

# 11. Interior-point methods

- inequality constrained minimization
- logarithmic barrier function and central path
- barrier method
- feasibility and phase I methods
- complexity analysis via self-concordance
- second-order cone and semidefinite programming

# Inequality constrained minimization

$$\begin{aligned} & \text{minimize} && f_0(x) \\ & \text{subject to} && f_i(x) \leq 0, \quad i = 1, \dots, m \\ & && Ax = b \end{aligned} \tag{1}$$

- $f_i$  convex, twice continuously differentiable
- $A \in \mathbf{R}^{p \times n}$  with  $\mathbf{rank} A = p$
- we assume  $p^\star$  is finite and attained
- we assume the problem is strictly feasible: there exists  $\tilde{x}$  with

$$\tilde{x} \in \text{dom } f_0, \quad f_i(\tilde{x}) < 0, \quad i = 1, \dots, m, \quad A\tilde{x} = b$$

hence, strong duality holds and dual optimum is attained

# Unconstrained (or equality-constrained) approximation

- write (1) as problem without inequality constraints:

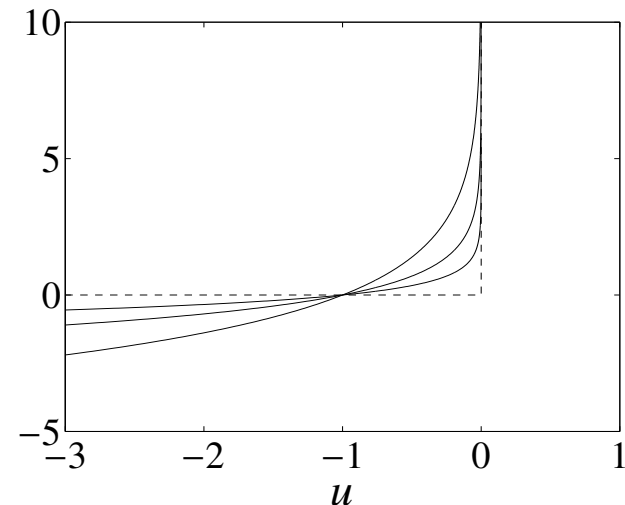
$$\begin{aligned} & \text{minimize} && f_0(x) + \sum_{i=1}^m h(f_i(x)) \\ & \text{subject to} && Ax = b \end{aligned}$$

where  $h$  is indicator function of  $\mathbf{R}_-$ :  $h(u) = 0$  if  $u \leq 0$  and  $h(u) = \infty$  otherwise

- approximate indicator function by logarithmic barrier:

$$\begin{aligned} & \text{minimize} && f_0(x) - (1/t) \sum_{i=1}^m \log(-f_i(x)) \\ & \text{subject to} && Ax = b \end{aligned}$$

- an equality constrained problem
- $t > 0$ , approximation improves as  $t \rightarrow \infty$



# Logarithmic barrier function

$$\phi(x) = -\sum_{i=1}^m \log(-f_i(x)), \quad \text{dom } \phi = \{x \mid f_1(x) < 0, \dots, f_m(x) < 0\}$$

- a convex function (follows from composition rules)
- twice continuously differentiable, with derivatives

$$\nabla \phi(x) = \sum_{i=1}^m \frac{1}{-f_i(x)} \nabla f_i(x)$$

$$\nabla^2 \phi(x) = \sum_{i=1}^m \frac{1}{f_i(x)^2} \nabla f_i(x) \nabla f_i(x)^T + \sum_{i=1}^m \frac{1}{-f_i(x)} \nabla^2 f_i(x)$$

# Central path

- for  $t > 0$ , define  $x^\star(t)$  as the solution of the *centering problem*

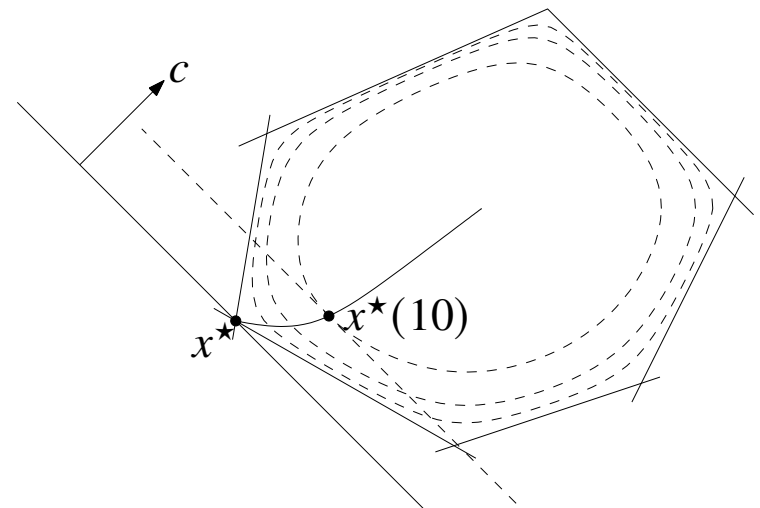
$$\begin{aligned} & \text{minimize} && t f_0(x) + \phi(x) \\ & \text{subject to} && Ax = b \end{aligned}$$

(for now, assume  $x^\star(t)$  exists and is unique for each  $t > 0$ )

- the set  $\{x^\star(t) \mid t > 0\}$  is called the *central path*

**Example:** central path for an LP

$$\begin{aligned} & \text{minimize} && c^T x \\ & \text{subject to} && a_i^T x \leq b_i, \quad i = 1, \dots, 6 \end{aligned}$$



hyperplane  $c^T x = c^T x^\star(t)$  is tangent to level curve of  $\phi$  through  $x^\star(t)$

