1.27**
powerlaw-200-2-08	358	200	**0.037**	0.038	0.454	**0.376**	3.75	**1.78**
random-100-02	1093	100	**0.193**	0.203	0.874	**0.742**	48.43	**0.67**
random-100-05	2479	100	**0.224**	0.234	0.938	**0.855**	168.40	**1.48**
random-150-02	2394	150	**0.209**	0.223	0.932	**0.809**	226.60	**3.98**
random-150-05	5675	150	**0.241**	0.250	0.965	**0.953**	580.59	**6.10**
random-200-02	4097	200	**0.218**	0.228	0.930	**0.887**	271.94	**7.68**
random-200-05	10023	200	**0.248**	0.255	**0.949**	0.952	1024.32	**11.43**
random-50-02	291	50	**0.143**	0.161	0.922	**0.638**	7.03	**0.17**
random-50-05	665	50	**0.195**	0.212	**0.836**	0.953	16.20	**0.23**
rnddegdist-100	2252	100	**0.223**	0.235	**0.929**	0.963	136.74	**1.48**
rnddegdist-150	5293	150	**0.240**	0.249	**0.939**	0.955	819.86	**3.91**
rnddegdist-30	174	30	**0.156**	0.179	0.767	**0.667**	2.26	**0.11**
rnddegdist-40	221	40	**0.156**	0.175	0.672	**0.628**	2.93	**0.17**
tripartite-100-02	4038	300	**0.198**	0.213	0.968	**0.737**	369.77	**10.39**
tripartite-100-05	10003	300	**0.227**	0.238	0.917	**0.729**	1150.35	**21.37**
tripartite-150-02	9061	450	**0.213**	0.227	0.956	**0.765**	2005.30	**32.43**
tripartite-150-05	22431	450	**0.235**	0.245	0.876	**0.751**	4687.28	**45.27**
tripartite-30-02	359	90	**0.106**	0.118	0.736	**0.547**	10.31	**0.37**
tripartite-50-02	995	150	**0.153**	0.173	0.958	**0.722**	38.55	**1.00**
tripartite-50-05	2519	150	**0.208**	0.220	0.849	**0.736**	160.43	**2.39**

Table: Performance on small sized graphs ($K = 2$)

Aspnes, J., Eren, T., Goldenberg, D., Morse, S., Whiteley, W., Yang, R., Anderson, B., and Belhumeur, P. (2006).
A theory of network localization.
IEEE Transactions on Mobile Computing, 5(12):1663–1678.

Beeker, N., Gaubert, S., Glusa, C., and Liberti, L. (2013).
Is the distance geometry problem in NP?
In Mucherino, A., Lavor, C., Liberti, L., and Maculan, N., editors, *Distance Geometry: Theory, Methods, and Applications*, pages 85–94. Springer, New York.

Kavitha, T., Liebchen, C., Mehlhorn, K., Michail, D., Rizzi, R., Ueckerdt, T., and Zweig, K. (2009).
Cycle bases in graphs: characterization, algorithms, complexity, and applications.
Computer Science Review, 3:199–243.

Lavor, C., Liberti, L., and Maculan, N. (2009).
Molecular distance geometry problem.
In Floudas, C. and Pardalos, P., editors, *Encyclopedia of Optimization*, pages 2305–2311. Springer, New York, second edition.

Paton, K. (1969).
An algorithm for finding a fundamental set of cycles of a graph.
Communications of the ACM, 12(9):514–518.

Saxe, J. (1979).
Embeddability of weighted graphs in k-space is strongly NP-hard.
Proceedings of 17th Allerton Conference in Communications, Control and Computing, pages 480–489.

Singer, A. (2011).
Angular synchronization by eigenvectors and semidefinite programming.
Applied and Computational Harmonic Analysis, 30:20–36.