Exploiting Sparsity in Matrix Inequality and Its Application to Polynomial SDP

Workshop of Computational Polynomial Optimization and Multilinear Algebra

The Hong Kong Polytechnic University
December 22-23, 2008

Masakazu Kojima

Tokyo Institute of Technology

- Joint work with S. Kim, Martin Mevissen and M. Yamashita

Motivation

A general framework for exploiting sparsity, which is characterized by

- a chordal graph structure, in
- nonlinear optimization problems involving matrix inequalities, via
- positive semidefinite matrix completion.

In this talk, we focus our attention to polynomial SDPs and their relaxation to linear SDPs.

Some preliminary numerical results on linear SDP relaxation of quadratic SDPs

Exploiting sparsity characterized by a chordal graph structure in polynomial SDPs via psd matrix completion

Polynomial SDP

each large matrix variable

↓ exploiting **domain-sparsity**

multiple smaller matrix variables

each large matrix inequality

↓ exploiting **range-sparsity**

multiple smaller matrix inequalities

↓ **sparse** SDP relaxation (Lasserre, **Kojima et al.**)

Linear SDP with multiple smaller matrix variables and multiple smaller matrix inequalities satisfying **correlative sparsity**

‖

sparsity of the Schur complement matrix *

* the positive definite coefficient matrix of a system of linear equations solved by the Cholesky factorization at each iteration of the p-d ipm for a search direction

Contents

Contents

1. An Extremely Sparse Example
2. Positive Semidefinite Matrix Completion
3. Duality in Positive Semidefinite Matrix Completion
4. Linear SDP relaxations of quadratic SDPs
5. Sensor network localization problems
6. Concluding Remarks

SDP: min $\sum_{|i-j|\leq 1} C_{ij}X_{ij}$ sub. to $\boldsymbol{M}(\boldsymbol{X}) \succeq \boldsymbol{O}$, $\boldsymbol{X} \succeq \boldsymbol{O}$, where

$$\boldsymbol{M}(\boldsymbol{X}) = \begin{pmatrix} 1 - X_{11} & 0 & \ldots & X_{12} \\ 0 & 1 - X_{22} & \ldots & X_{23} \\ \vdots & \vdots & \ddots & \vdots \\ X_{21} & X_{32} & \ldots & 1 - X_{nn} \end{pmatrix}$$

Two kinds of sparsity

(a) $\boldsymbol{M}(\boldsymbol{X})$ does not involve any X_{ij} with $|i-j| > 1$
— domain-space sparsity.

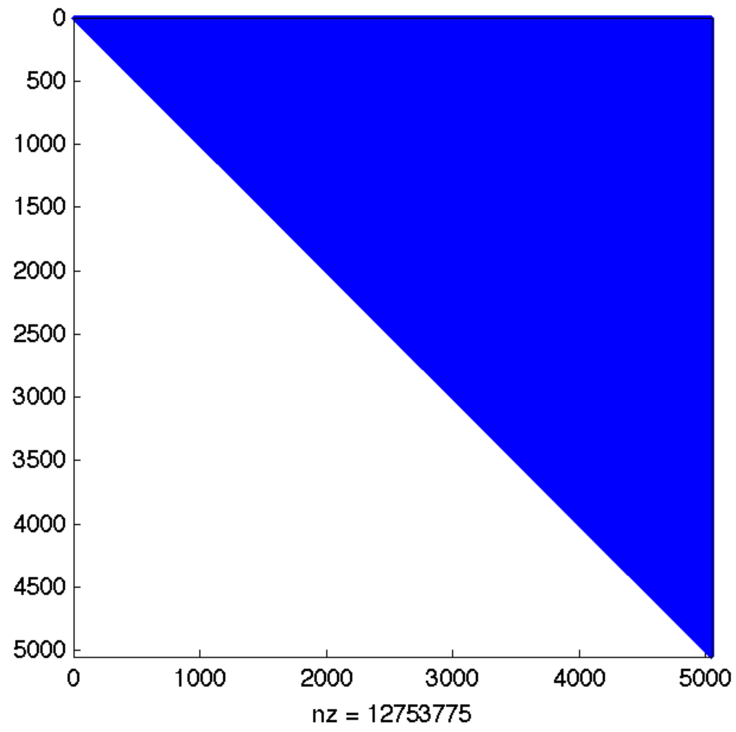
(b) $\boldsymbol{M}(\boldsymbol{X})$ is "diagonal + bordered" — range-space sparsity.

	SeDuMi cpu time in second for SDP relaxation (Size of Schur comp. mat., max. block size)	
n	dense reformulation	sparse reformulation
50	29.07 (1275, 50)	0.61 (147, 2)
100	⇒1797.49 (5050, 100)	0.97 (297, 2)
1000		6.62 (2997, 2)
10000		⇒192.02 (29997, 2)
	S.comp.mat. : fully dense	S.comp.mat. : sparse

The Cholesky factorization of the Schur complement matrix

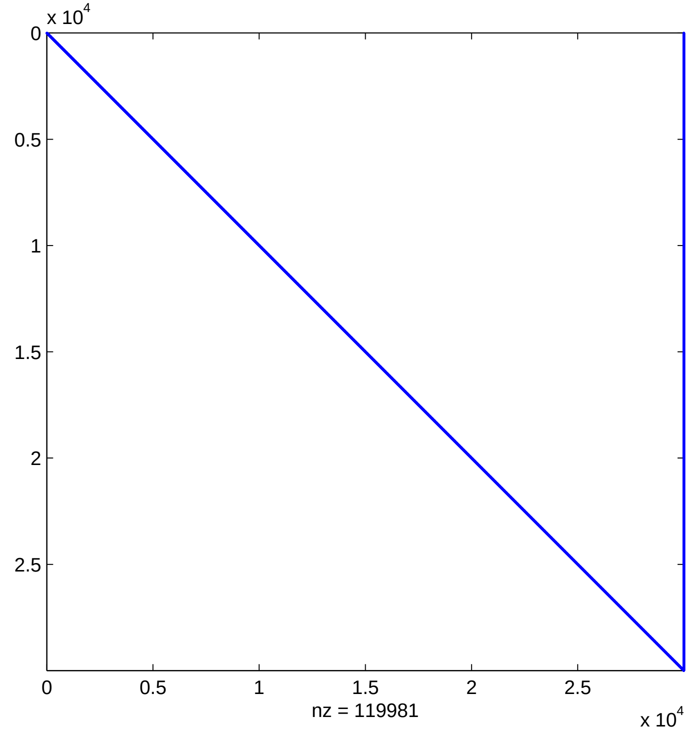
dense reformulation vs sparse reformulation

$(n = 100)$ $(n = 10,000)$



$5,050 \times 5,050$

nonzeros = 12,753,775



$29,997 \times 29,997$

nonzeros = 119,981

Contents

$G(N,E)$	:	a graph with $N=\{1,\ldots,n\}$, $E \subseteq N \times N$; $(i,i) \notin E$, $(i,j)=(j,i) \in E,\ E^\bullet = E \cup \{(i,i) : i \in N\}$.
$\mathbb{S}^n\ (\mathbb{S}^n_+)$	$=$	the set of $n \times n$ (psd) symmetric matrices.
$\mathbb{S}^n(E,?)$	$=$	the set of $n \times n$ partial symmetric matrices with entries specified in $E^\bullet$.
$\mathbb{S}^n_+(E,?)$	$=$	$\{\boldsymbol{X} \in \mathbb{S}^n(E,?)$ which can be psd$\}$
	$=$	$\{\boldsymbol{X} \in \mathbb{S}^n(E,?) : \exists \overline{\boldsymbol{X}} \in \mathbb{S}^n_+; \overline{X}_{ij} = X_{ij}$ if $(i,j) \in E^\bullet\}$.

PSD Matrix Completion Problem: Complete $\boldsymbol{X} \in \mathbb{S}^n(E,?)$ to an $\overline{\boldsymbol{X}} \in \mathbb{S}^n_+$ satisfying $\overline{X}_{ij} = X_{ij}$ $((i,j) \in E$ if it exists.

Example:

$$N=\{1,2,3\},\ E=\{(1,2),\ (2,3)\},\ \boldsymbol{X} = \begin{pmatrix} 3 & 3 & ? \\ 3 & 3 & 2 \\ ? & 2 & 2 \end{pmatrix} \in \mathbb{S}^3(E,?),$$

which is completed to $\overline{\boldsymbol{X}} = \begin{pmatrix} 3 & 3 & 2 \\ 3 & 3 & 2 \\ 2 & 2 & 2 \end{pmatrix} \in \mathbb{S}^3_+$. $\quad ? = 0 \not\Rightarrow$ psd.

$G(N,E)$: a graph with $N = \{1,\ldots,n\}$, $E \subseteq N \times N$; $(i,i) \notin E$, $(i,j) = (j,i) \in E$, $E^\bullet = E \cup \{(i,i) : i \in N\}$.

$\mathbb{S}^n$ ($\mathbb{S}^n_+$) = the set of $n \times n$ (psd) symmetric matrices.

$\mathbb{S}^n(E,?)$ = the set of $n \times n$ partial symmetric matrices with entries specified in $E^\bullet$.

$\mathbb{S}^n_+(E,?)$ = $\{\boldsymbol{X} \in \mathbb{S}^n(E,?)$ which can be psd$\}$.