## Dual points on central path

- optimality condition for centering problem:  $Ax = b$  and there exists a  $w$  such that

$$\begin{aligned} 0 &= t \nabla f_0(x) + \nabla \phi(x) + A^T w \\ &= t \nabla f_0(x) + \sum_{i=1}^m \frac{1}{-f_i(x)} \nabla f_i(x) + A^T w \end{aligned}$$

- point on central path  $x^\star(t)$  minimizes the Lagrangian of the original problem

$$L(x, \lambda, \nu) = f_0(x) + \sum_{i=1}^m \lambda_i f_i(x) + \nu^T (Ax - b)$$

for  $\lambda, \nu$  given by

$$\lambda_i^\star(t) = \frac{1}{-t f_i(x^\star(t))}, \quad i = 1, \dots, m, \quad \nu^\star(t) = w/t$$

centering gives a strictly primal feasible  $x^\star(t)$  and a dual feasible  $\lambda^\star(t), \nu^\star(t)$

## Duality gap on central path

- value of dual objective function at  $\lambda^\star(t)$ ,  $\nu^\star(t)$  is

$$\begin{aligned}g(\lambda^\star(t), \nu^\star(t)) &= \inf_x L(x, \lambda^\star(t), \nu^\star(t)) \\&= L(x^\star(t), \lambda^\star(t), \nu^\star(t)) \\&= f_0(x^\star) + \sum_{i=1}^m \lambda_i^\star(t) f_i(x^\star(t)) + \nu^{\star T} (Ax^\star - b) \\&= f_0(x^\star(t)) - \frac{m}{t}\end{aligned}$$

- this confirms the intuitive idea that  $f_0(x^\star(t)) \rightarrow p^\star$  if  $t \rightarrow \infty$ :

$$f_0(x^\star(t)) - p^\star \leq \frac{m}{t}$$

# Interpretation via KKT conditions

$x = x^*(t)$ ,  $\lambda = \lambda^*(t)$ ,  $\nu = \nu^*(t)$  satisfy

1. primal constraints:  $f_i(x) \leq 0$ ,  $i = 1, \dots, m$ ,  $Ax = b$

2. dual inequality:  $\lambda \succeq 0$

3. approximate complementary slackness:

$$\lambda_i f_i(x) = -\frac{1}{t}, \quad i = 1, \dots, m$$

4. gradient of Lagrangian with respect to  $x$  vanishes:

$$\nabla f_0(x) + \sum_{i=1}^m \lambda_i \nabla f_i(x) + A^T \nu = 0$$

difference with KKT conditions is that condition 3 replaces  $\lambda_i f_i(x) = 0$

# Force field interpretation

**Centering problem** (for problem with no equality constraints)

$$\text{minimize } t f_0(x) - \sum_{i=1}^m \log(-f_i(x))$$

## Force field interpretation

- $t f_0(x)$  is potential of force field

$$F_0(x) = -t \nabla f_0(x)$$

- $-\log(-f_i(x))$  is potential of force field

$$F_i(x) = (1/f_i(x)) \nabla f_i(x)$$

- the forces balance at  $x^\star(t)$ :

$$F_0(x^\star(t)) + \sum_{i=1}^m F_i(x^\star(t)) = 0$$

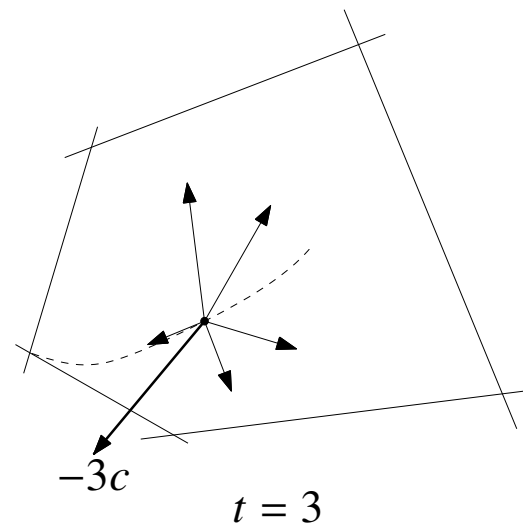
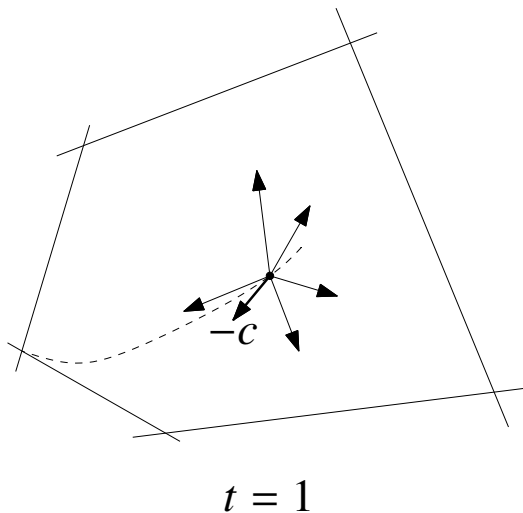
# Example

$$\begin{aligned} & \text{minimize} && c^T x \\ & \text{subject to} && a_i^T x \leq b_i, \quad i = 1, \dots, m \end{aligned}$$

- objective force field is constant:  $F_0(x) = -tc$
- constraint force field decays as inverse distance to constraint hyperplane:

$$F_i(x) = \frac{-a_i}{b_i - a_i^T x}, \quad \|F_i(x)\|_2 = \frac{1}{d(x, \mathcal{H}_i)}$$

where  $d(x, \mathcal{H}_i)$  is distance of  $x$  to hyperplane  $\mathcal{H}_i = \{x \mid a_i^T x = b_i\}$



