Exploiting Structure in SDPs with Chordal Sparsity

Antonis Papachristodoulou

Department of Engineering Science, University of Oxford

Joint work with Yang Zheng, Giovanni Fantuzzi, Paul Goulart and Andrew Wynn



CDC 2016 Pre-conference Workshop



OUTLINE



2 ADMM for Primal and Dual Sparse SDPs

3 CDCS: Cone Decomposition Conic Solver



Chordal Graphs

A graph $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ is chordal if every cycle of length \geq 4 has a chord.



Can recognise chordal graphs in $O(|\mathcal{V}| + |\mathcal{E}|)$ time

Chordal Graphs

A graph $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ is chordal if every cycle of length \geq 4 has a chord.



Can recognise chordal graphs in $O(|\mathcal{V}| + |\mathcal{E}|)$ time

Maximal Cliques

A maximal clique is a clique that is not a subset of another clique. e.g. $C_1 = \{1, 2, 6\}$.

Can find the maximal cliques of a chordal graph in $O(|\mathcal{V}|+|\mathcal{E}|)$ time.



Examples of Chordal Graphs



$\mathbb{S}^n_{\scriptscriptstyle +}(\mathcal{E},0) = \{ Z \in \mathbb{S}^n(\mathcal{E},0) \, | \, Z \succeq 0 \}$

$Z \in \mathbb{S}^{n}(\mathcal{E}, 0) = \left\{ Z \in \mathbb{S}^{n} \mid Z_{ij} = 0, \forall (i, j) \notin \mathcal{E} \right\}$



Agler's Theorem

Theorem: Let $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ be a chordal graph with set of maximal cliques $\mathcal{C} = \{\mathcal{C}_1, \mathcal{C}_2, ..., \mathcal{C}_p\}$. Suppose that $Z \in \mathbb{S}^n(\mathcal{E}, 0)$. Then $Z \in \mathbb{S}^n_+(\mathcal{E}, 0)$ if and only if there exists a set of matrices $\{Z^1, Z^2, ..., Z^p\}$ such that $Z = \sum_{k=1}^p Z^k$, $Z^k \succeq 0$, $Z_{ij}^k = 0$ if $(i, j) \notin \mathcal{E}_k \times \mathcal{E}_k$.









$Z \succeq 0 \iff \exists A, B, C, D$ such that $A + B + C + D = Z, A, B, C, D \succeq 0$



$Z \succeq 0 \iff \exists A, B, C, D$ such that $A + B + C + D = Z, A, B, C, D \succeq 0$



$Z \succeq 0 \iff \exists A, B, C, D \text{ such that}$ $A + B + C + D = Z, A, B, C, D \succeq 0$



$Z \succeq 0 \iff \exists A, B, C, D$ such that $A + B + C + D = Z, A, B, C, D \succeq 0$



SDPs with Chordal Sparsity



- Applications: control theory, fluid mechanics, machine learning, polynomial optimization, combinatorics, operations research, finance, etc.
- Second-order solvers : SeDuMi, SDPA, SDPT3
- Large-scale cases: exploit the inherent structure of the instances (De Klerk, 2010):
 - Low Rank and Algebraic Symmetry.
 - Chordal Sparsity:
 - ✓ Second-order methods: Fukuda et al., 2001; Nakata et al., 2003; Andersen et al., 2010;
 - ✓ First-order methods: Madani *et al.*, 2015; Sun *et al.*, 2014.

Decomposing Sparse SDPs



SDPs with Chordal Sparsity

$$\mathbb{S}^{n}(\mathcal{E},0) = \left\{ X \in \mathbb{S}^{n} \mid X_{ij} = 0, \forall (i,j) \notin \mathcal{E} \right\}$$
$$\mathbb{S}^{n}_{+}(\mathcal{E},0) = \left\{ X \in \mathbb{S}^{n}(\mathcal{E},0) \mid X \succeq 0 \right\}$$



 $\mathbb{S}^{n}(\mathcal{E},?) = n \times n \text{ partial symmetric matrices with entries defined on } \mathcal{E}$ $\mathbb{S}^{n}_{+}(\mathcal{E},?) = \{X \in \mathbb{S}^{n}(\mathcal{E},?) \mid \exists M \ge 0, M_{ij} = X_{ij}, \forall (i,j) \in \mathcal{E}\}$

 $\mathbb{S}^{n}_{+}(\mathcal{E},?)$ and $\mathbb{S}^{n}_{+}(\mathcal{E},0)$ are dual cones of each other

Grone's theorem

Consider a choral graph $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ with a set of maximal cliques $\mathcal{C}_1, \ldots, \mathcal{C}_p$.

