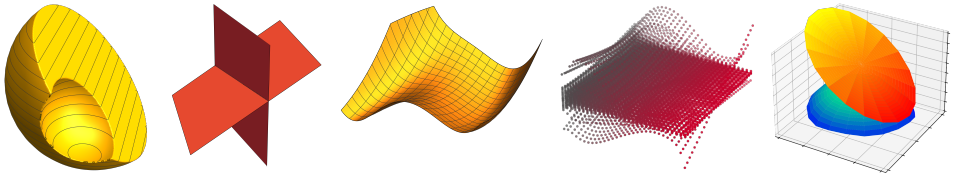
# Accurately and efficiently solving structured nonconvex optimization problems

**Alex L. Wang**
Carnegie Mellon University



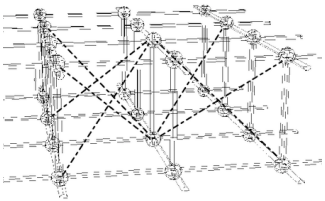
These slides are publicly available at `cs.cmu.edu/~alw1`

# Collaborators

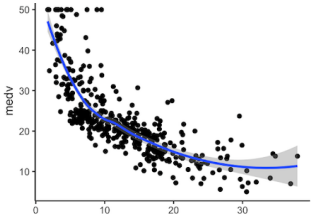
Carnegie Mellon University (OR, Math, CS),
Northwestern University, Fudan University,
Peking University

# Convex optimization

- Convex optimization is influential in many different fields



**Engineering**
Controller stability, power allocation, truss design, +




**Statistics**
(Linear) Regression, parameter estimation, +

**Finance**
Portfolio optimization, risk analysis, +

- Convex optimization is accurate and efficient

# Convex optimization, meet nonconvex problems

- Unfortunately, many practical optimization problems are nonconvex
- Example: Low-rank matrix completion (**Netflix problem**)

Movies

Users

$$\begin{pmatrix} 5 & 4 & ? & ? & 4 \\ 3 & ? & ? & 3 & ? \\ ? & 2 & 4 & 1 & 1 \\ ? & 3 & ? & ? & 4 \end{pmatrix}$$

$$\min_{X \in \mathbb{R}^{n \times k}} \{\mathrm{rank}(X) : X \text{ agrees with revealed entries}\}$$

- Rank constraints, binary constraints, sparsity constraints
- Generally hard, **but not always!**
- Some nonconvex problems can be solved using convex optimization

# Research goal

## Long-term research goal

Understand **structures within nonconvex problems** that enable us to solve them "well" using **convex optimization**

- **Completed work**:
  - Nonconvex problems: quadratically constrained quadratic programs (QCQPs)
  - Convex relaxations: semidefinite programs (SDPs)

## Today's questions

Understand **structures within QCQPs** that enable us to solve them exactly and efficiently **using SDPs**

# Today's outline

- **Preliminaries**
  QCQPs and their applications, the SDP relaxation
- **Understand structures within QCQPs that enable us to solve them**...
  - **exactly** [IPCO 20], [Math. Prog. 21], [Math. Prog. *under review*]
    Objective value, convex hull exactness, applications
  - **efficiently** [Math. Prog. 20], [SIAM J. Optim. *under review*], [Ongoing]
    The generalized trust-region subproblem and regular QCQPs
- **Conclusion and future directions**

1 Preliminaries

2 Objective value exactness, convex hull exactness, applications

3 Efficient algorithms for structured QCQPs

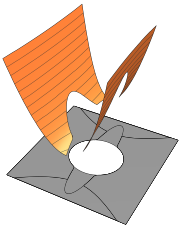
4 Conclusion and future directions

# Quadratically constrained quadratic programs (QCQPs)

- $q_{\text{obj}}, q_1, \ldots, q_m : \mathbb{R}^n \to \mathbb{R}$ quadratic (possibly nonconvex!)

$$q_i(x) = x^\mathsf{T} A_i x + 2b_i^\mathsf{T} x + c_i$$

$$\text{Opt} := \inf_{x \in \mathbb{R}^n} \left\{ q_{\text{obj}}(x) : \; q_i(x) \leq 0, \, \forall i \in [m] \right\}$$

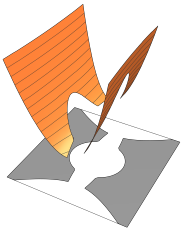


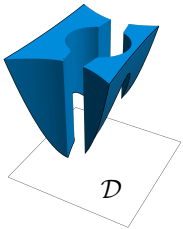
- Highly expressive:
  - MAX-CUT, MAX-CLIQUE, pooling, truss design, facility location, production planning
  - binary programs $x_1(1 - x_1) = 0$
  - polynomial optimization problems $x_1 x_2 = z_{12}$
- NP-hard in general

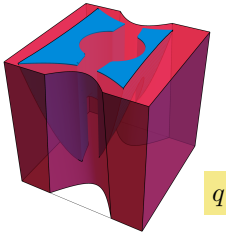
# The QCQP epigraph

- QCQP epigraph

$$\mathcal{D} := \left\{ (x,t) \in \mathbb{R}^{n+1} : \begin{array}{l} q_{\text{obj}}(x) \le t \\ q_i(x) \le 0,\ \forall i \in [m] \end{array} \right\}$$





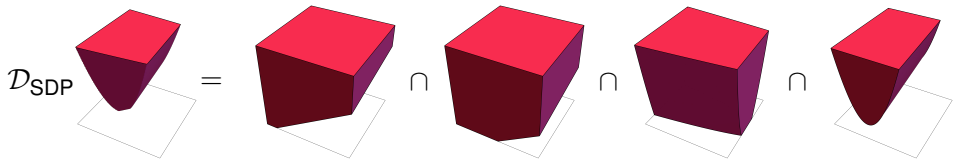




- How can we derive ~~convex~~ relaxations of $\mathcal{D}$?
- If $\gamma \in \mathbb{R}^m_+$, then

$$\forall (x,t) \in \mathcal{D}, \quad \underbrace{q_{\text{obj}}(x) + \sum_{i=1}^{m} \gamma_i q_i(x)}_{=: q(\gamma, x)} \le t$$

# The SDP relaxation

- SDP relaxation = impose all convex aggregated inequalities!