(Nonlinear) SDP: min $f_0(\boldsymbol{y},\boldsymbol{X})$ sub. to $\boldsymbol{f}(\boldsymbol{y},\boldsymbol{X}) \in \boldsymbol{\Omega}$, $\boldsymbol{X} \in \mathbb{S}^n_+$

$\Downarrow$ Values of X_{ij} ($(i,j) \notin E^\bullet$) are not relevant except $\boldsymbol{X} \in \mathbb{S}^n_+$

- Replace the psd condition "$\boldsymbol{X} \in \mathbb{S}^n_+$" by "$\boldsymbol{X} \in \mathbb{S}^n_+(E,?)$".

$\Downarrow$ G is chordal ($\forall$ cycle with more than 3 edges has a chord)

- "$\boldsymbol{X} \in \mathbb{S}^n_+(E,?)$" $\Leftrightarrow$ "$\boldsymbol{X}(C_k) \succeq \boldsymbol{O}$ $(k = 1,\ldots,\ell)$", where C_k $(k = 1,\ldots,\ell)$ denotes the max. cliques of G, and $\boldsymbol{X}(C_k)$ the submatrix of X_{ij} ($(i,j) \in C_k$) (Grone et. al 1984).
- "$\boldsymbol{X} \in \mathbb{S}^n_+$" in SDP $\Rightarrow$ "$\boldsymbol{X}(C_k) \succeq \boldsymbol{O}$ $(k = 1,\ldots,\ell)$".

$G(N,E)$: a graph with $N = \{1,\ldots,n\}$, $E \subseteq N \times N$; $(i,i) \notin E$, $(i,j) = (j,i) \in E,\ \ E^{\bullet} = E \cup \{(i,i) : i \in N\}$.

$\mathbb{S}^n$ ($\mathbb{S}^n_+$) = the set of $n \times n$ (psd) symmetric matrices.

$\mathbb{S}^n(E,?)$ = the set of $n \times n$ partial symmetric matrices with entries specified in $E^{\bullet}$.

$\mathbb{S}^n_+(E,?)$ = $\{\boldsymbol{X} \in \mathbb{S}^n(E,?)$ which can be psd$\}$.

---

min $\boldsymbol{A}_0 \bullet \boldsymbol{X}$ sub.to $\boldsymbol{A}_p \bullet \boldsymbol{X} = b_p\ (p = 1,2,\ldots,m),\ \ \boldsymbol{X} \succeq \boldsymbol{O}$

$N = \{1,2,3,4\}$
$E = \{(i,j) \in N \times N : |i-j| = 1\}$

the sparsity pattern of $\forall \boldsymbol{A}_0,\ \boldsymbol{A}_p = \begin{pmatrix} \star & \star & 0 & 0 \\ \star & \star & \star & 0 \\ 0 & \star & \star & \star \\ 0 & 0 & \star & \star \end{pmatrix}$

$\Downarrow\ \ E^{\bullet} = \{(i,j) \in N \times N : |i-j| \leq 1\}$

$G(N,E)$: 

chordal, max cliques $C_1 = \{1,2\}, C_2 = \{2,3\}, C_3 = \{3,4\}$

$\bullet\ \bullet\ \bullet\ \bullet\ \boldsymbol{X} \in \mathbb{S}^n_+(E,?) \implies \bullet\ \bullet\ \bullet\ \bullet\ \boldsymbol{X}(C_k) \succeq \boldsymbol{O}\ (k = 1,2,3)$

- Tridiagonal case $\Rightarrow$ chordal.

$G(N,E)$: a graph with $N = \{1,\ldots,n\}$, $E \subseteq N \times N$; $(i,i) \notin E$, $(i,j) = (j,i) \in E$, $E^\bullet = E \cup \{(i,i) : i \in N\}$.

$\mathbb{S}^n$ ($\mathbb{S}^n_+$) = the set of $n \times n$ (psd) symmetric matrices.

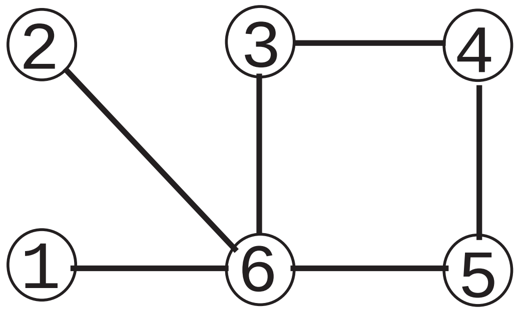
$\mathbb{S}^n(E,?)$ = the set of $n \times n$ partial symmetric matrices with entries specified in $E^\bullet$.

$\mathbb{S}^n_+(E,?)$ = $\{\boldsymbol{X} \in \mathbb{S}^n(E,?)$ which can be psd$\}$.

---

min $\boldsymbol{A}_0 \bullet \boldsymbol{X}$ sub.to $\boldsymbol{A}_p \bullet \boldsymbol{X} = b_p \ (p = 1,2,\ldots,m)$, $\boldsymbol{X} \succeq \boldsymbol{O}$

$$\boldsymbol{A}_0,\ \boldsymbol{A}_p \sim \begin{pmatrix} \star & 0 & 0 & 0 & 0 & \star \\ 0 & \star & 0 & 0 & 0 & \star \\ 0 & 0 & \star & \star & 0 & \star \\ 0 & 0 & \star & \star & \star & 0 \\ 0 & 0 & 0 & \star & \star & \star \\ \star & \star & \star & 0 & \star & \star \end{pmatrix}$$



not chordal

$G(N,E)$: a graph with $N = \{1,\ldots,n\}$, $E \subseteq N \times N$; $(i,i) \notin E$, $(i,j) = (j,i) \in E,\ E^\bullet = E \cup \{(i,i) : i \in N\}$.

$\mathbb{S}^n$ ($\mathbb{S}^n_+$) = the set of $n \times n$ (psd) symmetric matrices.

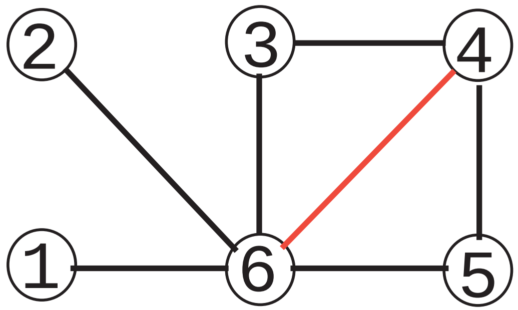
$\mathbb{S}^n(E,?)$ = the set of $n \times n$ partial symmetric matrices with entries specified in $E^\bullet$.

$\mathbb{S}^n_+(E,?)$ = $\{\boldsymbol{X} \in \mathbb{S}^n(E,?)$ which can be psd$\}$.

---

min $\boldsymbol{A}_0 \bullet \boldsymbol{X}$ sub.to $\boldsymbol{A}_p \bullet \boldsymbol{X} = b_p \ (p = 1,2,\ldots,m),\ \boldsymbol{X} \succeq \boldsymbol{O}$

$$\boldsymbol{A}_0,\ \boldsymbol{A}_p \sim \begin{pmatrix} \star & 0 & 0 & 0 & 0 & \star \\ 0 & \star & 0 & 0 & 0 & \star \\ 0 & 0 & \star & \star & 0 & \star \\ 0 & 0 & \star & \star & \star & \star \\ 0 & 0 & 0 & \star & \star & \star \\ \star & \star & \star & \star & \star & \star \end{pmatrix}$$



chordal

$\{3,4,6\},\ \{4,5,6\},\ \{1,6\},$ $\{2,6\}$: max. cliques C_k.

- $\boldsymbol{X} \in \mathbb{S}^n_+(E,?) \Leftrightarrow \boldsymbol{X}(C_1),\ \boldsymbol{X}(C_2),\ \boldsymbol{X}(C_3),\ \boldsymbol{X}(C_4) \succeq \boldsymbol{O}$

$G(N,E)$: a graph with $N = \{1, \ldots, n\}$, $E \subseteq N \times N$; $(i,i) \notin E$,
$(i,j) = (j,i) \in E,\ E^{\bullet} = E \cup \{(i,i) : i \in N\}$.

$\mathbb{S}^n$ ($\mathbb{S}^n_+$) = the set of $n \times n$ (psd) symmetric matrices.

$\mathbb{S}^n(E,?)$ = the set of $n \times n$ partial symmetric matrices with entries specified in $E^{\bullet}$.

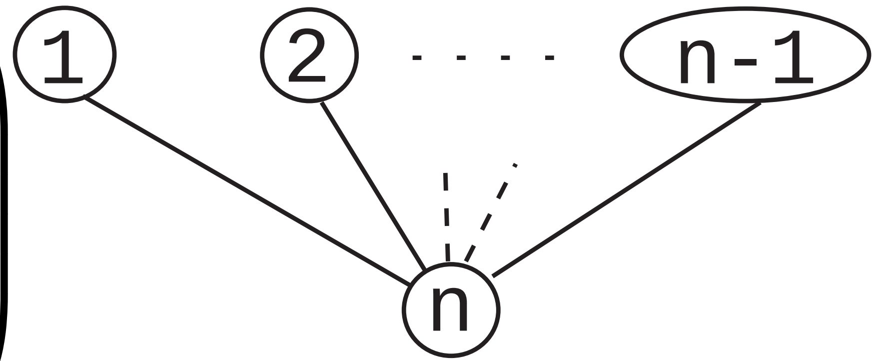
$\mathbb{S}^n_+(E,?)$ = $\{\boldsymbol{X} \in \mathbb{S}^n(E,?)$ which can be psd$\}$.