Grone's Theorem:

 $X \in \mathbb{S}^{n}_{+}(\mathcal{E},?)$ if and only if $X(\mathcal{C}_{k}) \succeq 0, k = 1,...,p$.



Cone Decomposition of Primal and Dual SDPs





OUTLINE



2 ADMM for Primal and Dual Sparse SDPs

3 CDCS: Cone Decomposition Conic Solver



An operator splitting method for a problem of the form

 $\min f(x) + g(z)$

s.t.
$$x = z$$

f, g may be nonsmooth.

Lagrangian:

Lagrangian:

$$\mathcal{L} = f(x) + g(z) + \frac{\rho}{2} \left\| x - z + \frac{1}{\rho} \lambda \right\|_{2}^{2}$$
ADMM:

$$x^{n+1} = \operatorname{argmin}_{x} \left(f(x) + \frac{\rho}{2} \left\| x - z^{n} + \frac{1}{\rho} \lambda^{n} \right\|_{2}^{2} \right)$$

$$z^{n+1} = \operatorname{argmin}_{z} \left(g(z) + \frac{\rho}{2} \left\| x^{n+1} - z + \frac{1}{\rho} \lambda^{n} \right\|_{2}^{2} \right)$$

$$\lambda^{n+1} = \lambda^{n} + \rho(x^{n+1} - z^{n+1})$$

Reformulation and decomposition of the PSD constraint

$$\min_{X} \langle C, X \rangle$$

s.t. $\mathcal{A}(X) = b$
 $X_{k} = E_{C_{k}} X E_{C_{k}}^{T}, k = 1, ..., p$
 $X_{k} \in \mathbb{S}_{+}^{|C_{k}|}, k = 1, ..., p$

$$\min_{\substack{x, x_1, \dots, x_p}} c^{\mathsf{T}} x$$

s.t. $Ax = b$
 $x_k = H_k x, \ k = 1, \dots, p$
 $x_k \in \mathcal{S}_k, \ k = 1, \dots, p$

Using indicator functions

$$\min_{x,x_1,\dots,x_p} c^{\mathsf{T}} x + \delta_0 (Ax - b) + \sum_{k=1}^p \delta_{s_k} (x_k)$$

s.t. $x_k = H_k x, \ k = 1,\dots,p$

Augmented Lagrangian

$$\mathcal{L} = c^{T} x + \delta_{0} (Ax - b) + \sum_{k=1}^{p} \left[\delta_{S_{k}} (x_{k}) + \frac{\rho}{2} \left\| x_{k} - H_{k} x + \frac{1}{\rho} \lambda_{k} \right\|^{2} \right]$$

• Regroup the variables

$$\mathcal{X} \triangleq \{\mathbf{x}\}; \quad \mathcal{Y} \triangleq \{\mathbf{x}_1, \dots, \mathbf{x}_p\}; \quad \mathcal{Z} \triangleq \{\lambda_1, \dots, \lambda_p\}$$

ADMM for Primal SDPs

$$\mathcal{L} = \mathbf{c}^{\mathsf{T}} \mathbf{x} + \delta_0 (\mathbf{A}\mathbf{x} - \mathbf{b}) + \sum_{k=1}^{p} \left[\delta_{s_k} (\mathbf{x}_k) + \frac{\rho}{2} \left\| \mathbf{x}_k - \mathbf{H}_k \mathbf{x} + \frac{1}{\rho} \lambda_k \right\|^2 \right]$$
$$\mathcal{X} \triangleq \{\mathbf{x}\}; \quad \mathcal{Y} \triangleq \{\mathbf{x}_1, \dots, \mathbf{x}_p\}; \quad \mathcal{Z} \triangleq \{\lambda_1, \dots, \lambda_p\}$$