- Formally,

$$\Gamma := \{\gamma \in \mathbb{R}^m_+ : \ q(\gamma, x) \text{ is convex in } x\} = \left\{\gamma \in \mathbb{R}^m_+ : \ A_{\text{obj}} + \sum_{i=1}^m \gamma_i A_i \succeq 0\right\}$$

$$\mathcal{D}_{\text{SDP}} := \bigcap_{\gamma \in \Gamma} \{(x,t) : \ q(\gamma, x) \le t\} = \left\{(x,t) \in \mathbb{R}^{n+1} : \ \sup_{\gamma \in \Gamma} q(\gamma, x) \le t\right\}$$

$$\text{Opt}_{\text{SDP}} := \inf_{x \in \mathbb{R}^n} \sup_{\gamma \in \Gamma} q(\gamma, x)$$

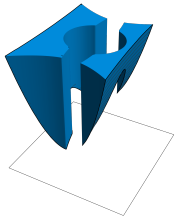
# The usual SDP relaxation

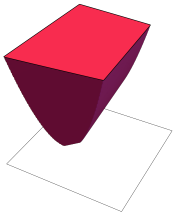
$$
\begin{aligned}
&\inf_{x\in\mathbb{R}^n} \{q_{\mathsf{obj}}(x) : q_i(x) \le 0,\ \forall i \in [m]\} \\
&\quad = \inf_{x\in\mathbb{R}^n, X\in\mathbb{S}^n} \left\{ \langle A_{\mathsf{obj}}, X\rangle + 2b_{\mathsf{obj}}^\top x + c_{\mathsf{obj}} : \begin{array}{l} X = xx^\top \\ \langle A_i, X\rangle + 2b_i^\top x + c_i \le 0,\ \forall i \in [m] \end{array} \right\} \\
&\quad \ge \inf_{x\in\mathbb{R}^n, X\in\mathbb{S}^n} \left\{ \langle A_{\mathsf{obj}}, X\rangle + 2b_{\mathsf{obj}}^\top x + c_{\mathsf{obj}} : \begin{array}{l} X - xx^\top \succeq 0 \\ \langle A_i, X\rangle + 2b_i^\top x + c_i \le 0,\ \forall i \in [m] \end{array} \right\} \\
&\quad = \inf_{x\in\mathbb{R}^n} \inf_{X\in\mathbb{S}^n} \dots \\
&\quad = \inf_{x\in\mathbb{R}^n} \sup_{\gamma\in\Gamma} q(\gamma, x)
\end{aligned}
$$

# Preliminaries recap

- Main objects of interest

| | nonconvex QCQP | convex SDP |
|---|---|---|
| Optimum value | $\mathrm{Opt}$ | $\mathrm{Opt}_{\mathrm{SDP}}$ |
| Epigraph | $\mathcal{D}$ | $\mathcal{D}_{\mathrm{SDP}}$ |





- Useful for analysis:
  - $q(\gamma, x) =$ Lagrangian function
  - $\Gamma =$ aggregation weights giving convex $q(\gamma, x)$

1 Preliminaries

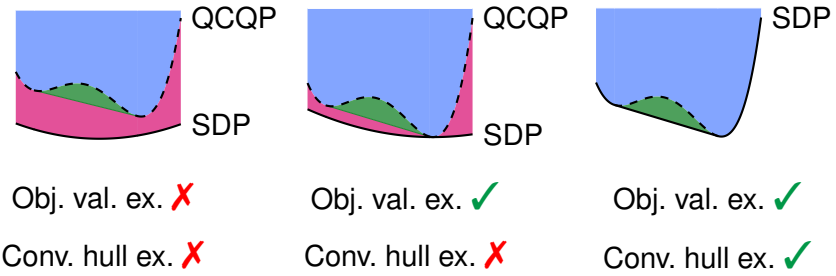
2 Objective value exactness, convex hull exactness, applications

3 Efficient algorithms for structured QCQPs

4 Conclusion and future directions

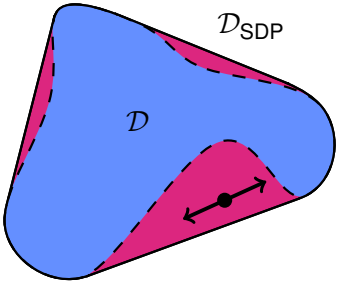
# Forms of exactness

- What does exactness mean?
  - Objective value exactness: $\mathrm{Opt} = \mathrm{Opt}_{\mathsf{SDP}}$
  - Convex hull exactness: $\mathrm{conv}(\mathcal{D}) = \mathcal{D}_{\mathsf{SDP}}$ $\leftarrow$ convexification of substructures



# Convex hull exactness

- $\mathrm{conv}(\mathcal{D}) \stackrel{?}{=} \mathcal{D}_{\mathsf{SDP}}$




$\mathrm{conv}(\mathcal{D}) = \mathcal{D}_{\mathsf{SDP}} \iff$ "Given any point in $\mathcal{D}_{\mathsf{SDP}} \setminus \mathcal{D}$, exists direction such that can move forward and backward inside $\mathcal{D}_{\mathsf{SDP}}$"

- Can carry out this idea for QCQPs!
- Sufficient conditions based on abstract properties $\longrightarrow$ concrete conditions

# Sufficient conditions based on abstract properties

## Theorem ([Math. Prog. 21])

Suppose $\Gamma$ polyhedral. If for every semidefinite face $\mathcal{F} \trianglelefteq \Gamma$,

$$\operatorname{aff}\left(\operatorname{Proj}_{\mathcal{V}(\mathcal{F})}\{b(\gamma) : \gamma \in \mathcal{F}\}\right) \neq \mathcal{V}(\mathcal{F}),$$

then $\operatorname{conv}(\mathcal{D}) = \mathcal{D}_{\text{SDP}}$.

## Theorem ([Math. Prog. *under review*])

If for every $(x,t) \in \mathcal{D}_{\text{SDP}} \setminus \mathcal{D}$,

$$\left\{ (x',t') \in \mathbb{R}^{n+1} : \begin{array}{l} x' \in \mathcal{V}(\mathcal{F}) \\ \langle A(\eta)x + b(\eta), x' \rangle - t' = 0,\ \forall (1,\eta) \in (\mathcal{F}^{\triangle})^{\perp} \end{array} \right\} \neq \{0\},$$

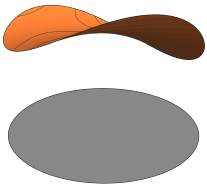
then $\operatorname{conv}(\mathcal{D}) = \mathcal{D}_{\text{SDP}}$.