# Barrier method

given: strictly feasible  $x$ ,  $t := t^{(0)} > 0$ ,  $\mu > 1$ , tolerance  $\epsilon > 0$

repeat

1. *centering step*: compute  $x^\star(t)$  by minimizing  $t f_0(x) + \phi(x)$  subject to  $Ax = b$
2. *update*:  $x := x^\star(t)$
3. *stopping criterion*: quit if  $m/t < \epsilon$
4. *increase  $t$* :  $t := \mu t$

- terminates with strictly feasible point that satisfies  $f_0(x) - p^\star \leq m/t < \epsilon$
- centering is usually done using Newton's method, starting at current  $x$
- an outer iteration loop (steps 1–4) and an inner (Newton) iteration loop (step 1)
- choice of  $\mu$  involves trade-off between number of outer and inner iterations
- typical values of  $\mu$  are 10–20
- several heuristics exist for choosing  $t^{(0)}$

# Convergence analysis

**Number of outer (centering) iterations:** exactly

$$\left\lceil \frac{\log(m/(\epsilon t^{(0)}))}{\log \mu} \right\rceil$$

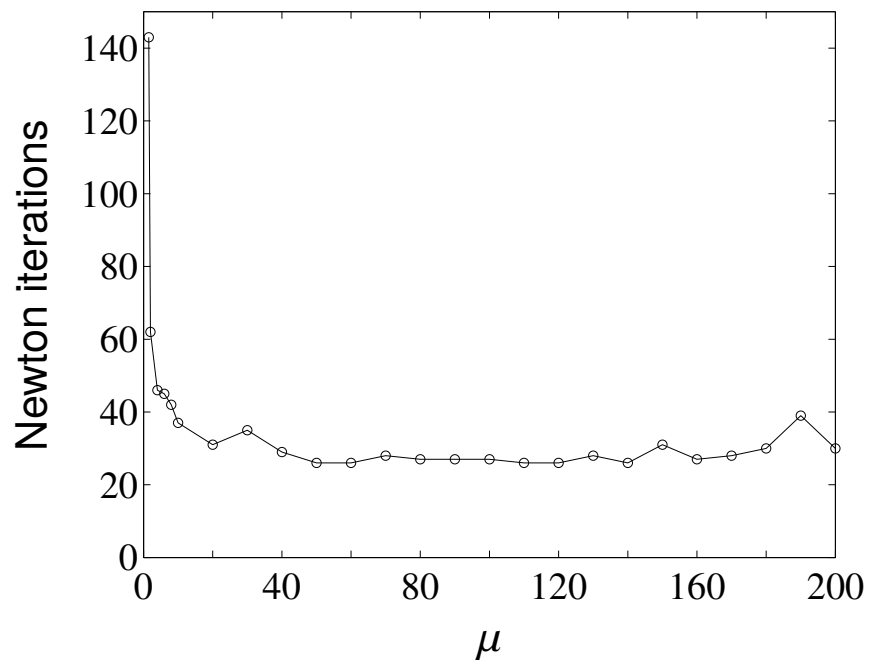
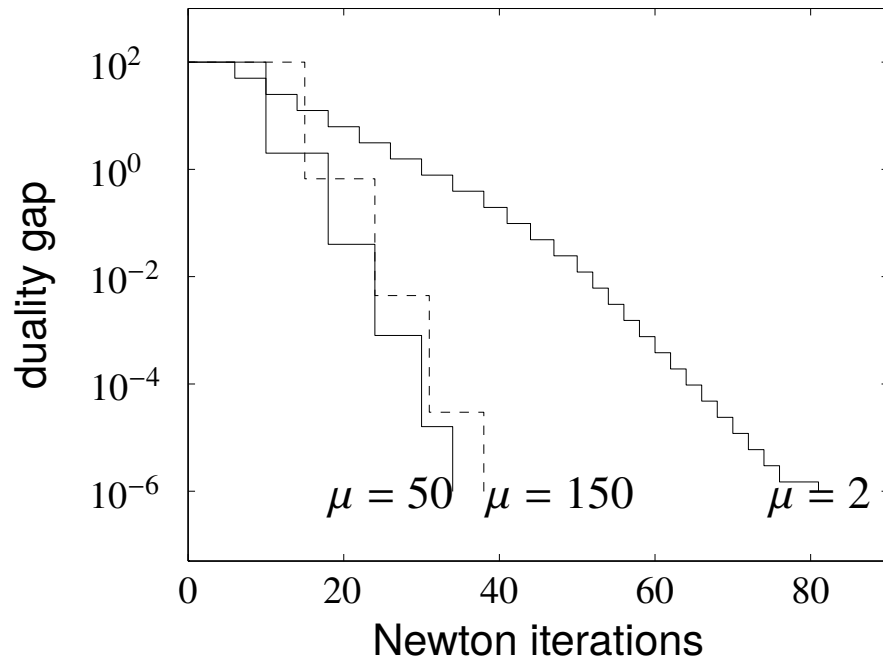
plus the initial centering step (to compute  $x^\star(t^{(0)})$ )

**Centering problem:** see convergence analysis of Newton's method

- $t f_0 + \phi$  must have closed sublevel sets for  $t \geq t^{(0)}$
- classical analysis requires strong convexity, Lipschitz continuity of Hessian
- analysis via self-concordance requires self-concordance of  $t f_0 + \phi$
- the additional assumptions also guarantee that solution exists and is unique