1) Minimization over ${\cal X}$

$$\min_{x} c^{T} x + \frac{\rho}{2} \sum_{k=1}^{p} \left\| x_{k}^{(n)} - H_{k} x + \frac{1}{\rho} \lambda_{k}^{(n)} \right\|^{2}$$

s.t.
$$Ax = b$$

2) Minimization over \mathcal{Y} $\min_{x_{k}} \left\| x_{k} - H_{k} x^{(n+1)} + \frac{1}{\rho} \lambda_{k}^{(n)} \right\|^{2}$ s.t. $x_{k} \in \mathcal{S}_{k}$ 3) Update multipliers $\lambda_{k}^{(n+1)} = \lambda_{k}^{(n)} + \rho \left(x_{k}^{(n+1)} - H_{k} x^{(n+1)} \right)$ QP with linear constraint

 ✓ the KKT system matrix only depends on the problem data.

Projections in parallel

ADMM for Primal and Dual SDPs



The duality between the primal and dual SDP is inherited by the decomposed problems by virtue of the duality between Grone's and Agler's theorems.

ADMM for the Homogeneous self-dual embedding

$$\begin{split} \min_{x,x_{k}} & c^{T}x & \min_{y,z_{k}} & -b^{T}y \\ \text{s.t.} & Ax = b & \\ & x_{k} = H_{k}x & \text{s.t.} & A^{T}y + \sum_{k=1}^{p} H_{k}^{T}v_{k} = c \\ & x_{k} \in \mathcal{S}_{k}, \ k = 1, \dots, p & z_{k} \in \mathcal{S}_{k}, \ k = 1, \dots, p \\ \end{split}$$
 $\bullet \text{ Notational simplicity} & s \triangleq \begin{bmatrix} x_{1} \\ \vdots \\ x_{p} \end{bmatrix}, \ z \triangleq \begin{bmatrix} z_{1} \\ \vdots \\ z_{p} \end{bmatrix}, \ v \triangleq \begin{bmatrix} v_{1} \\ \vdots \\ v_{p} \end{bmatrix}, \ H \triangleq \begin{bmatrix} H_{1} \\ \vdots \\ H_{p} \end{bmatrix}$

KKT conditions
 Primal feasible

Dual feasible

Zero-duality gap

$$Ax^{*} - r^{*} = b, \quad r^{*} = 0,$$

$$s^{*} + w^{*} = Hx^{*}, \quad w^{*} = 0, \quad s^{*} \in S$$

$$A^{T}y^{*} + H^{T}v^{*} + h^{*} = c, \quad h^{*} = 0,$$

$$z^{*} - v^{*} = 0, \quad z^{*} \in S$$

$$c^{T}x^{*} - b^{T}y^{*} = 0$$

ADMM for the Homogeneous self-dual embedding

$$\begin{bmatrix} h \\ z \\ r \\ w \\ \kappa \end{bmatrix} = \begin{bmatrix} 0 & 0 & -A^{\mathsf{T}} & -H^{\mathsf{T}} & c \\ 0 & 0 & 0 & I & 0 \\ A & 0 & 0 & 0 & -b \\ H & -I & 0 & 0 & 0 \\ -c^{\mathsf{T}} & 0 & b^{\mathsf{T}} & 0 & 0 \end{bmatrix} \begin{bmatrix} x \\ s \\ y \\ v \\ \tau \end{bmatrix}$$

 au, κ are two non-negative and complementary variables

Notational simplicity

$$v \triangleq \begin{bmatrix} h \\ z \\ r \\ w \\ \kappa \end{bmatrix}, \ u \triangleq \begin{bmatrix} x \\ s \\ y \\ v \\ \tau \end{bmatrix}, \ Q \triangleq \begin{bmatrix} 0 & 0 & -A^{T} & -H^{T} & c \\ 0 & 0 & 0 & I & 0 \\ A & 0 & 0 & 0 & -b \\ H & -I & 0 & 0 & 0 \\ -c^{T} & 0 & b^{T} & 0 & 0 \end{bmatrix}, \ \mathcal{K} = \mathbb{R}^{n^{2}} \times \mathcal{S} \times \mathbb{R}^{m} \times \mathbb{R}^{n_{d}} \times \mathbb{R}_{+}$$