Based on: **[IPCO 19]**, **[Math. Prog. 21]**, **[Math. Prog. *under review*]**

# Example: the trust-region subproblem

- Convex hull exactness in the case of single ball constraint

$$\text{Opt} = \inf_{x \in \mathbb{R}^n} \left\{ q_{\text{obj}}(x) : \|x\|^2 \leq 1 \right\}$$



- **Applications**:
  - Nonlinear minimization (trust-region methods), combinatorial optimization, robust optimization

Based on: **[IPCO 19]**, **[Math. Prog. 20]**
Related: Yakubovich [1971], Yıldıran [2009], Ho-Nguyen and Kılınç-Karzan [2017]

# Example: QCQPs with symmetry

- Convex hull exactness in the case of "highly symmetric" QCQPs
- Suppose $A_{\text{obj}} = I_k \otimes \mathbb{A}_{\text{obj}}$, $A_i = I_k \otimes \mathbb{A}_i$ for all $i \in [m]$

$$I_k \otimes \mathbb{A} = \begin{pmatrix} \mathbb{A} & & & \\ & \mathbb{A} & & \\ & & \ddots & \\ & & & \mathbb{A} \end{pmatrix}$$

  and $k \geq m$
- **Applications**:
  - Robust least squares, sphere packing, QCQPs with spherical constraints, orthogonal Procrustes problem

Based on: **[Math. Prog. *under review*]**
Related: Beck [2007], Beck et al. [2012]

# Example: QCQPs with sign-definite linear terms

- Objective value exactness in the case of diagonal, sign-definite QCQPs
- Suppose $A_{\text{obj}}, \ldots, A_m$ diagonal and

$$\forall j \in [n], \left\{(b_{\text{obj}})_j, (b_1)_j, \ldots, (b_m)_j\right\} \text{ have the same sign}$$

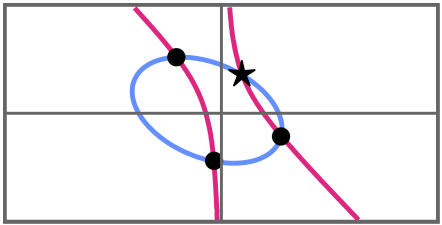
- Obj. val. exactness for

$$\min_{x \in \mathbb{R}^n} \left\{ x^\top A_{\text{obj}} x + 2 b_{\text{obj}}^\top x + c_{\text{obj}} : \|x\|_2 \leq 1, \|x\|_\infty \leq \alpha \right\}$$

Based on: **[Tut. Oper. Res. 21]**
Related: Burer and Ye [2019], Sojoudi and Lavaei [2014]

# Example: Random underconstrained quadratic systems

- Obj. val. exactness in the case of random underconstrained quadratic systems
- Solve $$\inf_{x \in \mathbb{R}^n} \left\{ \|x\|^2 : \; q_i(x) = 0, \; \forall i \in [m] \right\}$$



- Fix $m$, let $n \to \infty$, if data generated "as Gaussians", then objective value exactness w.p. $1 - o(1)$ (in fact, optimizer exactness)

Based on: **[Math. Prog. *under review*]**
Related: Burer and Ye [2019], Locatelli [2020]

# Summary of Part 1

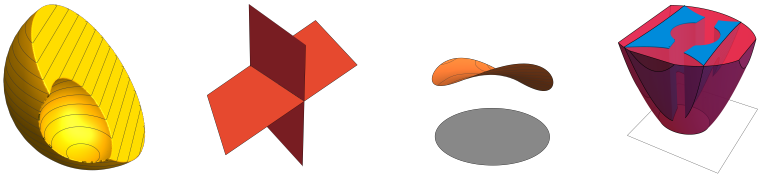
- Sufficient conditions for convex hull exactness
- **Necessary and sufficient** if $\Gamma$ is polyhedral (dual facially exposed)
- Sufficient conditions for objective value exactness
- Rank-one-generated (ROG) cones: QCQP-SDP analogue of integral polyhedra
- **Applications**:
  - Semi-random QCQPs, ratios of quadratic functions

| Exactness | ROG Cones |
| --- | --- |
| [IPCO 20], [Math. Prog. 21], [Math. Prog. *under review*] | [Tut. Oper. Res. 21], [Math. Oper. Res. 21] |

## Long-term research goal

Understand **structures within nonconvex problems** that enable us to solve them "well" using **convex optimization**

**SDPs provide exact reformulations for broad classes of QCQPs!**



1 Preliminaries

2 Objective value exactness, convex hull exactness, applications

3 Efficient algorithms for structured QCQPs

4 Conclusion and future directions

# Revisiting the SDP relaxation

- Usual SDP relaxation in $x \in \mathbb{R}^n$ and $X \in \mathbb{S}^n \implies \Theta(n^2)$ variables
  - Interior point method $\implies \Theta\left(n^2 + m^2\right)$ storage
- Can we solve SDP more efficiently/using less storage if it is exact?
- Our view:

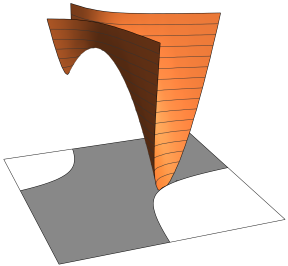
$$\mathrm{Opt}_{\mathsf{SDP}} \coloneqq \inf_{x \in \mathbb{R}^n} \left( \sup_{\gamma \in \Gamma} q(\gamma, x) \right)$$

  is a minimization problem in the original space $\implies n$ variables

# The generalized trust-region subproblem (GTRS)

- Special setting with single constraint ($\leq$ or $=$)

$$\text{Opt} := \inf_{x \in \mathbb{R}^n} \{q_{\text{obj}}(x) : q_1(x) \leq 0\}$$



- **TRS Applications**:
  - nonlinear programming (trust-region methods), combinatorial optimization, robust optimization

- **GTRS Applications**:
  - minimizing quartics of the form $q(x, p(x))$

$$\inf_{x \in \mathbb{R}^n, \alpha} \{q(x, \alpha) : \alpha = p(x)\}$$

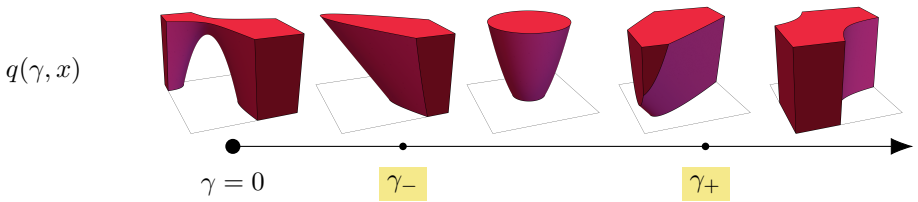
(source localization, constrained rank-one approximation), regression with adversarial data, iterative QCQP solvers

Based on: **[Math. Prog. 20]**, **[SIAM. J. Optim. *under review*]**, **[ICML *under review*]**

# Linear-time algorithm for the GTRS

- Convex hull exactness holds $\quad \operatorname{conv}(\mathcal{D}) = \mathcal{D}_{\text{SDP}} = \left\{ (x,t) : \sup_{\gamma \in \Gamma} q(\gamma, x) \leq t \right\}$