---

min $\boldsymbol{A}_0 \bullet \boldsymbol{X}$ sub.to $\boldsymbol{A}_p \bullet \boldsymbol{X} = b_p\ (p = 1, 2, \ldots, m),\ \boldsymbol{X} \succeq \boldsymbol{O}$

Diagonal-bordered case

$$\boldsymbol{A}_0,\ \boldsymbol{A}_p \sim \begin{pmatrix} \star & 0 & \ldots & \star \\ 0 & \star & \ldots & \star \\ \ldots & \ldots & \ddots & \ldots \\ \star & \star & \ldots & \star \end{pmatrix}$$



$G = (N,E)$

- $C_k = \{k, n\}\ (k = 1, \ldots, n-1)$: the max. cliques.
- $\boldsymbol{X} \in \mathbb{S}^n_+(E,?) \Leftrightarrow \boldsymbol{X}(C_k) \succeq \boldsymbol{O}\ (k = 1, \ldots, n-1)$ — 2×2.

$G(N,E)$: a graph with $N = \{1,\ldots,n\}$, $E \subseteq N \times N$; $(i,i) \notin E$, $(i,j) = (j,i) \in E$, $E^\bullet = E \cup \{(i,i) : i \in N\}$.

$\mathbb{S}^n$ ($\mathbb{S}^n_+$) = the set of $n \times n$ (psd) symmetric matrices.

$\mathbb{S}^n(E,?)$ = the set of $n \times n$ partial symmetric matrices with entries specified in $E^\bullet$.

$\mathbb{S}^n_+(E,?)$ = $\{\boldsymbol{X} \in \mathbb{S}^n(E,?)$ which can be psd$\}$.

Nonlinear SDP : min $f_0(\boldsymbol{y},\boldsymbol{X})$ sub. to $\boldsymbol{f}(\boldsymbol{y},\boldsymbol{X}) \in \boldsymbol{\Omega}$, $\boldsymbol{X} \in \mathbb{S}^n_+$

- Values of X_{ij} ($(i,j) \notin E^\bullet$) are not relevant except $\boldsymbol{X} \in \mathbb{S}^n_+$
- G is chordal. C_k ($k = 1,\ldots,\ell$) : the max. cliques of G

$$\Downarrow$$

SDP' : min $f_0(\boldsymbol{y},\boldsymbol{X})$
sub. to $\boldsymbol{f}(\boldsymbol{y},\boldsymbol{X}) \in \boldsymbol{\Omega}$, $\boldsymbol{X}(C_k) \succeq \boldsymbol{O}$ ($k = 1,\ldots,\ell$).

- $\boldsymbol{X}(C_k) \succeq \boldsymbol{O}$ ($k = 1,\ldots,\ell$) are not indep.; $\boldsymbol{X}(C_j) \succeq \boldsymbol{O}$ and $\boldsymbol{X}(C_k) \succeq \boldsymbol{O}$ share common X_{ij} if $i,j \in C_j \cap C_k \neq \emptyset$.
- Further conversion to make them independent.

Nonlinear SDP : min $f_0(\boldsymbol{y}, \boldsymbol{X})$ sub. to $\boldsymbol{f}(\boldsymbol{y}, \boldsymbol{X}) \in \boldsymbol{\Omega}$, $\boldsymbol{X} \in \mathbb{S}^n_+$

$$\Downarrow$$

SDP' : min $f_0(\boldsymbol{y}, \boldsymbol{X})$ sub. to $\boldsymbol{f}(\boldsymbol{y}, \boldsymbol{X}) \in \boldsymbol{\Omega}$ and

$$\begin{pmatrix} X_{11} & X_{12} \\ X_{21} & X_{22} \end{pmatrix}, \begin{pmatrix} X_{22} & X_{23} & X_{24} \\ X_{32} & X_{33} & X_{34} \\ X_{42} & X_{43} & X_{44} \end{pmatrix}, \begin{pmatrix} X_{33} & X_{34} & X_{35} \\ X_{43} & X_{44} & X_{45} \\ X_{53} & X_{54} & X_{55} \end{pmatrix} \succeq \boldsymbol{O}$$

(an SDP with smaller SDP cones and shared variables) $\Longrightarrow$

- Conversion to separate the shared variables — 2 ways

1. d-space conversion method-ae
— Identify common variables by adding equalities

$$\begin{pmatrix} Y^1_{11} & Y^1_{12} \\ Y^1_{21} & Y^1_{22} \end{pmatrix}, \begin{pmatrix} Y^2_{11} & Y^2_{12} & Y^2_{13} \\ Y^2_{21} & Y^2_{22} & Y^2_{23} \\ Y^2_{31} & Y^2_{32} & Y^2_{33} \end{pmatrix}, \begin{pmatrix} Y^3_{11} & Y^3_{12} & Y^3_{13} \\ Y^3_{21} & Y^3_{22} & Y^3_{23} \\ Y^3_{31} & Y^3_{32} & Y^3_{33} \end{pmatrix} \succeq \boldsymbol{O},$$

$$Y^1_{22} = Y^2_{11}, \; Y^2_{22} = Y^3_{11}, \; Y^2_{23} = Y^3_{12}, \; Y^2_{33} = Y^3_{22}.$$

Nonlinear SDP : min $f_0(\boldsymbol{y},\boldsymbol{X})$ sub. to $\boldsymbol{f}(\boldsymbol{y},\boldsymbol{X}) \in \boldsymbol{\Omega}$, $\boldsymbol{X} \in \mathbb{S}^n_+$

$$\Downarrow$$

SDP' : min $f_0(\boldsymbol{y},\boldsymbol{X})$ sub. to $\boldsymbol{f}(\boldsymbol{y},\boldsymbol{X}) \in \boldsymbol{\Omega}$ and

$$\begin{pmatrix} X_{11} & X_{12} \\ X_{21} & X_{22} \end{pmatrix}, \begin{pmatrix} X_{22} & X_{23} & X_{24} \\ X_{32} & X_{33} & X_{34} \\ X_{42} & X_{43} & X_{44} \end{pmatrix}, \begin{pmatrix} X_{33} & X_{34} & X_{35} \\ X_{43} & X_{44} & X_{45} \\ X_{53} & X_{54} & X_{55} \end{pmatrix} \succeq \boldsymbol{O}$$

(an SDP with smaller SDP cones and shared variables) $\Longrightarrow$

2. d-space conversion method-br — represent $\boldsymbol{X}(C_k)$ using basis $\boldsymbol{E}_{ij}$ $((i,j \in C_k, i \leq j)$ of $\mathbb{S}^{C_k}$; $\boldsymbol{X}(C_k) = \sum_{i,j\in C_k, i\leq j} \boldsymbol{E}_{ij} X_{ij}$;

$$\begin{pmatrix} X_{11} & X_{12} \\ X_{21} & X_{22} \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} X_{11} + \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} X_{12} + \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} X_{22}$$

SDP" : min $f_0(\boldsymbol{y},\boldsymbol{X})$ sub. to $\boldsymbol{f}(\boldsymbol{y},\boldsymbol{X}) \in \boldsymbol{\Omega}$ and

$$\sum_{i,j\in C_k, i\leq j} \boldsymbol{E}_{ij} X_{ij} \succeq \boldsymbol{O} \ (k = 1,2,3),$$

where $C_1 = \{1,2\},\ C_2 = \{2,3,4\},\ C_3 = \{3,4,5\}$.

Summary of domain-space conversion

$G(N,E)$: a graph with $N = \{1,\ldots,n\}$, $E \subseteq N \times N$; $(i,i) \notin E$, $(i,j) = (j,i) \in E$, $E^\bullet = E \cup \{(i,i) : i \in N\}$.

$\mathbb{S}^n$ ($\mathbb{S}^n_+$) = the set of $n \times n$ (psd) symmetric matrices.

$\mathbb{S}^n(E,?)$ = the set of $n \times n$ partial symmetric matrices with entries specified in $E^\bullet$.

$\mathbb{S}^n_+(E,?)$ = $\{\boldsymbol{X} \in \mathbb{S}^n(E,?)$ which can be psd$\}$.

Nonlinear SDP : min $f_0(\boldsymbol{y},\boldsymbol{X})$ sub. to $\boldsymbol{f}(\boldsymbol{y},\boldsymbol{X}) \in \boldsymbol{\Omega}$, $\boldsymbol{X} \in \mathbb{S}^n_+$

- Values of X_{ij} ($(i,j) \notin E^\bullet$) are not relevant except $\boldsymbol{X} \in \mathbb{S}^n_+$
- G is chordal. C_k $(k = 1,\ldots,\ell)$: the max. cliques of G

$$\Downarrow$$

SDP' : min f_0 sub. to $\boldsymbol{f} \in \boldsymbol{\Omega}$, $\boldsymbol{X}(C_k) \succeq \boldsymbol{O}$ $(k = 1,\ldots,\ell)$.

- $\boldsymbol{X}(C_k) \succeq \boldsymbol{O}$ $(k = 1,\ldots,\ell)$ are not independent
  conv.method-ae (with additional equalities) $\Leftarrow$ Fukuda et.al
  conv.method-br (with the use of basis representation)
  $\Leftarrow$ Kim et.al for sensor network localization problems

Contents

$G(N,E)$: a chordal graph, $E^{\bullet} = E \cup \{(i,i) : i \in N\}$.

$C_1, \ldots, C_\ell$: the max. cliques of G. $N = \{1, \ldots, n\}$.

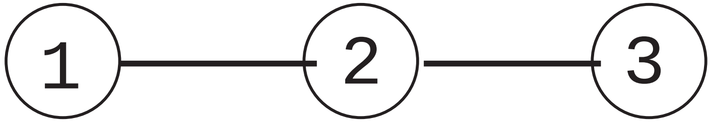
$\mathbb{S}^n(E,0) = \{\boldsymbol{Y} \in \mathbb{S}^n : Y_{ij} = 0 \ (i,j) \notin E^{\bullet}\}$.