# Example: inequality form LP

LP with  $m = 100$  inequalities,  $n = 50$  variables

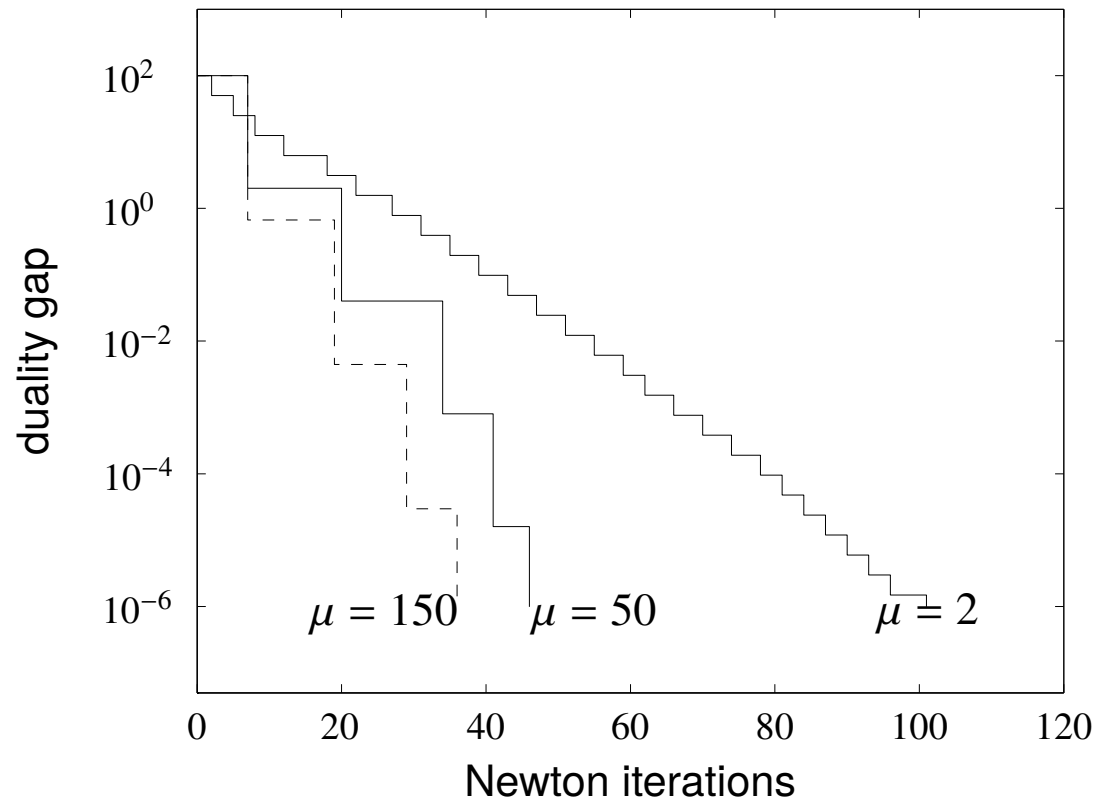


- starts with  $x$  on central path ( $t^{(0)} = 1$ , duality gap 100)
- terminates when  $t = 10^8$  (gap  $10^{-6}$ )
- centering uses Newton's method with backtracking
- total number of Newton iterations not very sensitive for  $\mu \geq 10$

# Example: geometric program

GP with  $m = 100$  inequalities and  $n = 50$  variables

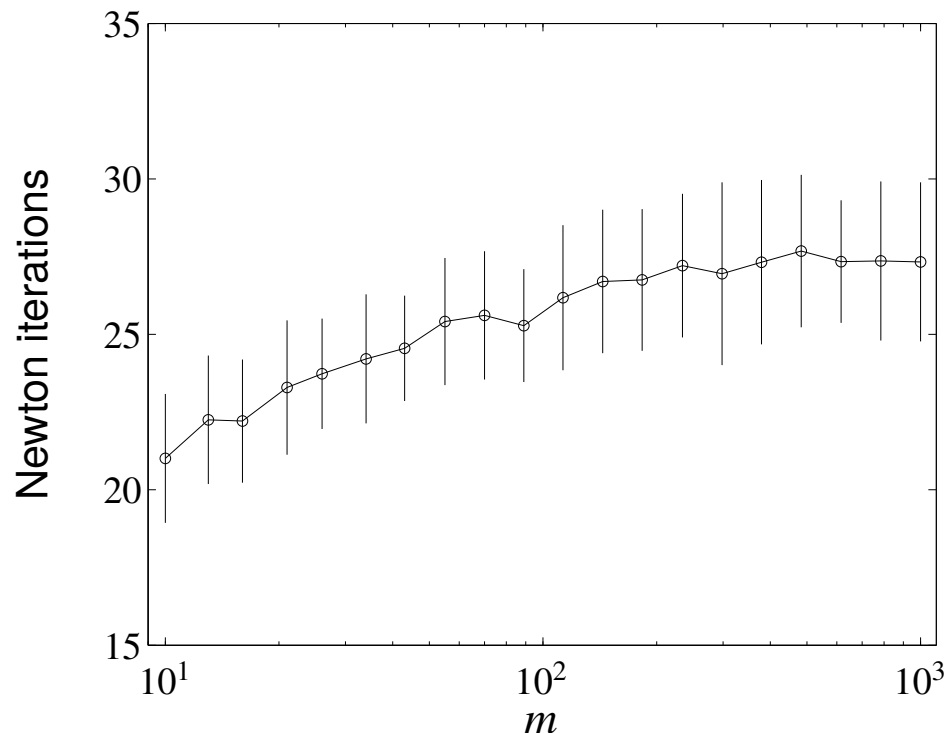
$$\begin{aligned} \text{minimize} \quad & \log\left(\sum_{k=1}^5 \exp(a_{0k}^T x + b_{0k})\right) \\ \text{subject to} \quad & \log\left(\sum_{k=1}^5 \exp(a_{ik}^T x + b_{ik})\right) \leq 0, \quad i = 1, \dots, m \end{aligned}$$



## Example: family of standard LPs

$$\begin{array}{ll} \text{minimize} & c^T x \\ \text{subject to} & Ax = b, \quad x \succeq 0 \end{array}$$

