An Active Set SQP Algorithm for Nonlinear Programming with Inexact Subproblem Solutions

- T. C. Johnson

- F. E. Curtis - D. P. Robinson - A. Wächter

Johns Hopkins, Lehigh, Northwestern

GOAL

$$\text{Solve:} \quad \begin{array}{ll} \min & f(x) \\ \text{s.t.} & c(x) \leq 0\ldots \end{array}$$

GOAL

Solve: $$\begin{aligned} \min \quad & f(x) \\ \text{s.t.} \quad & c(x) \leq 0 \ldots \end{aligned}$$

with **inexact** *subproblem* solutions

Problem $\mapsto$ Subproblems

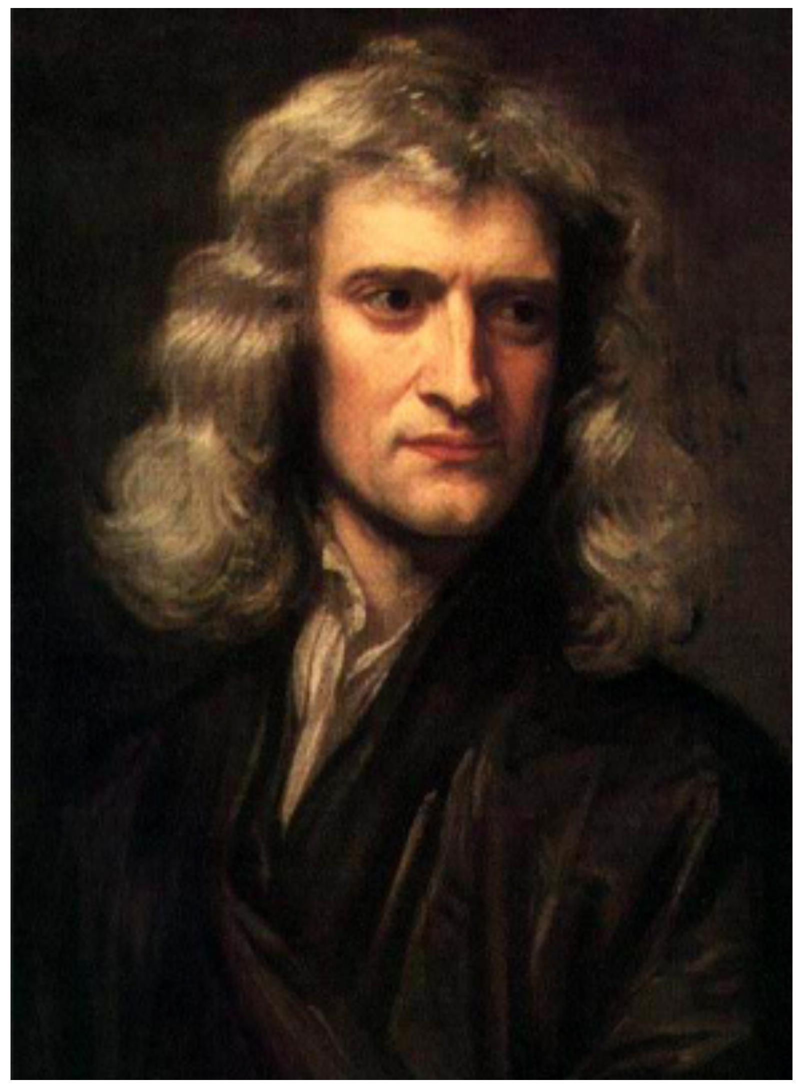
Problem $\mapsto$ Subproblem**s**

Problem:

$$F(z) = 0$$

Problem: Subproblem:

$$F(z) = 0 \longrightarrow F'(z)d = -F(z)$$



Problem: Subproblem:

$$F(z) = 0 \longrightarrow F'(z)d = -F(z)$$

Direct Methods:

Problem: Subproblem:

$$F(z) = 0 \longrightarrow F'(z)d = -F(z)$$

Direct Methods: $\|F(z) + F'(z)d\| = 0$

Problem: Subproblem:

$$F(z) = 0 \longrightarrow F'(z)d = -F(z)$$

Direct Methods: $$\|F(z) + F'(z)d\| = 0$$

$$\|F(z) + F(z)d\| \leq \eta \|F(z)\|$$

$$0 < \eta < 1$$

Problem: Subproblem:

$$F(z) = 0 \longrightarrow F'(z)d = -F(z)$$

Direct Methods: $\|F(z) + F'(z)d\| = 0$

??: $\|F(z) + F(z)d\| \leq \eta\|F(z)\|$

$0 < \eta < 1$

Problem: Subproblem:

$$F(z) = 0 \longrightarrow F'(z)d = -F(z)$$

Direct Methods: $\|F(z) + F'(z)d\| = 0$

??: $\|F(z) + F(z)d\| \leq \eta \|F(z)\|$

$0 < \eta < 1$

Dembo et al 82

Problem: Subproblem:

$$F(z) = 0 \longrightarrow F'(z)d = -F(z)$$

Direct Methods: $\|F(z) + F'(z)d\| = 0$

(any): $\|F(z) + F(z)d\| \leq \eta \|F(z)\|$

$0 < \eta < 1$

Dembo et al 82

Inexact "trick" for Constrained Optimization

Equality Constrained:

Problem:

$$\begin{aligned} \min \quad & f(x) \\ \text{s.t.} \quad & c(x) = 0 \end{aligned}$$

Equality Constrained:

Problem:

$$\begin{array}{ll} \min & f(x) \\ \text{s.t.} & c(x) = 0 \end{array}$$

Subproblem:

$$\begin{bmatrix} W & A^T \\ A & 0 \end{bmatrix} \begin{pmatrix} d \\ \lambda^+ \end{pmatrix} = - \begin{pmatrix} g \\ c \end{pmatrix}$$

Inequality Constrained Interior Point:

Problem:

$$\begin{aligned} \min \quad & f(x) \\ \text{s.t.} \quad & c(x) = 0 \\ & x \geq 0 \end{aligned}$$

Inequality Constrained Interior Point:

Problem:

$$\min \quad f(x) - \sum \ln x$$

$$\text{s.t.} \quad c(x) = 0$$

Inequality Constrained Interior Point:

Problem:

$$\begin{aligned} \min \quad & f(x) - \sum \ln x \\ \text{s.t.} \quad & c(x) = 0 \end{aligned}$$

Subproblem:

$$\begin{bmatrix} W & A^T \\ A & 0 \end{bmatrix} \begin{pmatrix} d \\ \lambda^+ \end{pmatrix} = - \begin{pmatrix} g \\ c \end{pmatrix}$$

Inequality Constrained SQP:

Problem:

$$\begin{aligned} \min \quad & f(x) \\ \text{s.t.} \quad & c(x) = 0 \\ & x \geq 0 \end{aligned}$$

Inequality Constrained SQP:

Problem:

$$\begin{aligned} \min \quad & f(x) \\ \text{s.t.} \quad & c(x) = 0 \\ & x \geq 0 \end{aligned}$$

Subproblem:

$$\begin{aligned} \min \quad & 1/2 d^T W d + g^T d \\ \text{s.t.} \quad & Ad + c = 0 \\ & x + d \geq 0 \end{aligned}$$

Inequality Constrained SQP:

Problem:

$$\begin{aligned} \min \quad & f(x) \\ \text{s.t.} \quad & c(x) = 0 \\ & x \geq 0 \end{aligned}$$

Subproblem:

$$\begin{aligned} \min \quad & 1/2 d^T W d + g^T d \\ \text{s.t.} \quad & Ad + c = 0 \\ & x + d \geq 0 \end{aligned}$$

$$\begin{aligned} \min \quad & 1/2 d^T W d + g^T d \\ & - \sum \ln(x + d) \\ \text{s.t.} \quad & Ad + c = 0 \end{aligned}$$

Inequality Constrained SQP:

Problem:

$$\begin{aligned} \min \quad & f(x) \\ \text{s.t.} \quad & c(x) = 0 \\ & x \geq 0 \end{aligned}$$

Subproblem:

$$\begin{aligned} \min \quad & 1/2 d^T W d + g^T d \\ \text{s.t.} \quad & Ad + c = 0 \\ & x + d \geq 0 \end{aligned}$$

$$\begin{bmatrix} W & A^T \\ A & 0 \end{bmatrix} \begin{pmatrix} d \\ \lambda^+ \end{pmatrix} = - \begin{pmatrix} g \\ c \end{pmatrix}$$