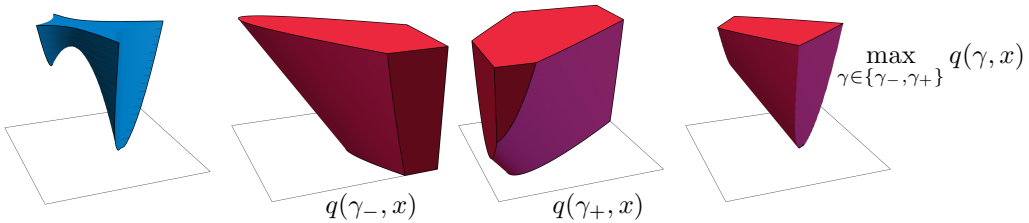
- Recall $\quad \Gamma = \{\gamma \in \mathbb{R}_+ : q(\gamma, x) \text{ is convex in } x\}$




Based on: **[Math. Prog. 20]**
Related: Hazan and Koren [2016], Ho-Nguyen and Kılınç-Karzan [2017], Jiang and Li [2019, 2020]

# Linear-time algorithm for the GTRS

- $\Gamma = [\gamma_-, \gamma_+] \implies \mathrm{Opt} = \mathrm{Opt}_{\mathrm{SDP}} = \inf_{x \in \mathbb{R}^n} \max_{\gamma \in \{\gamma_-, \gamma_+\}} q(\gamma, x)$




- **Algorithmic idea**
  - Compute $\gamma_-$ and $\gamma_+$ to some accuracy
  - Apply accelerated gradient descent

$$\implies \tilde{O}\left(\frac{N}{\sqrt{\epsilon}} \log\left(\frac{n}{p}\right) \log\left(\frac{1}{\epsilon}\right)\right) \approx \frac{1}{\sqrt{\epsilon}}$$

Based on: **[Math. Prog. 20]**
Related: Hazan and Koren [2016], Ho-Nguyen and Kılınç-Karzan [2017], Jiang and Li [2019, 2020]

# Is this running time optimal?

- GTRS:

$$\inf_{x \in \mathbb{R}^n} \left\{ q_{\text{obj}}(x) : q_1(x) \leq 0 \right\} \qquad \tilde{O}\left( \frac{N}{\sqrt{\epsilon}} \log\left( \frac{n}{p} \right) \log\left( \frac{1}{\epsilon} \right) \right)$$

- Minimum eigenvalue:

$$\inf_{x \in \mathbb{R}^n} \left\{ x^\mathsf{T} A x : \|x\|^2 = 1 \right\} \qquad O\left( \frac{N}{\sqrt{\epsilon}} \log\left( \frac{n}{p} \right) \right)$$

- Smooth convex quadratic:

$$\inf_{x \in \mathbb{R}^n} q_{\text{obj}}(x) \qquad O\left( \frac{N}{\sqrt{\epsilon}} \right)$$

- But, these are hardest problems within the GTRS!
- Can do better when regularity $\mu^* > 0$

Based on: **[Math. Prog. 20]**
Related: Kuczynski and Wozniakowski [1992], Nesterov [2018]

# Linear convergence for regular GTRS

- $\mu^* > 0$ holds for most GTRS
- Dual problem

$$\mathrm{Opt}_{\mathsf{SDP}} := \inf_{x \in \mathbb{R}^n} \sup_{\gamma \in \Gamma} q(\gamma, x) = \sup_{\gamma \in \Gamma} \inf_{x \in \mathbb{R}^n} q(\gamma, x)$$

## Definition

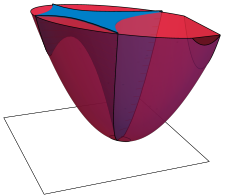
Let $\gamma^*$ be dual optimizer. Define $\mu^* := \lambda_{\min}(A_{\mathsf{obj}} + \gamma^* A_1)$. GTRS instance is regular if $\mu^* > 0$.

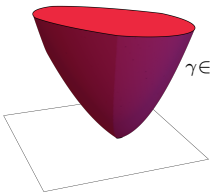
- Minimum eigenvalue problem is not regular
- Some instances of smooth quadratic minimization are not regular

Based on: **[SIAM J. Optim. *under review*]**
Related: Carmon and Duchi [2018]

# Linear convergence for regular GTRS

- Suppose $\mu^* > 0$





$$\max_{\gamma \in \{\tilde{\gamma}_-, \tilde{\gamma}_+\}} q(\gamma, x)$$

- Suppose $\gamma^* \in [\tilde{\gamma}_-, \tilde{\gamma}_+] \subseteq \Gamma \implies \mathrm{Opt} = \inf_{x \in \mathbb{R}^n} \max_{\gamma \in \{\tilde{\gamma}_-, \tilde{\gamma}_+\}} q(\gamma, x)$
- Suffices to estimate $\gamma^*$ roughly and can exploit strong convexity

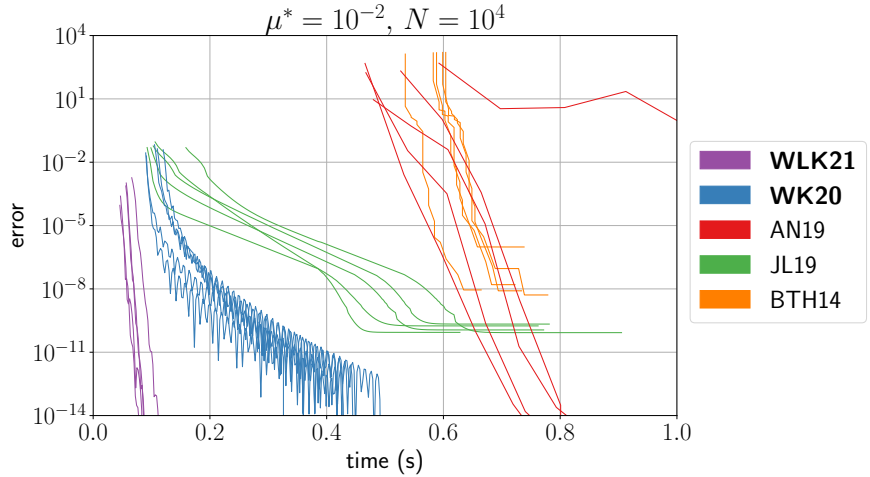
$$\tilde{O}\left(\frac{N}{\sqrt{\mu^*}} \log\left(\frac{1}{\mu^*}\right) \log\left(\frac{n}{p}\right) \log\left(\frac{1}{\epsilon}\right)\right) \approx \log\left(\frac{1}{\epsilon}\right)$$