• Feasibility problem

find
$$(u,v)$$

s.t. $v = Qu$
 $(u,v) \in \mathcal{K} \times \mathcal{K}^*$

ADMM for the Homogeneous self-dual embedding

find
$$(u,v)$$

s.t. $v = Qu$, $(u,v) \in \mathcal{K} \times \mathcal{K}^*$

• ADMM steps (similar to solver SCS [1])

 $\hat{u}^{k+1} = (I+Q)^{-1}(u^{k}+v^{k}) \longrightarrow \text{Projection to a subspace}$ $u^{k+1} = \mathsf{P}_{\kappa}(\hat{u}^{k+1}-v^{k}) \longrightarrow \text{Projection to cones}$ $V^{k+1} = V^k - \hat{U}^{k+1} + U^{k+1}$ Q is highly structured and sparse $Q \triangleq \begin{bmatrix} 0 & 0 & -A^{T} & -H^{T} & c \\ 0 & 0 & 0 & I & 0 \\ A & 0 & 0 & 0 & -b \\ H & -I & 0 & 0 & 0 \\ -c^{T} & 0 & b^{T} & 0 & 0 \end{bmatrix} \checkmark \text{Block elimination can be applied}$ here to speed up the projection; $\checkmark \text{ Then, the per-iteration cost is the same as applying a splitting method to the primal or dual alone}$ to the primal or dual alone.

[1] O'Donoghue, B., Chu, E., Parikh, N. and Boyd, S. (2016). Conic optimization via operator splitting and homogeneous self-dual embedding. Journal of Optimization Theory and Applications, 169(3), 1042–1068

OUTLINE



2 ADMM for Primal and Dual Sparse SDPs

3 CDCS: Cone Decomposition Conic Solver



- An open source MATLAB solver for partially decomposable conic programs;
- CDCS supports constraints on the following cones:
 - ✓ Free variables
 - ✓ non-negative orthant
 - \checkmark second-order cone
 - \checkmark the positive semidefinite cone.
- Input-output format is aligned with SeDuMi;
- Works with latest YALMIP release.

Syntax:

[x,y,z,info] = cdcs(At,b,c,K,opts);

Download from https://github.com/OxfordControl/CDCS

Random SDPs with block-arrow pattern

- Block size: *d*
- Number of Blocks: /
- Arrow head: h
- Number of constraints: m



Numerical Comparison

- SeDuMi
- sparseCoLO+SeDuMi
- SCS
- sparseCoLO+SCS

CDCS and SCS $\epsilon_{tol} = 10^{-3}$

Numerical Result 10^{4} 10^{4} 10^{4} 10^{3} 10^{3} 103 (s) 10² 10¹ 10^{1} 10^{1} 10^{0} 10^{0} 10 600 1000 1500 2000 2500 2050 100 150 200 10 3500 10 2030 40400Number of constraints, mSize of each block, dNumber of blocks, l

 \square SeDuMi + SparseCoLo+SeDuMi \bigcirc SCS \triangle SparseCoLo+SCS \diamondsuit CDCS (primal) ∇ CDCS (dual)

CPU time for SDPs with block-arrow patterns. Left to right: varying number of constraints; varying number of blocks; varying block size.

Benchmark problems in SDPLIB

Three sets of benchmark problems in SDPLIB (Borchers, 1999):

- 1) Four small and medium-sized SDPs (theta1, theta2, qap5 and qap9);
- 2) Four large-scale sparse SDPs (maxG11, maxG32, qpG11 and qpG51);
- 3) Two infeasible SDPs (infp1 and infd1).

	Small and medium-size $(n \leq 100)$			Large-scale and sparse $(n \ge 800)$				Infeasible		
	theta1	theta2	qap5	qap9	maxG11	maxG32	qpG11	qpG51	infp1	infd1
Original cone size, n	50	100	26	82	800	2000	1600	2000	30	30
Affine constraints, m	104	498	136	748	800	2000	800	1000	10	10
Number of cliques, p	1	1	1	1	598	1499	1405	1675	1	1
Maximum clique size	50	100	26	82	24	60	24	304	30	30
Minimum clique size	50	100	26	82	5	5	1	1	30	30

Table 1. Details of the SDPLIB problems considered in this work.