Inequality Constrained SQP:

Problem:

$$\begin{aligned} \min \quad & f(x) \\ \text{s.t.} \quad & c(x) = 0 \\ & x \geq 0 \end{aligned}$$

Subproblem:

$$\begin{aligned} \min \quad & 1/2 d^T W d + g^T d \\ \text{s.t.} \quad & Ad + c = 0 \\ & x + d \geq 0 \end{aligned}$$

Inequality Constrained SQP:

Problem:

$$\begin{aligned} \min \quad & f(x) \\ \text{s.t.} \quad & c(x) = 0 \\ & x \geq 0 \end{aligned}$$

Subproblem:

$$\begin{aligned} \min \quad & 1/2 d^T W d + g^T d \\ \text{s.t.} \quad & Ad + c = 0 \\ & x + d \geq 0 \end{aligned}$$

Staying SQP critical for:

Inequality Constrained SQP:

Problem:

$$\begin{aligned} \min \quad & f(x) \\ \text{s.t.} \quad & c(x) = 0 \\ & x \geq 0 \end{aligned}$$

Subproblem:

$$\begin{aligned} \min \quad & 1/2 d^T W d + g^T d \\ \text{s.t.} \quad & Ad + c = 0 \\ & x + d \geq 0 \end{aligned}$$

Staying SQP critical for:

- Mixed integer nonlinear programming

Inequality Constrained SQP:

Problem:

$$\begin{aligned} \min \quad & f(x) \\ \text{s.t.} \quad & c(x) = 0 \\ & x \geq 0 \end{aligned}$$

Subproblem:

$$\begin{aligned} \min \quad & 1/2 d^T W d + g^T d \\ \text{s.t.} \quad & Ad + c = 0 \\ & x + d \geq 0 \end{aligned}$$

Staying SQP critical for:

- Mixed integer nonlinear programming
- Nonlinear model predictive control

Inequality Constrained SQP:

Problem:

$$\begin{aligned} \min \quad & f(x) \\ \text{s.t.} \quad & c(x) = 0 \\ & x \geq 0 \end{aligned}$$

Subproblem:

$$\begin{aligned} \min \quad & 1/2 d^T W d + g^T d \\ \text{s.t.} \quad & Ad + c = 0 \\ & x + d \geq 0 \end{aligned}$$

Staying SQP critical for:

- Mixed integer nonlinear programming
- Nonlinear model predictive control
- ...

Inexact methods:

- Reduce residual in subproblem
- Are lazy
- Exploit problem structure

Plan

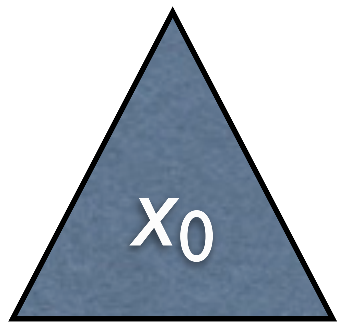
1. Review SQP & Sℓ_1QP
2. Inexactness and Penalty Steering
3. Theory & Numerical Results

SQP REVIEW

SQP & Line Search

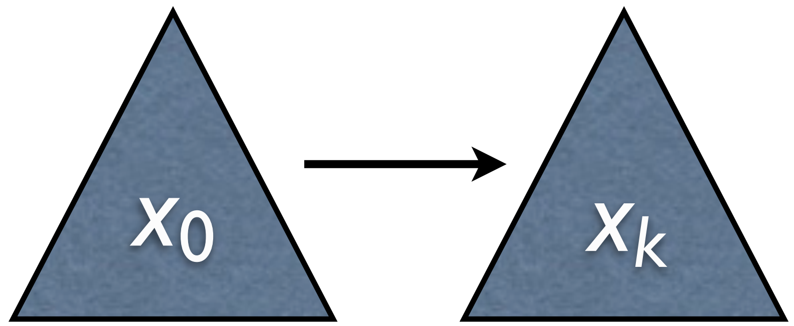
$$\begin{array}{ll} \min & f(x) \\ \text{s.t.} & c(x) \leq 0 \end{array}$$