- Linear in $N$ and $\log(1/\epsilon)$

---

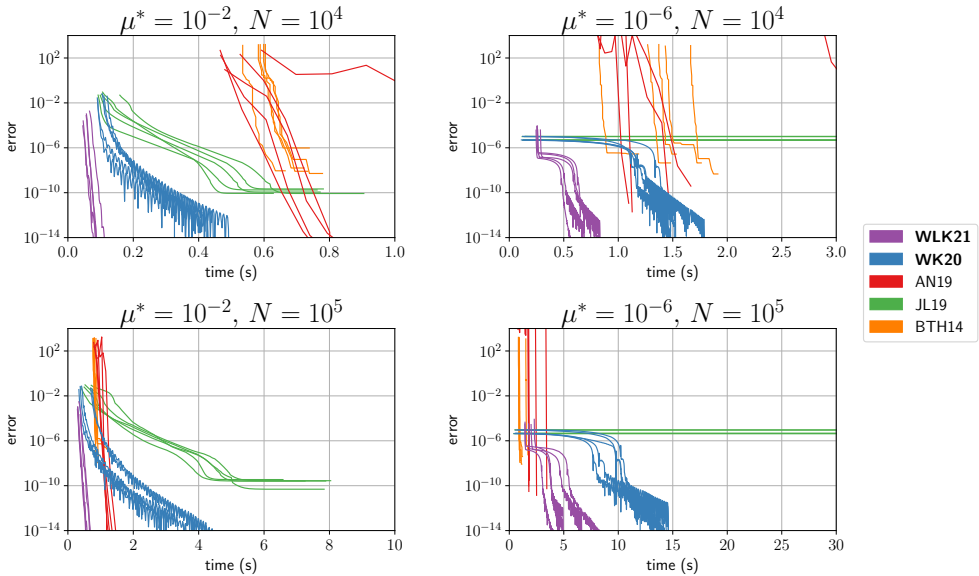
Based on: **[SIAM J. Optim. *under review*]**
Related: Carmon and Duchi [2018]

# Numerical experiments, n = 1,000

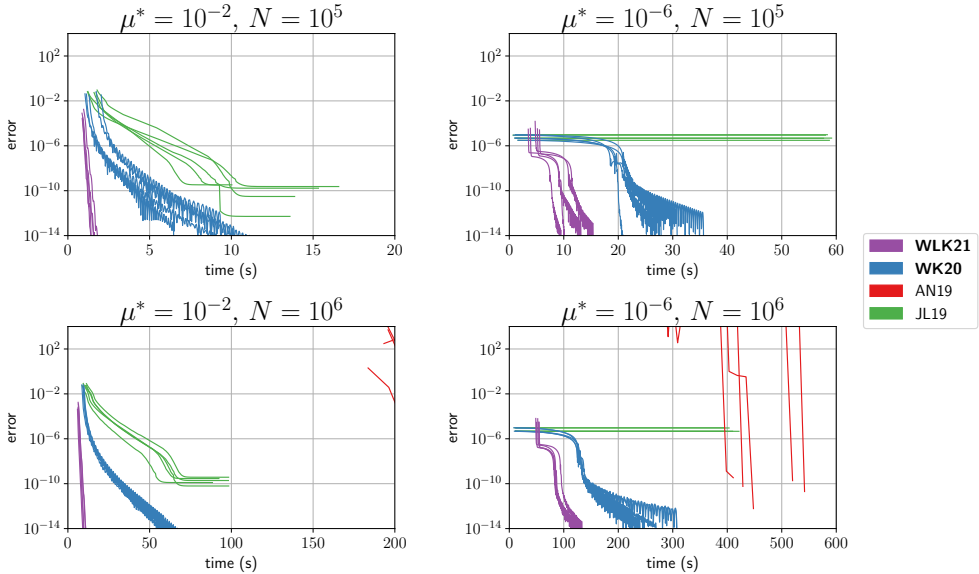



Based on: **[Math. Prog. 20]**, **[SIAM. J. Optim. *under review*]**
Related: Ben-Tal and den Hertog [2014], Jiang and Li [2019], Adachi and Nakatsukasa [2019]

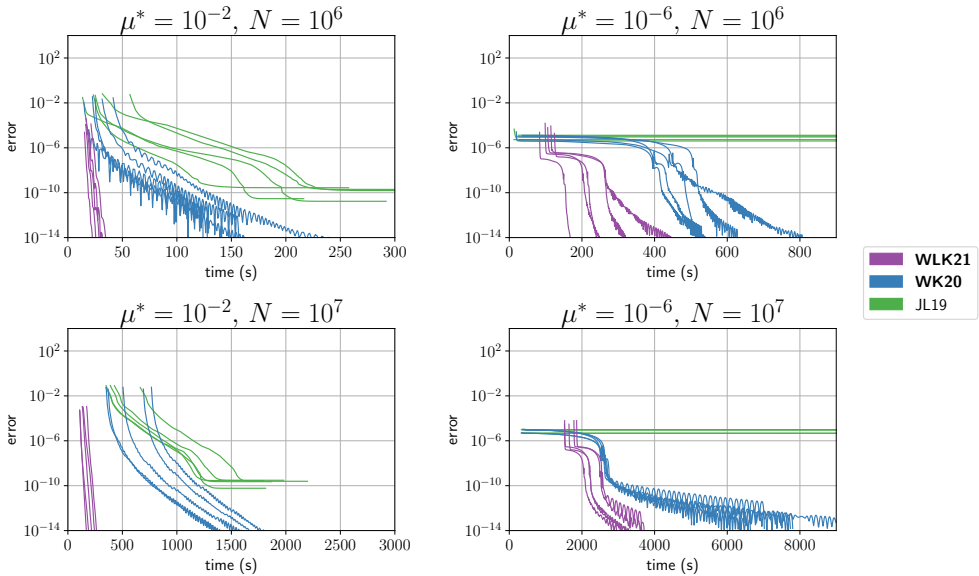
# Numerical experiments, n = 1,000



# Numerical experiments, n = 10,000



# Numerical experiments, n = 100,000



# Efficient algorithms for regular QCQPs

- Regularity can also be defined for general QCQPs! (low-rank SDPs)

## Definition

Let $\gamma^*$ be dual optimizer. Define $\mu^* := \lambda_{\min}\left(A_{\text{obj}} + \sum_{i=1}^m \gamma_i^* A_i\right)$. QCQP instance is regular if $\mu^* > 0$.