Result: small and medium-sized instances

		SeDuMi	SparseCoLO+ SeDuMi	SCS	CDCS (primal)	$\begin{array}{c} \mathrm{CDCS} \\ \mathrm{(dual)} \end{array}$	Self-dual
theta1	Total time (s) Pre- time (s) Iterations Objective	0.262 0 14 2.300×10^{1}	$0.279 \\ 0.005 \\ 14 \\ 2.300 \times 10^{1}$	$\begin{array}{c} 0.145 \\ 0.011 \\ 240 \\ 2.300 imes 10^1 \end{array}$	$\begin{array}{c} 0.751 \\ 0.013 \\ 317 \\ 2.299 imes 10^1 \end{array}$	$\begin{array}{c} 0.707 \\ 0.010 \\ 320 \\ 2.299 \times 10^1 \end{array}$	$0.534 \\ 0.012 \\ 230 \\ 2.303 \times 10^{1}$
theta2	Total time (s) Pre- time (s) Iterations Objective	1.45 0 15 3.288×10^{1}	1.55 0.014 15 3.288×10^{1}	0.92 0.018 500 3.288×10^{1}	$ \begin{array}{r} 1.45 \\ 0.046 \\ 287 \\ 3.288 \times 10^1 \end{array} $	$ \begin{array}{r} 1.30 \\ 0.036 \\ 277 \\ 3.288 \times 10^1 \end{array} $	$\begin{array}{r} 0.60 \\ 0.031 \\ 110 \\ 3.287 \times 10^1 \end{array}$
qap5	Total time (s) Pre- time (s) Iterations Objective	0.365 0 12 -4.360×10 ²	0.386 0.006 12 -4.360×10^{2}	0.412 0.026 320 -4.359×10^{2}	$0.879 \\ 0.011 \\ 334 \\ -4.360 \times 10^2$	0.748 0.009 332 -4.364×10^{2}	$1.465 \\ 0.009 \\ 783 \\ -4.362 \times 10^2$
qap9	Total time (s) Pre- time (s) Iterations Objective	6.291 0 25 -1.410×10 ³	$6.751 \\ 0.012 \\ 25 \\ -1.410 \times 10^3$	3.261 0.010 2000 -1.409×10 ³	7.520 0.064 2000 -1.407×10^{3}	7.397 0.036 2000 -1.409×10 ³	$1.173 \\ 0.032 \\ 261 \\ -1.410 \times 10^3$

Table 2. Results for some small and medium-sized SDPs in SDPLIB.

Result: large-sparse instances

		SeDuMi	SparseCoLO+ SeDuMi	SCS	CDCS (primal)	$\begin{array}{c} \mathrm{CDCS} \\ \mathrm{(dual)} \end{array}$	Self-dual
maxG11	Total time (s) Pre- time (s) Iterations Objective	92.0 0 13 6.292×10^2	9.83 2.39 15 6.292×10^2	$160.5 \\ 0.07 \\ 1860 \\ 6.292 \times 10^2$	$\begin{array}{c} 126.6 \\ 3.33 \\ 1317 \\ 6.292 \times 10^2 \end{array}$	$114.1 \\ 4.28 \\ 1306 \\ 6.292 \times 10^2$	23.9 2.45 279 6.295×10^2
maxG32	Total time (s) Pre- time (s) Iterations Objective	$\begin{array}{c} 1.385 \times 10^{3} \\ 0 \\ 14 \\ 1.568 \times 10^{3} \end{array}$	577.4 7.63 15 1.568×10 ³	$\begin{array}{r} 2.487 \times 10^{3} \\ 0.589 \\ 2000 \\ 1.568 \times 10^{3} \end{array}$	520.0 53.9 1796 1.568×10 ³	273.8 55.6 943 $1.568 imes 10^3$	$\begin{array}{r} 87.4 \\ 30.5 \\ 272 \\ 1.568 \times 10^3 \end{array}$
qpG11	Total time (s) Pre- time (s) Iterations Objective	$\begin{array}{r} 675.3 \\ 0 \\ 14 \\ 2.449 \times 10^3 \end{array}$	27.3 11.2 15 2.449×10^{3}	$\begin{array}{r} 1.115 \times 10^{3} \\ 0.57 \\ 2000 \\ 2.449 \times 10^{3} \end{array}$	273.6 6.26 1355 2.449×10 ³	92.5 6.26 656 2.449×10 ³	$\begin{array}{r} 32.1 \\ 3.85 \\ 304 \\ 2.450 \times 10^3 \end{array}$
qpG51	Total time (s) Pre- time (s) Iterations Objective	$\begin{array}{c} 1.984\!\times\!10^{3} \\ 0 \\ 22 \\ 1.182\!\times\!10^{3} \end{array}$	- - -	2.290×10^{3} 0.90 2000 1.288×10^{3}	$\begin{array}{c} 1.627{\times}10^{3} \\ 10.82 \\ 2000 \\ 1.183{\times}10^{3} \end{array}$	$\begin{array}{r} 1.635\!\times\!10^3 \\ 12.77 \\ 2000 \\ 1.186\!\times\!10^3 \end{array}$	538.1 7.89 716 1.181×10 ³