$\mathbb{S}^n_+(E,0) = \{\boldsymbol{Y} \in \mathbb{S}^n(E,0) : \boldsymbol{Y} \in \mathbb{S}^n_+\}$.

$\mathbb{S}^C_+ = \{\boldsymbol{Y} \in \mathbb{S}^n_+ : Y_{ij} = 0$ if $(i,j) \notin C \times C\}$ for $\forall C \subseteq N$.

$\mathbb{S}^n(E,?) =$ partial sym. matrices with entries specified in $E^{\bullet}$.

$\mathbb{S}^n_+(E,?) = \{\boldsymbol{X} \in \mathbb{S}^n(E,?)$ which can be psd$\}$.

$G(N,E)$ 

$$\boldsymbol{Y} = \begin{pmatrix} 2 & 2 & 0 \\ 2 & 3 & 1 \\ 0 & 1 & 2 \end{pmatrix} \in \mathbb{S}^3(E,0),\ \boldsymbol{X} = \begin{pmatrix} 3 & 3 & \\ 3 & 3 & 2 \\ & 2 & 2 \end{pmatrix} \in \mathbb{S}^3(E,?)$$

- $\boldsymbol{X} \bullet \boldsymbol{Y} \equiv \sum_{(i,j),(j,i) \in E^{\bullet}} X_{ij} Y_{ij}$ for $\forall\ \boldsymbol{Y} \in \mathbb{S}^n(E,0)$, $\boldsymbol{X} \in \mathbb{S}^n(E,?)$.

- $\boldsymbol{Y} \in \mathbb{S}^n_+(E,0) \Leftrightarrow \boldsymbol{X} \bullet \boldsymbol{Y} \geq 0$ for $\forall \boldsymbol{X} \in \mathbb{S}^n_+$
  $\Leftrightarrow \boldsymbol{X} \bullet \boldsymbol{Y} \geq 0$ for $\forall \boldsymbol{X} \in \mathbb{S}^n_+(E,?)$ $\quad \Rightarrow$ Duality.

$G(N,E)$: a chordal graph, $E^{\bullet} = E \cup \{(i,i) : i \in N\}$.

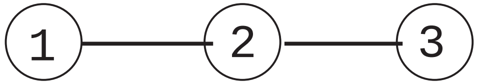
$C_1, \ldots, C_\ell$: the max. cliques of G. $N = \{1, \ldots, n\}$.

$\mathbb{S}^n(E,0) = \{\boldsymbol{Y} \in \mathbb{S}^n : Y_{ij} = 0 \ (i,j) \notin E^{\bullet}\}$.

$\mathbb{S}^n_+(E,0) = \{\boldsymbol{Y} \in \mathbb{S}^n(E,0) : \boldsymbol{Y} \in \mathbb{S}^n_+\}$.

$\mathbb{S}^C_+ = \{\boldsymbol{Y} \in \mathbb{S}^n_+ : Y_{ij} = 0$ if $(i,j) \notin C \times C\}$ for $\forall C \subseteq N$.

Theorem. Let $\boldsymbol{Y} \in \mathbb{S}^n(E,0)$. $\boldsymbol{Y} \in \mathbb{S}^n_+(E,0)$ if and only if $\boldsymbol{Y} = \boldsymbol{Y}^1 + \boldsymbol{Y}^2 + \cdots + \boldsymbol{Y}^\ell$ for some $\boldsymbol{Y}^k \in \mathbb{S}^{C_k}_+ \ (k = 1, \ldots, \ell)$.



$C_1 = \{1,2\}$, $C_2 = \{2,3\}$.

$$\begin{pmatrix} 2 & 2 & 0 \\ 2 & 3 & 1 \\ 0 & 1 & 2 \end{pmatrix} = \begin{pmatrix} 2 & 2 & 0 \\ 2 & 2 & 0 \\ 0 & 0 & 0 \end{pmatrix} + \begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & 1 \\ 0 & 1 & 2 \end{pmatrix}$$

$\in \mathbb{S}^3_+(E,0)$ $\qquad \in \mathbb{S}^{C_1}_+$ $\qquad \in \mathbb{S}^{C_2}_+$

$G(N,E)$: a chordal graph, $E^\bullet = E \cup \{(i,i) : i \in N\}$.

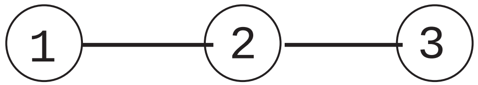
$C_1, \ldots, C_\ell$: the max. cliques of G. $N = \{1,\ldots,n\}$.

$\mathbb{S}^n(E,0) = \{\boldsymbol{Y} \in \mathbb{S}^n : Y_{ij} = 0 \ (i,j) \notin E^\bullet\}$.

$\mathbb{S}^n_+(E,0) = \{\boldsymbol{Y} \in \mathbb{S}^n(E,0) : \boldsymbol{Y} \in \mathbb{S}^n_+\}$.

$\mathbb{S}^C_+ = \{\boldsymbol{Y} \in \mathbb{S}^n_+ : Y_{ij} = 0$ if $(i,j) \notin C \times C\}$ for $\forall C \subseteq N$.

Let $\boldsymbol{M} : \mathbb{R}^s \to \mathbb{S}^n(E,0)$ and $\boldsymbol{u} \in \mathbb{R}^s$. Then $\boldsymbol{M}(\boldsymbol{u}) \in \mathbb{S}^n_+(E,0)$ iff $\boldsymbol{M}(\boldsymbol{u}) = \boldsymbol{Y}^1 + \boldsymbol{Y}^2 + \cdots + \boldsymbol{Y}^\ell$ for some $\boldsymbol{Y}^k \in \mathbb{S}^{C_k}_+$ $(k = 1,\ldots,\ell)$.



$C_1 = \{1,2\}$, $C_2 = \{2,3\}$. $\boldsymbol{M} : \mathbb{R}^m \to \mathbb{S}^3(E,0)$.

$$\boldsymbol{M}(\boldsymbol{u}) \in \mathbb{S}^3_+(E,0) \qquad \boldsymbol{M}(\boldsymbol{u}) = \begin{pmatrix} Y^1_{11} & Y^1_{12} & 0 \\ Y^1_{12} & Y^1_{22} & 0 \\ 0 & 0 & 0 \end{pmatrix} + \begin{pmatrix} 0 & 0 & 0 \\ 0 & Y^2_{22} & Y^2_{23} \\ 0 & Y^2_{32} & Y^2_{33} \end{pmatrix}$$

$$\Updownarrow$$

$$\left.\begin{array}{l} M_{11} = Y^1_{11}, M_{12} = Y^1_{12}, \\ M_{22} = Y^1_{22} + Y^2_{22}, \\ M_{23} = Y^2_{23}, M_{33} = Y^2_{33}, \\ \square \succeq \boldsymbol{O}, \ \square \succeq \boldsymbol{O} \end{array}\right\} \Leftrightarrow \left\{\begin{array}{l} \begin{pmatrix} M_{11}(\boldsymbol{u}) & M_{12}(\boldsymbol{u}) \\ M_{21}(\boldsymbol{u}) & Y^1_{22} \end{pmatrix} \succeq \boldsymbol{O}, \\ \begin{pmatrix} M_{22}(\boldsymbol{u}) - Y^1_{22} & M_{23}(\boldsymbol{u}) \\ M_{32}(\boldsymbol{u}) & M_{33}(\boldsymbol{u}) \end{pmatrix} \succeq \boldsymbol{O} \end{array}\right.$$

$G(N,E)$: a chordal graph, $E^{\bullet} = E \cup \{(i,i) : i \in N\}$.

$C_1,\ldots,C_\ell$: the max. cliques of G. $N = \{1,\ldots,n\}$.

$\mathbb{S}^n(E,0) = \{\boldsymbol{Y} \in \mathbb{S}^n : Y_{ij} = 0 \ (i,j) \notin E^{\bullet}\}$.

$\mathbb{S}^n_+(E,0) = \{\boldsymbol{Y} \in \mathbb{S}^n(E,0) : \boldsymbol{Y} \in \mathbb{S}^n_+\}$.

$\mathbb{S}^C_+ = \{\boldsymbol{Y} \in \mathbb{S}^n_+ : Y_{ij} = 0$ if $(i,j) \notin C \times C\}$ for $\forall C \subseteq N$.

Let $\boldsymbol{M} : \mathbb{R}^s \to \mathbb{S}^n(E,0)$ and $\boldsymbol{u} \in \mathbb{R}^s$. Then $\boldsymbol{M}(\boldsymbol{u}) \in \mathbb{S}^n_+(E,0)$ iff $\boldsymbol{M}(\boldsymbol{u}) = \boldsymbol{Y}^1 + \boldsymbol{Y}^2 + \cdots + \boldsymbol{Y}^\ell$ for some $\boldsymbol{Y}^k \in \mathbb{S}^{C_k}_+ \ (k = 1,\ldots,\ell)$.

- r-space conv. method-amv (with the use of auxiliary matrix variables) — the left method in the previous slide
- r-space conv. method-arv (with the use of auxiliary real variables) — the right method in the previous slide

Summary of d-space conversion and r-space conversion

- sparsity characterized by a chordal graph structure.