- $A \in \mathbf{R}^{m \times 2m}$  with  $m = 10, \dots, 1000$
- for each  $m$ , solve 100 randomly generated instances



number of iterations grows very slowly as  $m$  ranges over a 100 : 1 ratio

# Feasibility and phase I methods

**Phase I:** computes a starting point  $x$  that is strictly feasible for

$$f_i(x) \leq 0, \quad i = 1, \dots, m, \quad Ax = b \quad (2)$$

## Basic phase I method

$$\begin{array}{ll} \text{minimize (over } x, s) & s \\ \text{subject to} & f_i(x) \leq s, \quad i = 1, \dots, m \\ & Ax = b \end{array} \quad (3)$$

- problem (3) is strictly feasible: take any  $x, s$  that satisfies

$$x \in \text{dom } f_i, \quad i = 1, \dots, m, \quad Ax = b, \quad s > \max_i f_i(x)$$

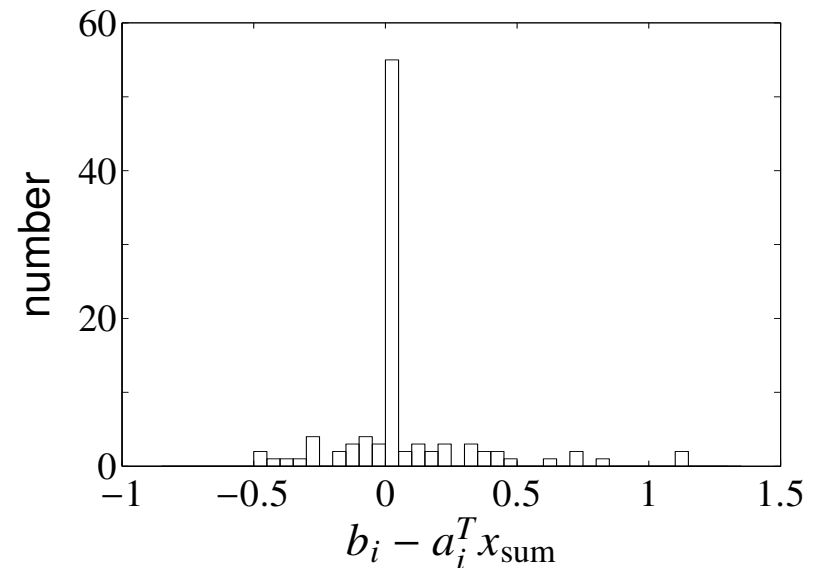
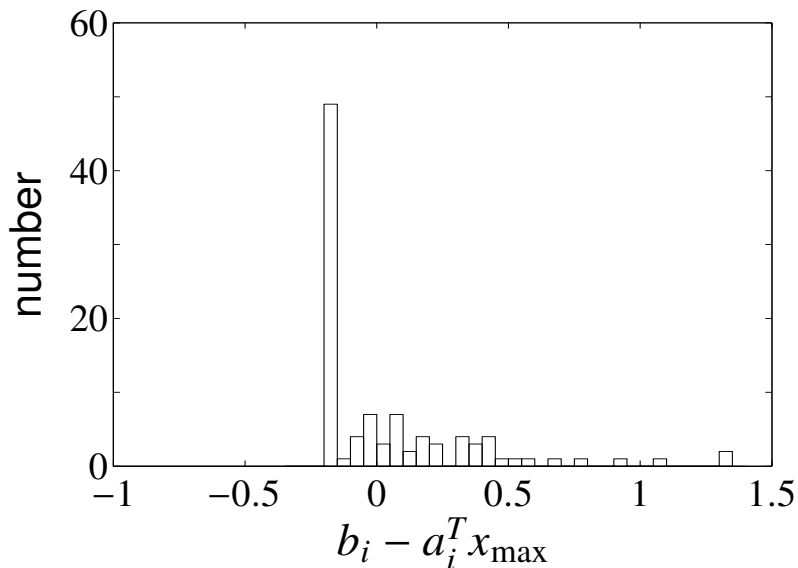
- if  $x, s$  are feasible for (3) with  $s < 0$ , then  $x$  is strictly feasible for (2)
- if optimal value  $\bar{p}^*$  of (3) is positive, then problem (2) is infeasible
- if  $\bar{p}^* = 0$  and attained, then problem (2) is feasible (but not strictly)
- if  $\bar{p}^* = 0$  and not attained, then problem (2) is infeasible

# Sum of infeasibilities phase I method

$$\begin{aligned} & \text{minimize} && \mathbf{1}^T s \\ & \text{subject to} && s \geq 0, \quad f_i(x) \leq s_i, \quad i = 1, \dots, m \\ & && Ax = b \end{aligned}$$

for infeasible problem, will find  $x$  that satisfies many more inequalities than (3)

**Example** (infeasible set of 100 linear inequalities in 50 variables)



- left: basic phase I solution; satisfies 39 inequalities
- right: sum of infeasibilities phase I solution; satisfies 79 inequalities

# Complexity analysis via self-concordance

same assumptions as on page 11.2, plus:

- sublevel sets (of  $f_0$ , on the feasible set) are bounded
- $t f_0 + \phi$  is self-concordant with closed sublevel sets
- second condition holds for LP, QP, QCQP
- may require reformulating the problem, *e.g.*,

$$\begin{array}{ll} \text{minimize} & \sum_{i=1}^n x_i \log x_i \\ \text{subject to} & Fx \preceq g \end{array} \quad \longrightarrow \quad \begin{array}{ll} \text{minimize} & \sum_{i=1}^n x_i \log x_i \\ \text{subject to} & Fx \preceq g, \quad x \succeq 0 \end{array}$$