Table 3. Results for some large-scale sparse SDPs in SDPLIB.

Result: Infeasible instances

		SeDuMi	SparseCoLO+ SeDuMi	SCS	$\begin{array}{c} \mathrm{CDCS} \\ \mathrm{(primal)} \end{array}$	$\begin{array}{c} \mathrm{CDCS} \\ \mathrm{(dual)} \end{array}$	Self-dual
	Total time (s)	0.063	0.083	0.062	*	*	0.18
info 1	Pre- time (s)	0	0.010	0.016	*	*	0.010
inipi	Iterations	2	2	20	*	*	104
	Status	Infeasible	Infeasible	Infeasible	*	*	Infeasible
	Total time (s)	0.125	0.140	0.050	*	*	0.144
infd1	Pre- time (s)	0	0.009	0.013	*	*	0.009
	Iterations	4	4	40	*	*	90
	Status	Infeasible	Infeasible	Infeasible	*	*	Infeasible

Table 4. Results for two infeasible SDPs in SDPLIB.

Result: CPU time per iteration

Table 5. CPU time per iteration (s) for some SDPs in SDPLIB

	SCS	CDCS (primal)	$\begin{array}{c} \mathrm{CDCS} \\ \mathrm{(dual)} \end{array}$	Self-dual	
theta1 theta2 qap5 qap9	6×10^{-4} 1.8×10^{-3} 1.2×10^{-3} 1.5×10^{-3}	2.3×10^{-3} 5.1×10^{-3} 2.6×10^{-3} 3.6×10^{-3}	2.2×10^{-3} 4.7×10^{-3} 2.2×10^{-3} 3.7×10^{-3}	2.3×10^{-3} 5.5×10^{-3} 1.9×10^{-3} 4.2×10^{-3}	
maxG11 maxG32 qpG11 qpG51	$\begin{array}{c} 0.086 \\ 1.243 \\ 0.557 \\ 1.144 \end{array}$	$0.094 \\ 0.260 \\ 0.198 \\ 0.808$	$\begin{array}{c} 0.084 \\ 0.231 \\ 0.132 \\ 0.811 \end{array}$	$\begin{array}{c} 0.077 \\ 0.209 \\ 0.093 \\ 0.741 \end{array}$	

✓ Our codes are currently written in MATLAB✓ SCS is implemented in C.

Conclusion



- Introduced a conversion framework for sparse SDPs
- Developed efficient ADMM algorithms

 \checkmark Primal and dual standard form;

✓ The homogeneous self-dual embedding;

suitable for firstorder methods

Ongoing work

- Develop ADMM algorithms for sparse SDPs arising in SOS.
- Applications in networked systems and power systems

Thank you for your attention!

 CDCS: Download from https://github.com/OxfordControl/CDCS

- Zheng, Y., Fantuzzi G., Papachristodoulou A., Goulart, P., and Wynn, A. (2016) Fast ADMM for Semidefinite Programs with Chordal Sparsity. *arXiv preprint arXiv:1609.06068*
- Zheng, Y., Fantuzzi, G., Papachristodoulou, A., Goulart, P., & Wynn, A. (2016) Fast ADMM for homogeneous self-dual embeddings of sparse SDPs. *arXiv preprint arXiv:1611.01828*