$G(N,E)$: a chordal graph behind the sparse structure
C_k : $(k=1,2,\ldots,\ell)$ be the max cliques.

d-space conversion: $\boldsymbol{X} \in \mathbb{S}^n_+ \Rightarrow \boldsymbol{X}(C_k) \in \mathbb{S}^{C_k}_+ \ (k=1,2,\ldots,\ell)$.
To make $\boldsymbol{X}(C_k) \in \mathbb{S}^{C_k}_+ \ (k=1,2,\ldots,\ell)$ independent, 2 methods
(d-ae) d-dpace conv. method-ae (additional equalities)
(d-br) d-space conv. method-br (basis representation)

r-space conversion: $\boldsymbol{M}(\boldsymbol{u}) \in \mathbb{S}^n_+(E,0) \Leftrightarrow$
$\boldsymbol{M}(\boldsymbol{u}) = \boldsymbol{Y}^1 + \boldsymbol{Y}^2 + \cdots + \boldsymbol{Y}^\ell$ for some $\boldsymbol{Y}^k \in \mathbb{S}^{C_k}_+ \ (k=1,\ldots,\ell)$.
(r-amv) r-space conv. method-amv (auxiliary mat. variables)
(r-arv) r-space conv. method-arv (auxiliary real variables)

- (d-ae) & (r-arv) ((d-br) & (r-amv)) are dual to each other.

P: min $\boldsymbol{M}(\boldsymbol{u}) \bullet \boldsymbol{X}$ sub. to $\boldsymbol{X} \succeq \boldsymbol{O}$
D: max 0 sub. to $\boldsymbol{M}(\boldsymbol{u}) \succeq \boldsymbol{O}$

Summary of d-space conversion and r-space conversion

- sparsity characterized by a chordal graph structure.

$G(N, E)$: a chordal graph behind the sparse structure
C_k : $(k = 1, 2, \ldots, \ell)$ be the max cliques.

d-space conversion: $\boldsymbol{X} \in \mathbb{S}^n_+ \Rightarrow \boldsymbol{X}(C_k) \in \mathbb{S}^{C_k}_+ \ (k = 1, 2, \ldots, \ell)$.
To make $\boldsymbol{X}(C_k) \in \mathbb{S}^{C_k}_+ \ (k = 1, 2, \ldots, \ell)$ independent, 2 methods
(d-ae) d-dpace conv. method-ae (additional equalities)
(d-br) d-space conv. method-br (basis representation)

r-space conversion: $\boldsymbol{M}(\boldsymbol{u}) \in \mathbb{S}^n_+(E, 0) \Leftrightarrow$
$\boldsymbol{M}(\boldsymbol{u}) = \boldsymbol{Y}^1 + \boldsymbol{Y}^2 + \cdots + \boldsymbol{Y}^\ell$ for some $\boldsymbol{Y}^k \in \mathbb{S}^{C_k}_+ \ (k = 1, \ldots, \ell)$.
(r-amv) r-space conv. method-amv (auxiliary mat. variables)
(r-arv) r-space conv. method-arv (auxiliary real variables)

- Efficient implementation and effective combination of the four methods $\Rightarrow$ future study
- Preliminary numerical results — later

Contents

Quadratic SDP ($\boldsymbol{u} \in \mathbb{R}^s$, $\boldsymbol{M}(\boldsymbol{u}) \in \mathbb{S}^n$ for every $\boldsymbol{u} \in \mathbb{R}^s$) :

min $\sum_{i=1}^s c_i u_i$ s.t. $\boldsymbol{M}(\boldsymbol{u}) \succeq \boldsymbol{O}, \ M_{ij}(\boldsymbol{u}) = \boldsymbol{Q}_{ij} \bullet \begin{pmatrix} 1 & \boldsymbol{u}^T \\ \boldsymbol{u} & \boldsymbol{u}\boldsymbol{u}^T \end{pmatrix}$.

SDP relaxation : ••• $\boldsymbol{W} \equiv \begin{pmatrix} 1 & \boldsymbol{u}^T \\ \boldsymbol{u} & \boldsymbol{X} \end{pmatrix} \succeq \boldsymbol{O}, \ \widehat{\boldsymbol{M}}(\boldsymbol{W}) \succeq \boldsymbol{O}$,
where $\widehat{M}_{ij}(\boldsymbol{W}) = \boldsymbol{Q}_{ij} \bullet \boldsymbol{W}$.

Some numerical results

1. Tridiagonal quadratic SDP (randomly generated)
2. Bordered block-diagnal quadratic SDP (randomly —)
3. Eigenvalue optimization of Structures (Kanno-Ohsaki ’07)

Comparison of 4 kinds of SDPs

(a) Original problem without exploiting any sparsity
(b) Exploiting doamin-sparsity
(c) Exploiting range-sparsity (d) (b) + (c)

All SDP problems were solved by SeDuMi on 2.66GHz Dual-Core Intel Xeon with 12GB Memory.

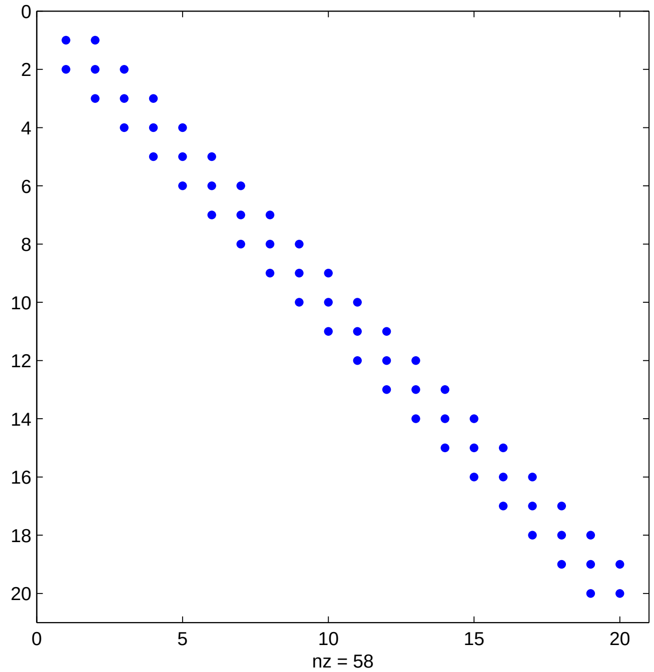
Quadratic SDP ($\boldsymbol{u} \in \mathbb{R}^s$, $\boldsymbol{M}(\boldsymbol{u}) \in \mathbb{S}^n$ for every $\boldsymbol{u} \in \mathbb{R}^s$) :

min $\sum_{i=1}^{s} c_i u_i$ s.t. $\boldsymbol{M}(\boldsymbol{u}) \succeq \boldsymbol{O}, \; M_{ij}(\boldsymbol{u}) = \boldsymbol{Q}_{ij} \bullet \begin{pmatrix} 1 & \boldsymbol{u}^T \\ \boldsymbol{u} & \boldsymbol{u}\boldsymbol{u}^T \end{pmatrix}.$

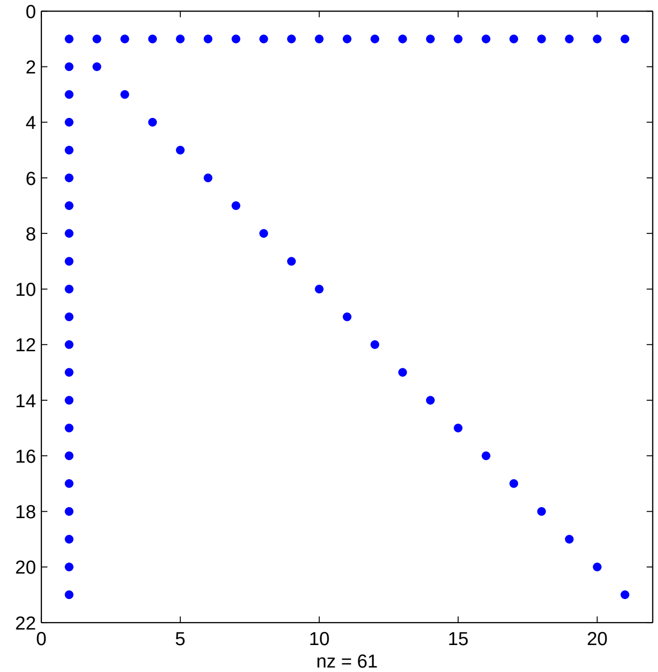
SDP relaxation : • • • $\boldsymbol{W} \equiv \begin{pmatrix} 1 & \boldsymbol{u}^T \\ \boldsymbol{u} & \boldsymbol{X} \end{pmatrix} \succeq \boldsymbol{O}, \; \widehat{\boldsymbol{M}}(\boldsymbol{W}) \succeq \boldsymbol{O}$,
where $\widehat{M}_{ij}(\boldsymbol{W}) = \boldsymbol{Q}_{ij} \bullet \boldsymbol{W}$.

Tridiagonal quadratic SDP

r-sparsity



d-sparsity



Quadratic SDP ($\boldsymbol{u} \in \mathbb{R}^s$, $\boldsymbol{M}(\boldsymbol{u}) \in \mathbb{S}^n$ for every $\boldsymbol{u} \in \mathbb{R}^s$) :

min $\sum_{i=1}^s c_i u_i$ s.t. $\boldsymbol{M}(\boldsymbol{u}) \succeq \boldsymbol{O},\ M_{ij}(\boldsymbol{u}) = \boldsymbol{Q}_{ij} \bullet \begin{pmatrix} 1 & \boldsymbol{u}^T \\ \boldsymbol{u} & \boldsymbol{u}\boldsymbol{u}^T \end{pmatrix}$.