- assumptions are needed for complexity analysis, not to run the barrier method

## Newton iterations per centering step

bound on effort of computing  $x^+ = x^\star(\mu t)$  starting at  $x = x^\star(t)$ :

$$\# \text{Newton iterations} \leq \frac{\mu t f_0(x) + \phi(x) - \mu t f_0(x^+) - \phi(x^+)}{\gamma} + c \quad (4)$$

- $\gamma, c$  are constants (depend only on algorithm parameters); see page 9.33
- upper bound on first term follows from duality:

$$\begin{aligned} & \mu t f_0(x) + \phi(x) - \mu t f_0(x^+) - \phi(x^+) \\ &= \mu t f_0(x) - \mu t f_0(x^+) + \sum_{i=1}^m \log(-\mu t \lambda_i f_i(x^+)) - m \log \mu \\ &\leq \mu t f_0(x) - \mu t f_0(x^+) - \mu t \sum_{i=1}^m \lambda_i f_i(x^+) - m - m \log \mu \\ &\leq \mu t f_0(x) - \mu t g(\lambda, \nu) - m - m \log \mu \\ &= m(\mu - 1 - \log \mu) \end{aligned}$$

where  $\lambda_i = \lambda_i^\star(t) = -1/(t f_i(x^\star(t)))$

# Total number of Newton iterations

- we exclude first centering step on page 11.11, assume we start at  $x^\star(t^{(0)})$
- bound on Newton iterations is number of outer iterations times (4)

$$\# \text{Newton iterations} \leq N = \left\lceil \frac{\log(m/(t^{(0)}\epsilon))}{\log \mu} \right\rceil \left( \frac{m(\mu - 1 - \log \mu)}{\gamma} + c \right)$$

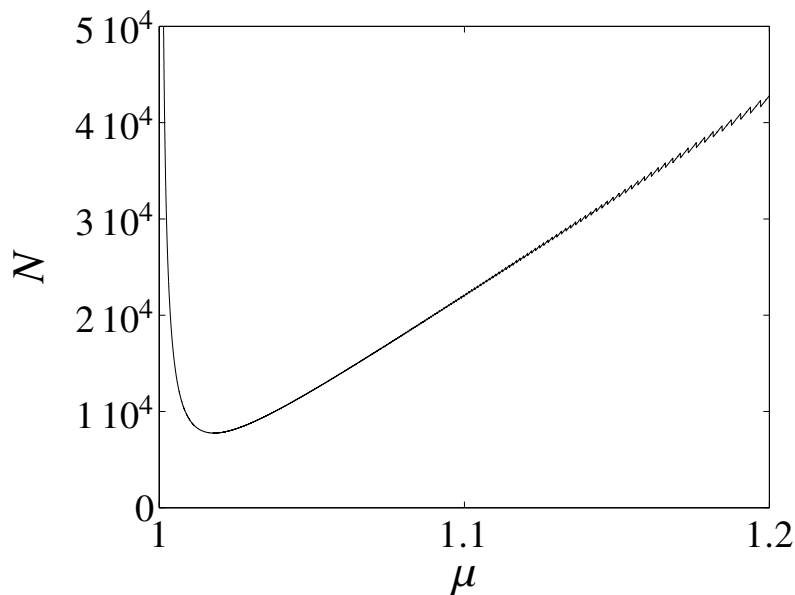


figure shows  $N$  for typical values of  $\gamma, c$ ,

$$m = 100, \quad \frac{m}{t^{(0)}\epsilon} = 10^5$$

- confirms trade-off in choice of  $\mu$
- in practice, #iterations is in the tens and not very sensitive for  $\mu \geq 10$

# Polynomial-time complexity of barrier method

- for  $\mu = 1 + 1/\sqrt{m}$ :

$$N = O\left(\sqrt{m} \log\left(\frac{m/t^{(0)}}{\epsilon}\right)\right)$$

- number of Newton iterations for fixed gap reduction is  $O(\sqrt{m})$
- multiply with cost of one Newton iteration to get bound on number of flops
- this choice of  $\mu$  optimizes worst-case complexity
- in practice we choose  $\mu$  fixed ( $\mu = 10, \dots, 20$ )

# Second-order cone programming

$$\begin{array}{ll} \text{minimize} & f^T x \\ \text{subject to} & \|A_i x + b_i\|_2 \leq c_i^T x + d_i, \quad i = 1, \dots, m \end{array}$$

- constraint functions are not differentiable
- barrier method for second-order cone programming uses barrier function

$$\begin{aligned} \phi(x) &= - \sum_{i=1}^m \log((c_i^T x + d_i)^2 - \|A_i x + b_i\|_2^2) \\ &= - \sum_{i=1}^m \log(c_i^T x + d_i) - \sum_{i=1}^m \log\left(c_i^T x + d_i - \frac{\|A_i x + b_i\|_2^2}{c_i^T x + d_i}\right) \end{aligned}$$

- equivalent to standard barrier method for reformulation with  $2m$  inequalities

$$\begin{array}{ll} \text{minimize} & f^T x \\ \text{subject to} & \frac{\|A_i x + b_i\|_2^2}{c_i^T x + d_i} \leq c_i^T x + d_i, \quad i = 1, \dots, m \\ & c_i^T x + d_i \geq 0, \quad i = 1, \dots, m \end{array}$$

# Semidefinite programming

**Primal and dual SDP** (with  $F_1, \dots, F_n, G \in \mathbf{S}^m$ )

$$\begin{array}{ll} \text{minimize} & c^T x \\ \text{subject to} & \sum_{i=1}^n x_i F_i \preceq G \end{array} \qquad \begin{array}{ll} \text{maximize} & -\text{tr}(GZ) \\ \text{subject to} & \text{tr}(F_i Z) + c_i = 0, \quad i = 1, \dots, n \\ & Z \succeq 0 \end{array}$$

## Logarithmic barrier

$$\phi(x) = -\log \det F(x), \quad \text{where } F(x) = G - \sum_{i=1}^n x_i F_i$$