- $\mu^* > 0 \implies$ objective value exactness
- $\mu^* > 0$ in a number of statistical recovery problems: phase-retrieval, clustering
- Contributions
  - **Conceptual**: If $\mu^* > 0$, then can construct strongly convex reformulation in time independent of $\epsilon$
  - **Algorithmic**: Can solve strongly convex reformulation efficiently $\Longleftarrow$ grad-map
  - **Practical**: Preliminary implementation and numerical experiments

Based on: **[Ongoing]**

# Constructing a strongly convex reformulation of regular QCQP

- Consider a subgradient method for **dual problem**

$$\sup_{\gamma \in \Gamma} \inf_{x \in \mathbb{R}^n} q(\gamma, x)$$

- Iterates $\gamma^{(1)}, \gamma^{(2)}, \ldots$ and bounds $\left\|\gamma^{(i)} - \gamma^*\right\| \leq \rho^{(i)}$
- Will stop once $B(\gamma^{(t)}, \rho^{(t)}) \subseteq \operatorname{int}(\Gamma)$
- Then, $$\operatorname{Opt} = \inf_{x \in \mathbb{R}^n} \sup_{\gamma \in B(\gamma^{(t)}, \rho^{(t)})} q(\gamma, x)$$
- Independent of $\epsilon$, can be done with $O(m+n)$ storage

Based on: **[ongoing]**
Related: Ding et al. [2021]

# Preliminary numerical experiments

- 10 synthetic instances: $n = 5000$, $m = 50$, density $= 0.001$, $\mu^* = 0.1$
- Primal-dual solver

| Section | ncalls | avg time | %tot | avg error |
|---|---|---|---|---|
| Total | 10 | 1740s | 100% | |
| dual_solve | 10 | 1691s | 97.2% | 1.15e-02 |
| eigsolve | 33.1k | 506ms | 96.0% | |
| primal_solve | 10 | 48.7s | 2.80% | 3.61e-12 |

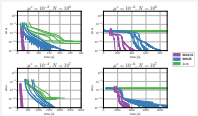
- Splitting Conic Solver (SCS): avg time 18150s

# Summary of Part 2

- Efficient algorithms for the GTRS
- Efficient algorithms for regular QCQPs (low-rank SDPs)
- Algorithms for diagonalizing QCQPs

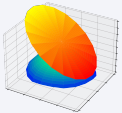
**GTRS**

[Math. Prog. 20],
[SIAM J. Optim. *under review*],
[ICML *under review*]



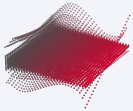
**Regular QCQPs**

[Ongoing]



**Diagonalizing QCQPs**

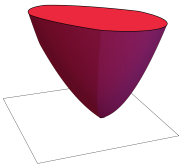
[Math. Prog. *under review*]

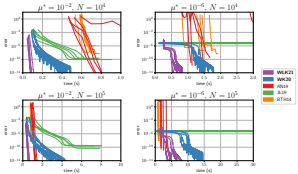


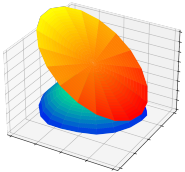
## Long-term research goal

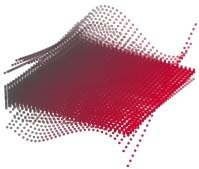
Understand **structures within nonconvex problems** that enable us to solve them "well" using **convex optimization**

**Some nonconvex problems can be solved efficiently via first-order methods!**







1 Preliminaries

2 Objective value exactness, convex hull exactness, applications

3 Efficient algorithms for structured QCQPs

4 Conclusion and future directions

# Future work

## Long-term research goal

Understand **structures within nonconvex problems** that enable us to solve them "well" using **convex optimization**

- **Solving nonconvex problems accurately**
  - **Completed**: SDPs provide exact reformulations for broad classes of QCQPs!
  - **Future**:
    - Can we understand approximation quality systematically within general framework?
    - Can we understand exactness/approximation for more powerful convex relaxations?

# Future work

## Long-term research goal

Understand **structures within nonconvex problems** that enable us to solve them "well" using **convex optimization**

- **Solving nonconvex problems efficiently**
  - **Completed**: Some nonconvex problems can be solved efficiently via first-order methods!
  - **Future**:
    - Exactness $\approx$ efficiency?
    - Can we develop efficient algorithms for semidefinite programs with low-rank solutions
    - Can we approximate "expensive" tools (e.g., SDPs) with cheap tools (e.g., linear programs, second-order cone programs)

# Summary of my research

## Exactness

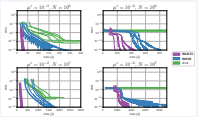
[IPCO 20], [Math. Prog. 21], [Math. Prog. *under review*]

## ROG Cones

[Tut. Oper. Res. 21], [Math. Oper. Res. 21]

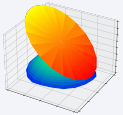
## GTRS

[Math. Prog. 20], [SIAM J. Optim. *under review*], [ICML *under review*]



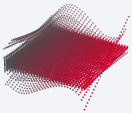
## Regular QCQPs

[Ongoing]



## Diagonalizing QCQPs

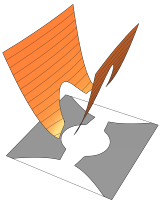
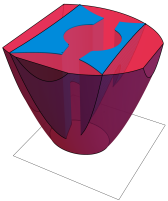
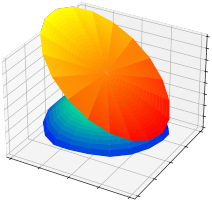
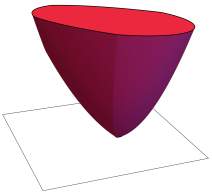
[Math. Prog. *under review*]



## Long-term research goal

Understand **structures within nonconvex problems** that enable us to solve them "well" using **convex optimization**

**Thank you! Questions?**

# References I


Adachi, S. and Nakatsukasa, Y. (2019). Eigenvalue-based algorithm and analysis for nonconvex QCQP with one constraint. *Math. Program.*, 173:79–116.

Argue, C., Kılınç-Karzan, F., and Wang, A. L. (2020). Necessary and sufficient conditions for rank-one generated cones. *arXiv preprint*, 2007.07433.

Beck, A. (2007). Quadratic matrix programming. *SIAM J. Optim.*, 17(4):1224–1238.

Beck, A., Drori, Y., and Teboulle, M. (2012). A new semidefinite programming relaxation scheme for a class of quadratic matrix problems. *Oper. Res. Lett.*, 40(4):298–302.

Ben-Tal, A. and den Hertog, D. (2014). Hidden conic quadratic representation of some nonconvex quadratic optimization problems. *Math. Program.*, 143:1–29.

Burer, S. and Ye, Y. (2019). Exact semidefinite formulations for a class of (random and non-random) nonconvex quadratic programs. *Math. Program.*, 181:1–17.

Carmon, Y. and Duchi, J. C. (2018). Analysis of Krylov subspace solutions of regularized nonconvex quadratic problems. pages 10728–10738.