SDP relaxation : ● ● ● $\boldsymbol{W} \equiv \begin{pmatrix} 1 & \boldsymbol{u}^T \\ \boldsymbol{u} & \boldsymbol{X} \end{pmatrix} \succeq \boldsymbol{O},\ \widehat{\boldsymbol{M}}(\boldsymbol{W}) \succeq \boldsymbol{O}$,
where $\widehat{M}_{ij}(\boldsymbol{W}) = \boldsymbol{Q}_{ij} \bullet \boldsymbol{W}$.

Tridiagonal quadratic SDP

		cpu time (Schur comp.size, max.mat.var.size)		
s	n	no sp.	d-br	d-br, r-arv
40	40	8.38 (860, 41)	0.97 (80, 40)	0.68 (118, 2)
80	80	384.43 (3320, 81)	11.72 (160, 80)	1.58 (238, 2)
320	320		100.36 (640, 320)	24.57 (958, 2)

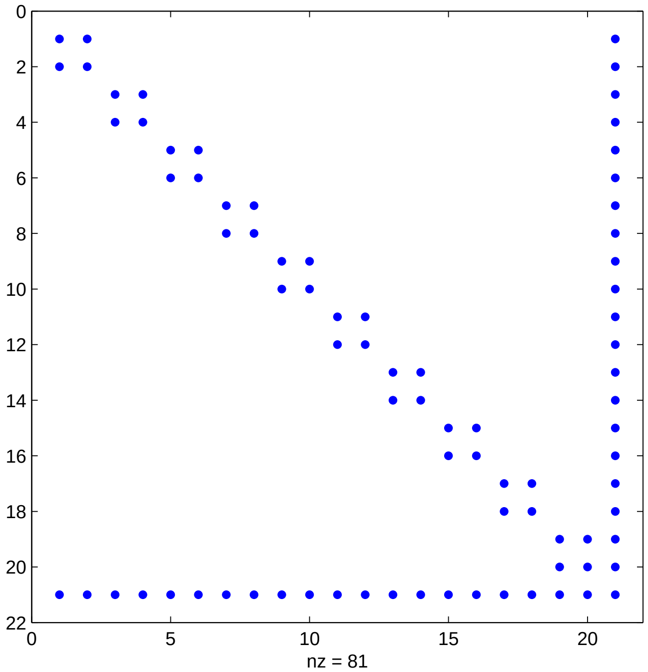
Quadratic SDP ($\boldsymbol{u} \in \mathbb{R}^s$, $\boldsymbol{M}(\boldsymbol{u}) \in \mathbb{S}^n$ for every $\boldsymbol{u} \in \mathbb{R}^s$) :

$$\text{min } \sum_{i=1}^{s} c_i u_i \text{ s.t. } \boldsymbol{M}(\boldsymbol{u}) \succeq \boldsymbol{O}, \; M_{ij}(\boldsymbol{u}) = \boldsymbol{Q}_{ij} \bullet \begin{pmatrix} 1 & \boldsymbol{u}^T \\ \boldsymbol{u} & \boldsymbol{u}\boldsymbol{u}^T \end{pmatrix}.$$

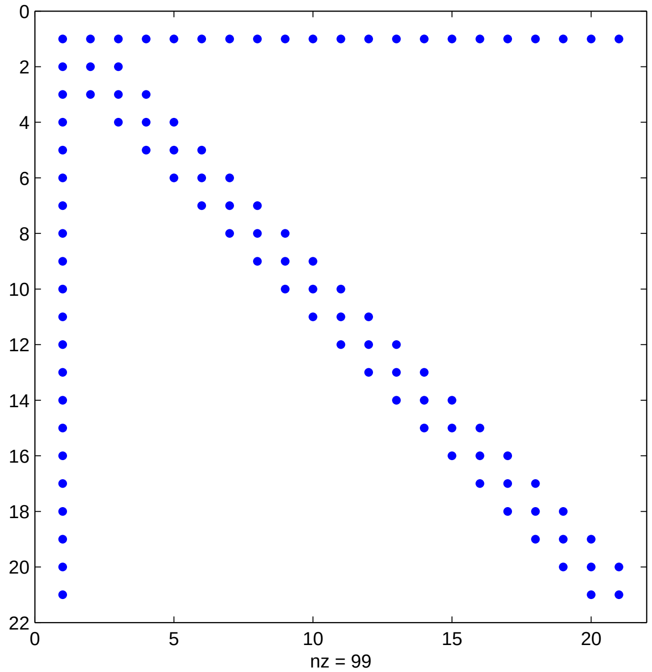
SDP relaxation : $\bullet\bullet\bullet$ $\boldsymbol{W} \equiv \begin{pmatrix} 1 & \boldsymbol{u}^T \\ \boldsymbol{u} & \boldsymbol{X} \end{pmatrix} \succeq \boldsymbol{O}, \; \widehat{\boldsymbol{M}}(\boldsymbol{W}) \succeq \boldsymbol{O}$,
where $\widehat{M}_{ij}(\boldsymbol{W}) = \boldsymbol{Q}_{ij} \bullet \boldsymbol{W}$.

Bordered block-diagonal quadratic SDP

r-sparsity



d-sparsity



Quadratic SDP ($\boldsymbol{u} \in \mathbb{R}^s$, $\boldsymbol{M}(\boldsymbol{u}) \in \mathbb{S}^n$ for every $\boldsymbol{u} \in \mathbb{R}^s$) :

min $\sum_{i=1}^{s} c_i u_i$ s.t. $\boldsymbol{M}(\boldsymbol{u}) \succeq \boldsymbol{O}, \; M_{ij}(\boldsymbol{u}) = \boldsymbol{Q}_{ij} \bullet \begin{pmatrix} 1 & \boldsymbol{u}^T \\ \boldsymbol{u} & \boldsymbol{u}\boldsymbol{u}^T \end{pmatrix}.$

SDP relaxation : • • • $\boldsymbol{W} \equiv \begin{pmatrix} 1 & \boldsymbol{u}^T \\ \boldsymbol{u} & \boldsymbol{X} \end{pmatrix} \succeq \boldsymbol{O}, \; \widehat{\boldsymbol{M}}(\boldsymbol{W}) \succeq \boldsymbol{O}$,
where $\widehat{M}_{ij}(\boldsymbol{W}) = \boldsymbol{Q}_{ij} \bullet \boldsymbol{W}$.

Bordered block-diagonal quadratic SDP

		cpu time (Schur comp.size, max.mat.var.size)			
s	n	no sp.	d-br	r-arv	d-br, r-arv
40	81	30.41 (860, 81)	12.26 (119, 81)	12.28 (899, 41)	0.94 (158,3)
40	161	38.71 (860, 161)	27.63 (119, 161)	9.22 (939, 41)	1.45 (198, 3)
40	641	591.10 (860, 641)	551.37 (119, 641)	24.10 (1179, 41)	8.13 (438, 3)

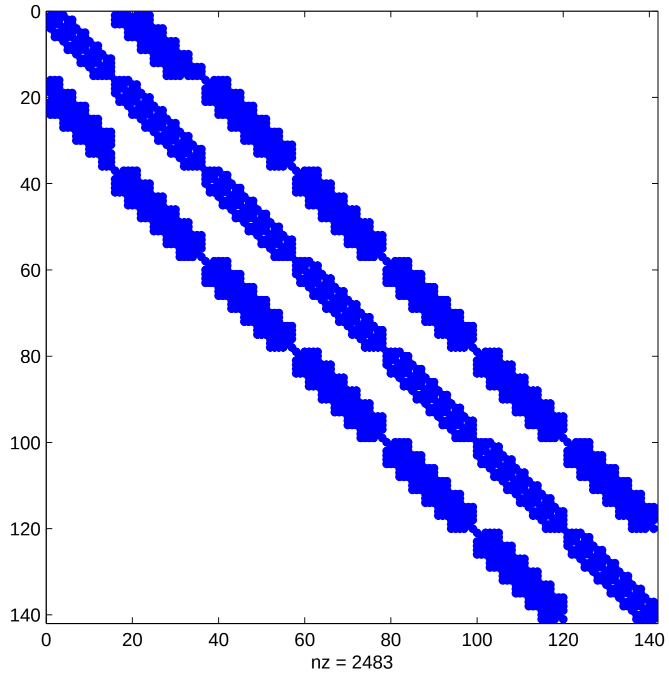
Quadratic SDP ($\boldsymbol{u} \in \mathbb{R}^s$, $\boldsymbol{M}(\boldsymbol{u}) \in \mathbb{S}^n$ for every $\boldsymbol{u} \in \mathbb{R}^s$) :

min $\sum_{i=1}^{s} c_i u_i$ s.t. $\boldsymbol{M}(\boldsymbol{u}) \succeq \boldsymbol{O}, \; M_{ij}(\boldsymbol{u}) = \boldsymbol{Q}_{ij} \bullet \begin{pmatrix} 1 & \boldsymbol{u}^T \\ \boldsymbol{u} & \boldsymbol{u}\boldsymbol{u}^T \end{pmatrix}.$

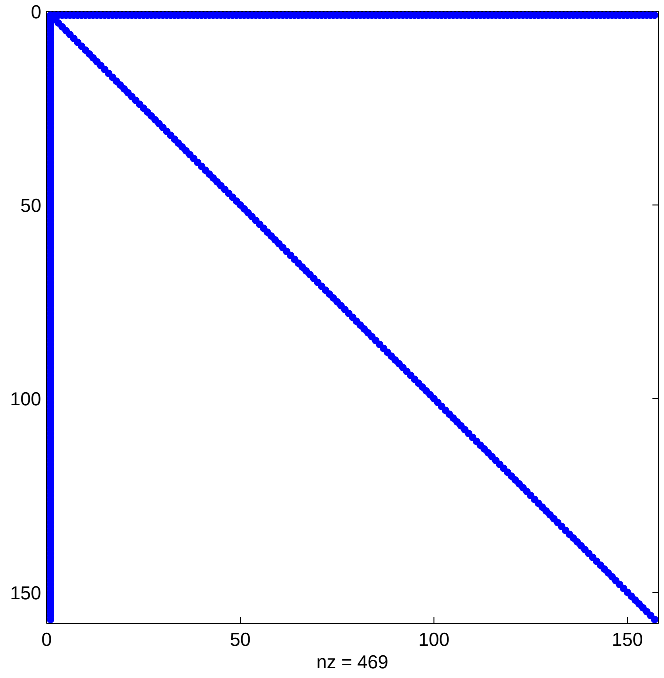
SDP relaxation : • • • $\boldsymbol{W} \equiv \begin{pmatrix} 1 & \boldsymbol{u}^T \\ \boldsymbol{u} & \boldsymbol{X} \end{pmatrix} \succeq \boldsymbol{O}, \; \widehat{\boldsymbol{M}}(\boldsymbol{W}) \succeq \boldsymbol{O},$
where $\widehat{M}_{ij}(\boldsymbol{W}) = \boldsymbol{Q}_{ij} \bullet \boldsymbol{W}$.