- a convex differentiable function, with domain  $\{x \mid F(x) \succ 0\}$
- gradient and Hessian are

$$\nabla \phi(x)_i = \text{tr}(F_i F(x)^{-1}), \quad \nabla^2 \phi(x)_{ij} = \text{tr}(F_i (F(x)^{-1} F_j F(x)^{-1})),$$

for  $i, j = 1, \dots, n$

# Central path

points on central path  $x^\star(t)$  for  $t > 0$  are minimizers of  $tc^T x + \phi(x)$

- optimality condition for centering problem:

$$0 = tc_i + \nabla\phi(x)_i = tc_i + \text{tr}(F_i F(x)^{-1}), \quad i = 1, \dots, n$$

- dual feasible point on central path:

$$Z^\star(t) = \frac{1}{t} F(x^\star(t))^{-1}$$

- corresponding duality gap:

$$\begin{aligned} c^T x^\star(t) + \text{tr}(GZ^\star(t)) &= \text{tr}\left(\left(-\sum_{i=1}^n x_i^\star(t) F_i + G\right) Z^\star(t)\right) \\ &= \text{tr}(F(x^\star(t)) Z^\star(t)) \\ &= m/t \end{aligned}$$

# Barrier method for semidefinite programming

given: strictly feasible  $x$ ,  $t := t^{(0)} > 0$ ,  $\mu > 1$ , tolerance  $\epsilon > 0$

repeat

1. *centering step*: compute  $x^\star(t)$  by minimizing  $tc^T x + \phi(x)$
2. *update*:  $x := x^\star(t)$
3. *stopping criterion*: quit if  $m/t < \epsilon$
4. *increase  $t$* :  $t := \mu t$

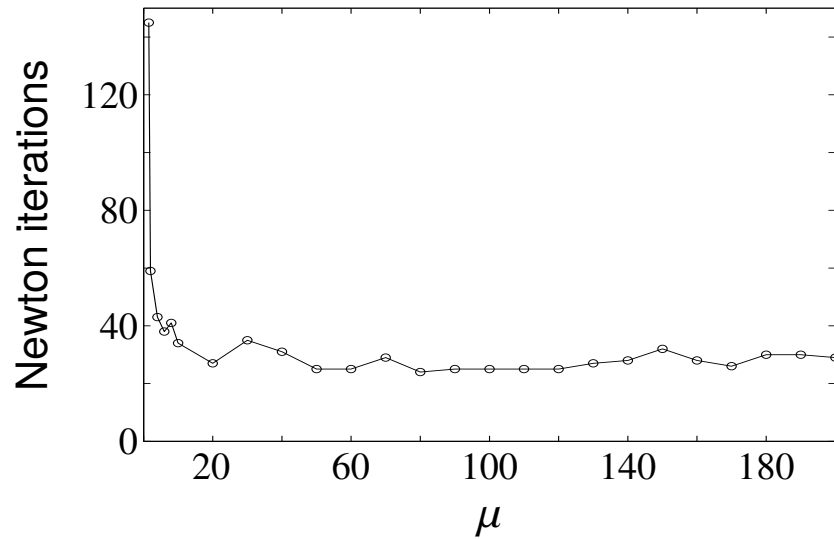
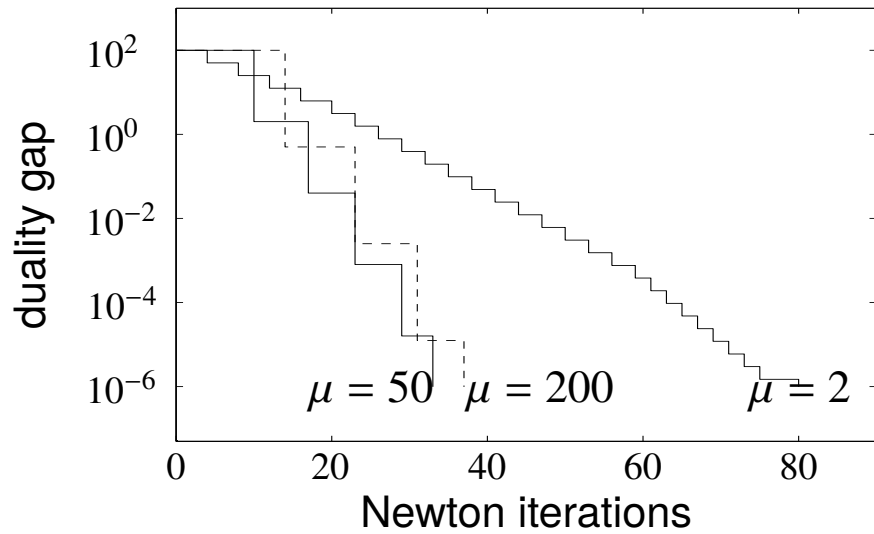
- number of outer iterations:

$$\left\lceil \frac{\log(m/(\epsilon t^{(0)}))}{\log \mu} \right\rceil$$

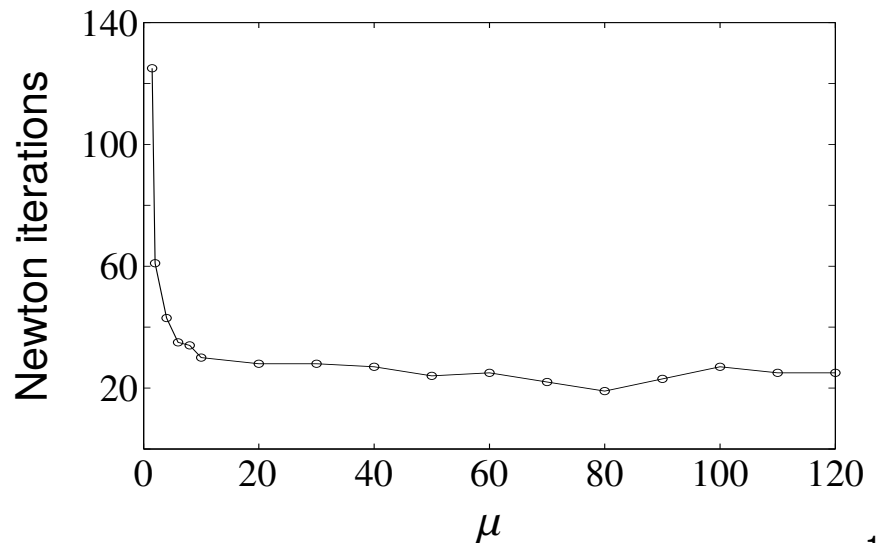
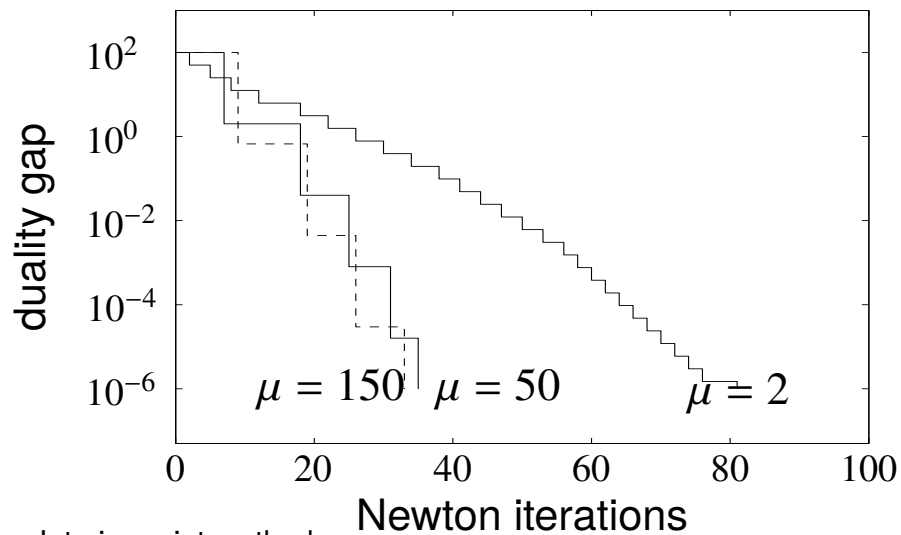
- complexity analysis via self-concordance also applies to SDP

# Examples

## Second-order cone program (50 variables, 50 SOC constraints in $\mathbf{R}^6$ )



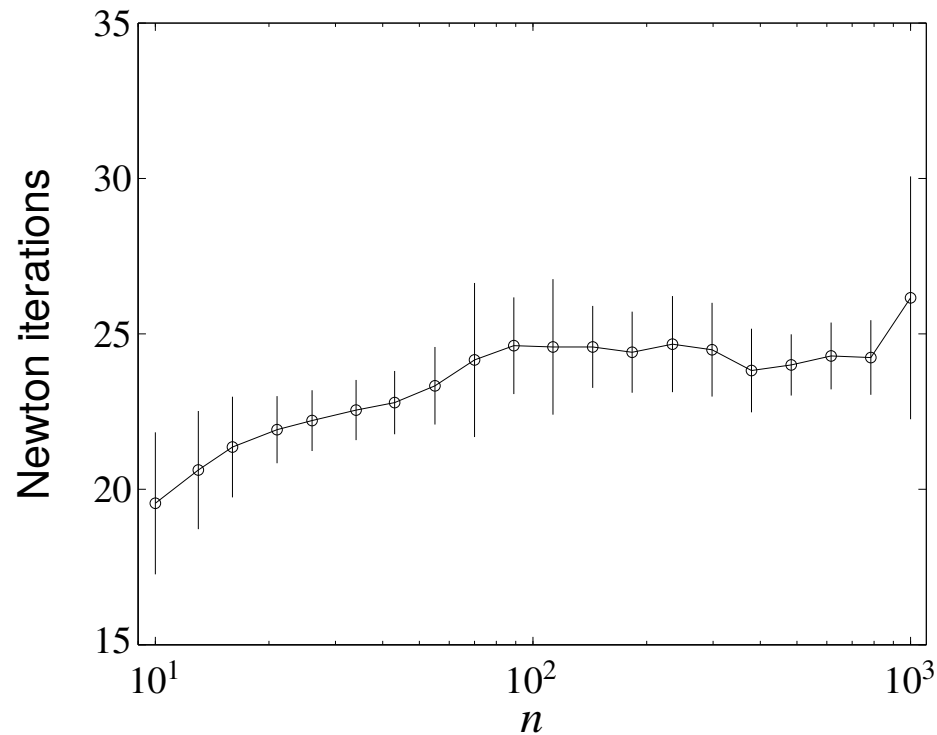
## Semidefinite program (100 variables, constraint in $\mathbf{S}^{100}$ )



## Family of SDPs ( $A \in \mathbf{S}^n, x \in \mathbf{R}^n$ )

$$\begin{array}{ll} \text{minimize} & \mathbf{1}^T x \\ \text{subject to} & A + \mathbf{diag}(x) \succeq 0 \end{array}$$

$n = 10, \dots, 1000$ , for each  $n$  solve 100 randomly generated instances



# Primal-dual interior-point methods

more efficient than barrier method when high accuracy is needed

- update primal and dual variables at each iteration
- no distinction between inner and outer iterations
- often exhibit superlinear asymptotic convergence
- steps can be interpreted as Newton iterates for modified KKT conditions
- can start at infeasible points
- cost per iteration same as barrier method