Ding, L., Yurtsever, A., Cevher, V., Tropp, J. A., and Udell, M. (2021). An optimal-storage approach to semidefinite programming using approximate complementarity. *SIAM J. Optim.*, 31(4):2695–2725.

# References II


Dutta, A., Vijayaraghavan, A., and Wang, A. L. (2017). Clustering stable instances of Euclidean k-means. In *Advances in Neural Information Processing Systems*, pages 6500–6509.

Hazan, E. and Koren, T. (2016). A linear-time algorithm for trust region problems. *Math. Program.*, 158:363–381.

Ho-Nguyen, N. and Kılınç-Karzan, F. (2017). A second-order cone based approach for solving the Trust Region Subproblem and its variants. *SIAM J. Optim.*, 27(3):1485–1512.

Jiang, R. and Li, D. (2019). Novel reformulations and efficient algorithms for the Generalized Trust Region Subproblem. *SIAM J. Optim.*, 29(2):1603–1633.

Jiang, R. and Li, D. (2020). A linear-time algorithm for generalized trust region problems. *SIAM J. Optim.*, 30(1):915–932.

Kılınç-Karzan, F. and Wang, A. L. (2021). Exactness in SDP relaxations of QCQPs: Theory and applications. Tut. in Oper. Res. INFORMS.

Kuczynski, J. and Wozniakowski, H. (1992). Estimating the largest eigenvalue by the power and Lanczos algorithms with a random start. *SIAM J. Matrix Anal. Appl.*, 13(4):1094–1122.

Locatelli, M. (2020). KKT-based primal-dual exactness conditions for the Shor relaxation. *arXiv preprint*, 2011.05033.

# References III


Miller, G. L., Walkington, N. J., and Wang, A. L. (2019). Hardy-Muckenhoupt bounds for Laplacian eigenvalues. In *Approximation, Randomization, and Combinatorial Optimization. Algorithms and Techniques (APPROX/RANDOM 2019)*, pages 8:1–8:19.

Nesterov, Y. (2018). *Lectures on convex optimization*, volume 137 of *Springer Optimization and Its Applications*. Springer.

Sojoudi, S. and Lavaei, J. (2014). Exactness of semidefinite relaxations for nonlinear optimization problems with underlying graph structure. *SIAM J. Optim.*, 24(4):1746–1778.

Wang, A. L. and Jiang, R. (2021). New notions of simultaneous diagonalizability of quadratic forms with applications to QCQPs. *arXiv preprint*, 2101.12141.

Wang, A. L. and Kılınç-Karzan, F. (2020a). The generalized trust region subproblem: solution complexity and convex hull results. *Math. Program.* Forthcoming.

Wang, A. L. and Kılınç-Karzan, F. (2020b). A geometric view of SDP exactness in QCQPs and its applications. *arXiv preprint*, 2011.07155.

Wang, A. L. and Kılınç-Karzan, F. (2020c). On convex hulls of epigraphs of QCQPs. In *Integer Programming and Combinatorial Optimization (IPCO 2020)*, pages 419–432. Springer.

# References IV


Wang, A. L. and Kılınç-Karzan, F. (2021). On the tightness of SDP relaxations of QCQPs. *Math. Program.* Forthcoming.

Wang, J., Huang, W., Jiang, R., Li, X., and Wang, A. L. (2022). Solving stackelberg prediction games with least squares loss via spherically constrained least squares.

Yakubovich, V. A. (1971). S-procedure in nonlinear control theory. *Vestnik Leningrad Univ. Math.*, pages 62–77.

Yıldıran, U. (2009). Convex hull of two quadratic constraints is an LMI set. *IMA J. Math. Control Inform.*, 26(4):417–450.

# Rank-one-generated cones

## Definition

Cone $\mathcal{S} \subseteq \mathbb{S}^n_+$ is rank-one-generated (ROG) if $\mathcal{S} = \operatorname{conv}\left(\mathcal{S} \cap \left\{xx^\top\right\}\right)$.

Compare: $P \subseteq [0,1]^n$ is integral if $P = \operatorname{conv}(P \cap \{0,1\}^n)$

- Given QCQP, if constraints correspond to ROG cone, then objective value exactness and convex hull exactness regardless of objective function
- Suppose $\mathcal{S} = \left\{X \in \mathbb{S}^n_+ : \langle M, X\rangle \leq 0, \forall M \in \mathcal{M}\right\}$

## Goal

What properties of $\mathcal{M} = \{M_1, \ldots, M_k\}$ imply $\mathcal{S}$ is ROG?

Based on: **[Math. Oper. Res. 21]**, **[Tut. Oper. Res. 21]**

**Theorem (Sufficient conditions)**

$\mathcal{S}$ is ROG if

- for all $i \neq j$, there exists $(\alpha, \beta) \neq (0, 0)$ such that $\alpha M_i + \beta M_j \succeq 0$, or
- there exists $a \in \mathbb{R}^n$ such that $M_i = ab_i^\top + b_i a^\top$.

**Theorem (Characterization of ROG for $|\mathcal{M}| = 2$)**

Suppose $\mathcal{M} = \{M_1, M_2\}$. Then sufficient condition above is also necessary.

Based on: **[Math. Oper. Res. 21]**, **[Tut. Oper. Res. 21]**

# ROG and ratios of quadratic functions

- $$\inf_{z\in\mathbb{R}^{n+1}} \left\{ \frac{z^\top M_{\text{obj}} z}{z^\top B z} : \begin{array}{l} z^\top M_i z \le 0,\ \forall i \in [m] \\ z^\top B z > 0 \\ z_{n+1}^2 = 1 \end{array} \right\}$$
$$= \inf_{\tilde{z}\in\mathbb{R}^{n+1}} \left\{ \tilde{z}^\top M_{\text{obj}} \tilde{z} : \begin{array}{l} \tilde{z}^\top M_i \tilde{z} \le 0,\ \forall i \in [m] \\ \tilde{z}^\top B \tilde{z} = 1 \\ \tilde{z}_{n+1}^2 > 0 \end{array} \right\}$$
$$\ge \inf_{Z\in\mathbb{S}_+^{n+1}} \left\{ \langle M_{\text{obj}}, Z\rangle : \begin{array}{l} \langle M_i, Z\rangle \le 0,\ \forall i \in [m] \\ \langle B, Z\rangle = 1 \end{array} \right\}$$

- Equality holds if $\mathcal{S}(\{M_1, \ldots, M_m\})$ is ROG (+ minor assumptions)
- Example: Regularized total least squares