Quadratic SDP from eigenvalue optimization of structures

r-sparsity



d-sparsity



Quadratic SDP ($\boldsymbol{u} \in \mathbb{R}^s$, $\boldsymbol{M}(\boldsymbol{u}) \in \mathbb{S}^n$ for every $\boldsymbol{u} \in \mathbb{R}^s$) :

min $\sum_{i=1}^{s} c_i u_i$ s.t. $\boldsymbol{M}(\boldsymbol{u}) \succeq \boldsymbol{O}, \; M_{ij}(\boldsymbol{u}) = \boldsymbol{Q}_{ij} \bullet \begin{pmatrix} 1 & \boldsymbol{u}^T \\ \boldsymbol{u} & \boldsymbol{u}\boldsymbol{u}^T \end{pmatrix}.$

SDP relaxation : • • • $\boldsymbol{W} \equiv \begin{pmatrix} 1 & \boldsymbol{u}^T \\ \boldsymbol{u} & \boldsymbol{X} \end{pmatrix} \succeq \boldsymbol{O}, \; \widehat{\boldsymbol{M}}(\boldsymbol{W}) \succeq \boldsymbol{O},$
where $\widehat{M}_{ij}(\boldsymbol{W}) = \boldsymbol{Q}_{ij} \bullet \boldsymbol{W}$.

Quadratic SDP from eigenvalue optimization of structures

		cpu time in second			
s	n	no sp.	d-br	r-amv	r-amv, d-ae
42	42	7.21	1.09	3.98	3.08
72	69	271.79	6.68	23.53	6.29
156	141		29.50	112.04	43.57
272	237		92.83	861.88	354.74

- Overheads in domain- and range-space conv. methods; adding equalities, real variables and/or matrix variables
- More efficient implementation? How do we combine them?

Contents

Sensor network localization problem: Let $s = 2$ or 3.

$$\boldsymbol{x}^p \in \mathbb{R}^s \quad : \quad \text{unknown location of sensors } (p = 1, 2, \ldots, m),$$

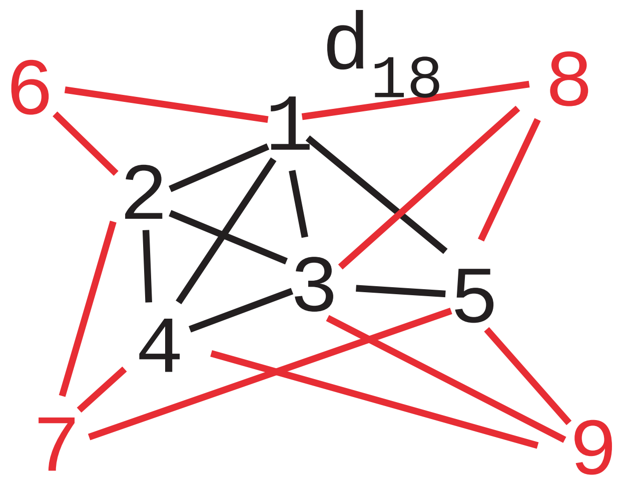
$$\boldsymbol{x}^r = \boldsymbol{a}^r \in \mathbb{R}^s \quad : \quad \text{known location of anchors } (r = m+1, \ldots, n),$$

$$d_{pq} = \|\boldsymbol{x}^p - \boldsymbol{x}^q\| + \epsilon_{pq} \text{ — given for } (p, q) \in \mathcal{N},$$

$$\mathcal{N} = \{(p, q) : \|\boldsymbol{x}^p - \boldsymbol{x}^q\| \leq \rho = \text{a given radio range}\}$$

Here ϵ_{pq} denotes a noise.

$m = 5, \ n = 9$.
$1, \ldots, 5$: sensors
$6, 7, 8, 9$: anchors



Anchors' positions are known.
A distance is given for $\forall$ edge.
Compute locations of sensors.

$\Rightarrow$ Some nonconvex QOPs

- SDP relaxation — FSDP by Biswas-Ye '06, ESDP by Wang et al '07, ... for $s = 2$.
- SOCP relaxation — Tseng '07 for $s = 2$.
- ...

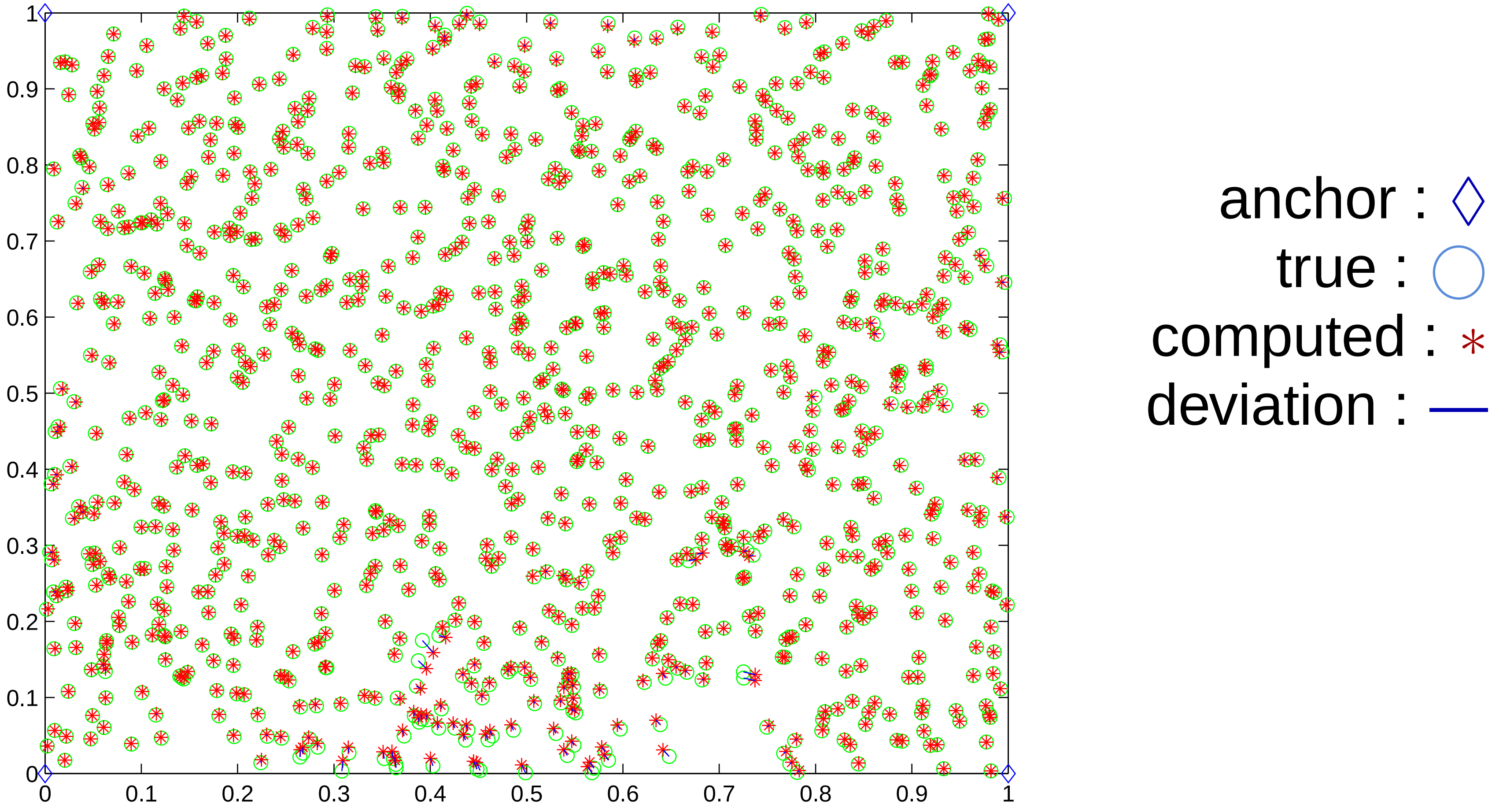
Numerical results on 3 methods (a), (b) and (c) applied to a sensor network localization problem with

$m =$ the number of sensors dist. randomly in $[0,1]^2$,
4 anchors located at the corner of $[0,1]^2$,
$\rho =$ radio distance $= 0.1,$ no noise.