Based on: **[Math. Oper. Res. 21]**, **[Tut. Oper. Res. 21]**

# Stackelberg prediction games with least squares losses

- True data $(x_i, \alpha_i, \beta_i)_{i=1}^m$
- Leader (learner) chooses $w \in \mathbb{R}^n$
- Follower (data provider) modifies $x_i \to \tilde{x}_i$ so that $\langle w, \tilde{x}_i \rangle \approx \beta_i$
- Leader has loss $(\langle w, \tilde{x}_i \rangle - \alpha_i)^2$
- $$\min_{w \in \mathbb{R}^n} \left\{ \sum_{i=1}^m (\langle w, \tilde{x}_i \rangle - \alpha_i)^2 : \tilde{x}_i \in \arg\min_{x \in \mathbb{R}^n} \gamma \|x - x_i\|^2 + (\langle w, x \rangle - \beta_i)^2 \right\}$$

Based on: **[ICML *under review*]**

# Stackelberg prediction games with least squares losses

- $$\begin{aligned}
&\min_{w\in\mathbb{R}^n}\left\{\sum_{i=1}^m \left(\langle w,\tilde{x}_i\rangle-\alpha_i\right)^2 :\ \tilde{x}_i\in\operatorname*{arg\,min}_{x\in\mathbb{R}^n}\gamma\left\|x-x_i\right\|^2+\left(\langle w,x\rangle-\beta_i\right)^2\right\}\\
&\quad=\min_{w\in\mathbb{R}^n}\left\{\left\|\tilde{X}^\top w-\alpha\right\|^2 :\ \tilde{X}=(\gamma I+ww^\top)^{-1}(\gamma X+w\beta^\top)\right\}\\
&\quad=\min_{w\in\mathbb{R}^n}\left\|\frac{\|w\|^2\beta+\gamma Xw}{\|w\|^2+\gamma}-\alpha\right\|^2\\
&\quad=\min_{w\in\mathbb{R}^n,t\in\mathbb{R}}\left\{\left\|\frac{t\beta+\gamma Xw}{t+\gamma}-\alpha\right\|^2 :\ t=\|w\|^2\right\}\\
&\quad=\min_{\tilde{w}\in\mathbb{R}^n,\tilde{t}\in\mathbb{R}}\left\{\left\|\frac{\tilde{t}}{2}\beta+\frac{\sqrt{\gamma}}{2}X\tilde{w}-\left(\alpha-\frac{\beta}{2}\right)\right\|^2 :\ \|\tilde{w}\|^2+\tilde{t}^2=1\right\}
\end{aligned}$$

Based on: **[ICML *under review*]**

# Diagonalizing QCQPs

## Definition

$\{A_i\} \subseteq \mathbb{S}^n$ is simultaneously diagonalizable via congruence (SDC) if there exists invertible $P \in \mathbb{R}^{n \times n}$ such that $P^\top A_i P$ is diagonal $\forall i$.

- Nice property because: SDP relaxation of diagonal QCQP is SOCP (faster), $\Gamma$ is polyhedral (better understanding of exactness)

## Goal

Most sets of matrices are not SDC, can we find other computationally variants of SDC and understand such properties?

Based on: **[Math. Prog. *under review*]**

# Diagonalizing QCQPs: variants of SDC

## Definition

$\{A_i\} \subseteq \mathbb{S}^n$ is almost SDC (ASDC) if for all $\epsilon > 0$, there exists $\|A_i' - A_i\| \leq \epsilon$ such that $\{A_i'\}$ is SDC.

- "Limit of SDC sets"

## Definition

$\{A_i\} \subseteq \mathbb{S}^n$ is $d$-restricted SDC ($d$-RSDC) if there exists $A_i' = \left(\begin{smallmatrix} A_i & * \\ * & * \end{smallmatrix}\right) \in \mathbb{S}^{n+d}$ such that $\{A_i'\}$ is SDC.

- "Restriction of SDC sets"

Based on: **[Math. Prog. *under review*]**

# Diagonalizing QCQPs: ASDC

**Theorem**

Let $\{A, B\} \subseteq \mathbb{S}^n$ and suppose $A$ invertible. Then $\{A, B\}$ is ASDC if and only if $A^{-1}B$ has real spectrum. (+ construction)

**Theorem**

Let $\{A, B\} \subseteq \mathbb{S}^n$. If $\{A, B\}$ is singular, then it is ASDC. (+ construction)

**Theorem**

Let $\{A, B, C\} \subseteq \mathbb{S}^n$ and suppose $A$ invertible. Then $\{A, B, C\}$ is ASDC if and only if $\left\{A^{-1}B, A^{-1}C\right\}$ commute and have real spectrum. (+ construction)

Based on: **[Math. Prog. *under review*]**

# Diagonalizing QCQPs: $d$-RSDC

**Theorem**

Let $\{A, B\} \subseteq \mathbb{S}^n$. If $A$ is invertible and $A^{-1}B$ has simple eigenvalues, then $\{A, B\}$ is $1$-RSDC. (+ construction)

- Condition holds generically

Based on: **[Math. Prog. *under review*]**

# “Fuzzy” spectral partitioning

- Connected graph $G = (V, E)$
- Vertex masses $\mu : V \to \mathbb{R}_{++}$ and edge weights $\kappa : E \to \mathbb{R}_{++}$
- Laplacian $L = D - A$ w.r.t. $\kappa$

**Theorem (Cheeger’s inequality)**

If $\mu_v = d_v$, then

$$\frac{\Phi^2}{2} \leq \lambda_2(L, M) \leq 2\Phi$$

- $\lambda_2(L, M)$ is first nontrivial generalized eigenvalue
- $\Phi$ is sparsest cut

Based on: **[APPROX 19]**

# “Fuzzy” spectral partitioning

- We define “Fuzzy cuts”

## Definition

$$\Psi \approx \min_{A,B} \left\{ \frac{\kappa_{\text{eff}}(A,B)}{\min\left(\mu(A), \mu(B)\right)},\ A, B \neq \varnothing,\ A \cap B = \varnothing \right\}$$

- $\Phi$ must partition, $\Psi$ may leave out. $\Psi = \Phi$ if $A, B$ is a partition.

## Theorem

$$\frac{\Psi}{4} \leq \lambda_2(L, M) \leq \Psi$$

Based on: **[APPROX 19]**

# Stable Euclidean $k$-means

- $k$-means clustering: $\{x_1, \ldots, x_n\} \subseteq \mathbb{R}^d$
- Suppose there exist true clustering that is unique optimum even if for all $i$, $x_i \mapsto x_i' \in B(x_i, \epsilon)$

## Theorem

**Two clusters.** There exists $c \geq 1$ such that for any fixed $\epsilon > 0$, we can recover true clustering in time $d \cdot n^{O(\epsilon^{-c})}$.

Additional results for $\geq 3$ clusters given an additional "separation" assumption

Based on: **[NeurIPS 17]**