(a) FSDP (Biswas-Ye '06)
(b) FSDP + d-br as strong as (a)
(c) FSDP + d-ae as strong as (a)

	cpu time for solving SDP by SeDuMi in second		
m	(a) FSDP	(b) FSDP + d-br	(c) FSDP + d-ae
500	389.1	35.0	69.5
1000	3345.2	60.4	178.8
2000		111.1	326.0
4000		182.1	761.0

(b) FSDP + d-br — cpu time 60.4 sec
(c) FSDP + d-ae — cpu time 178.8 sec

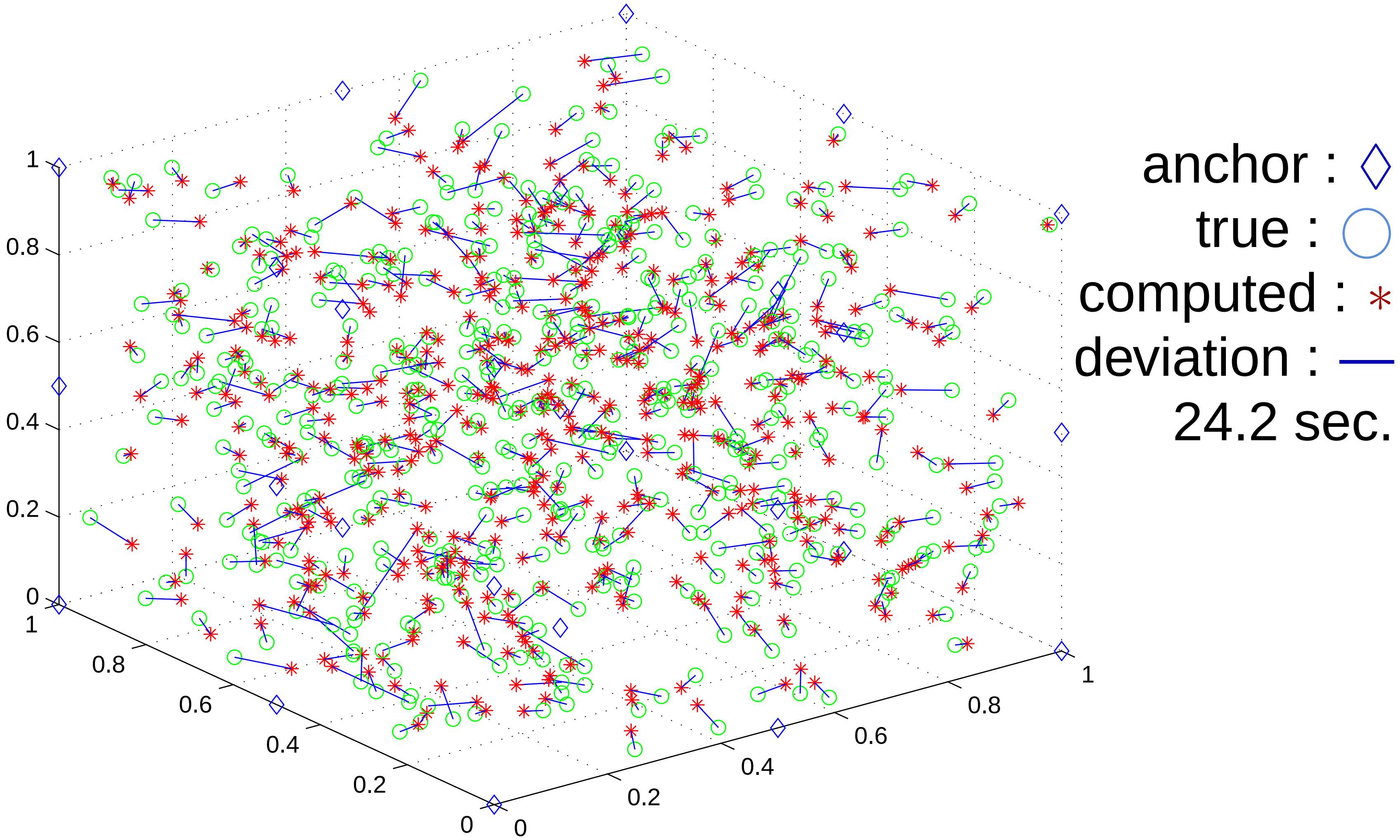


3 dim, 500 sensors, 27 anchors, r.range = 0.3, noise $\leftarrow$ N(0,0.1);

(estimated dist.) $\hat{d}_{pq} = (1 + \epsilon_{pq})d_{pq}$ (true unknown dist.),

$\epsilon_{pq} \leftarrow N(0, 0.1)$

(b) FSDP + d-br

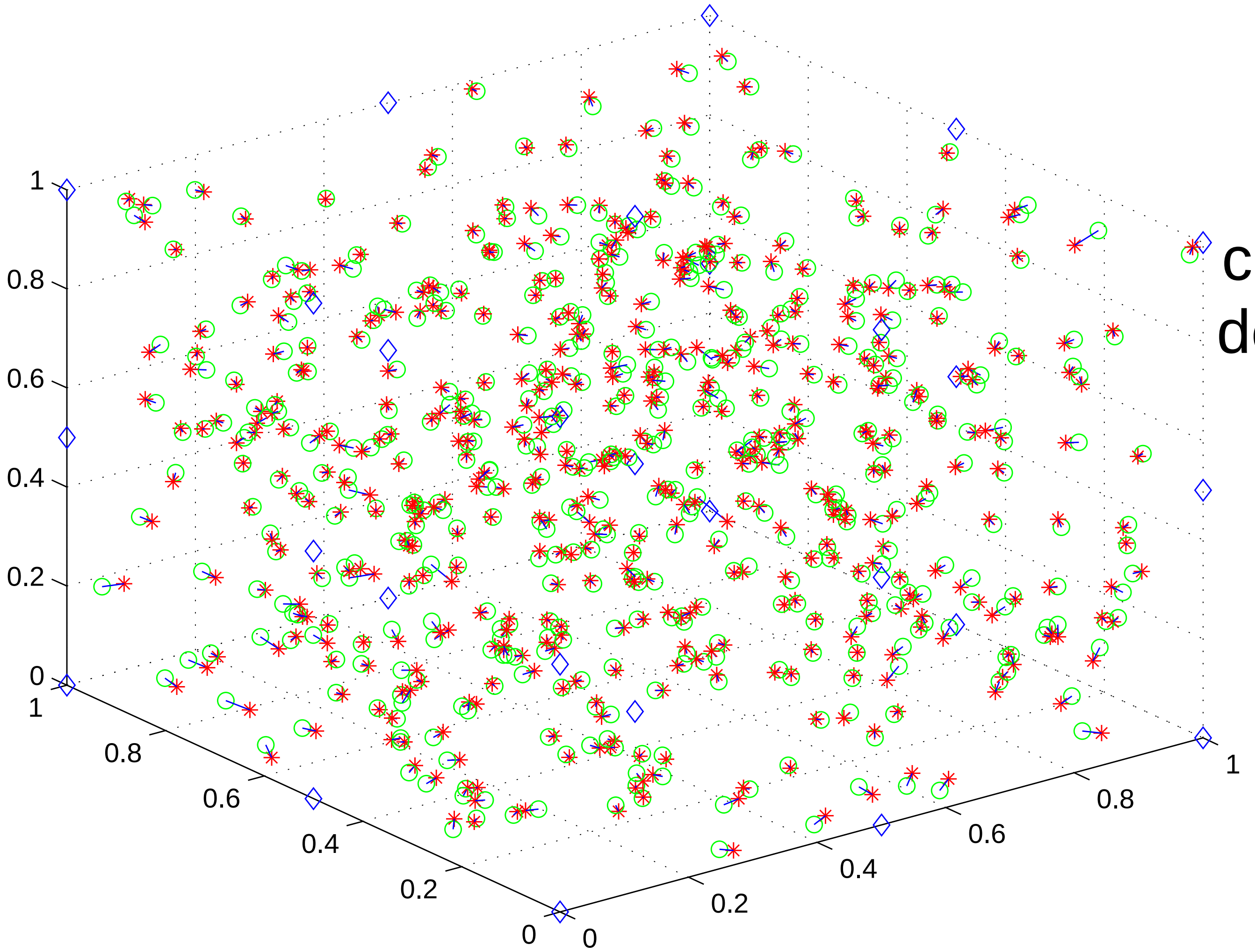


3 dim, 500 sensors, 27 anchors, r.range = 0.3, noise $\leftarrow$ N(0,0.1);

(estimated dist.) $\hat{d}_{pq} = (1 + \epsilon_{pq})d_{pq}$ (true unknown dist.),

$\epsilon_{pq} \leftarrow N(0, 0.1)$

(b) FSDP + d-br + Gradient method



anchor : ◊
true : ○
computed : *
deviation : —
24.2 sec.
+ 8.4 sec.

Contents

Exploiting sparsity characterized by a chordal graph structure in polynomial SDPs via psd matrix completion

Polynomial SDP

each large matrix variable

↓ exploiting **domain-sparsity**

multiple smaller matrix variables

each large matrix inequality

↓ exploiting **range-sparsity**

multiple smaller matrix inequalities

↓ **sparse** SDP relaxation (Lasserre, **Kojima et. al**)

Linear SDP with multiple smaller matrix variables and multiple smaller matrix inequalities satisfying **correlative sparsity**

‖

sparsity of the Schur complement matrix *

- Overheads in domain- and range-space conv. methods; adding equalities, real variables and/or matrix variables
- More efficient implementation? How do we combine them?