SQP & Line Search



$$\begin{aligned} \min \quad & f(x) \\ \text{s.t.} \quad & c(x) \leq 0 \end{aligned}$$

SQP & Line Search



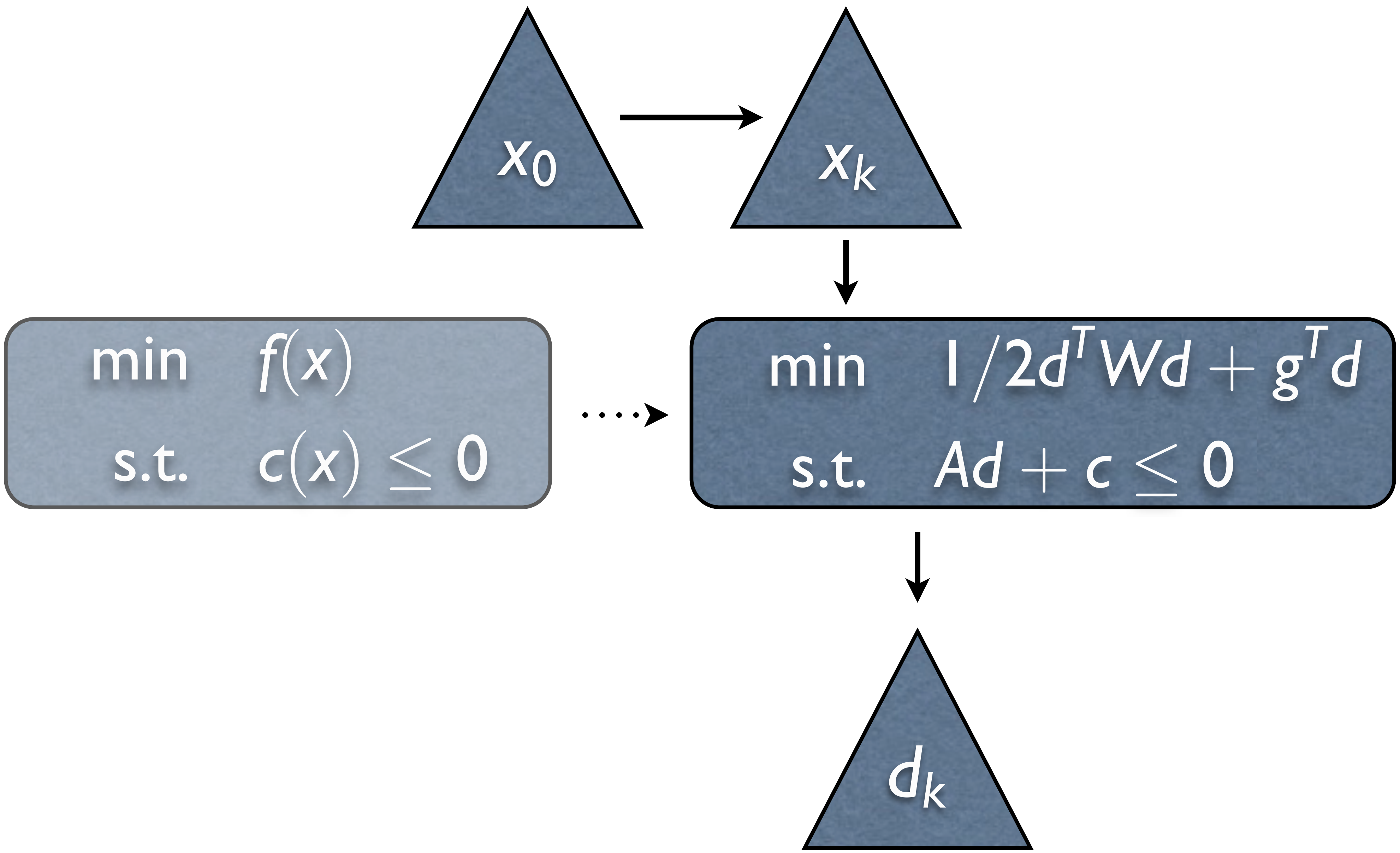
$$\begin{aligned} \min \quad & f(x) \\ \text{s.t.} \quad & c(x) \leq 0 \end{aligned}$$

SQP & Line Search

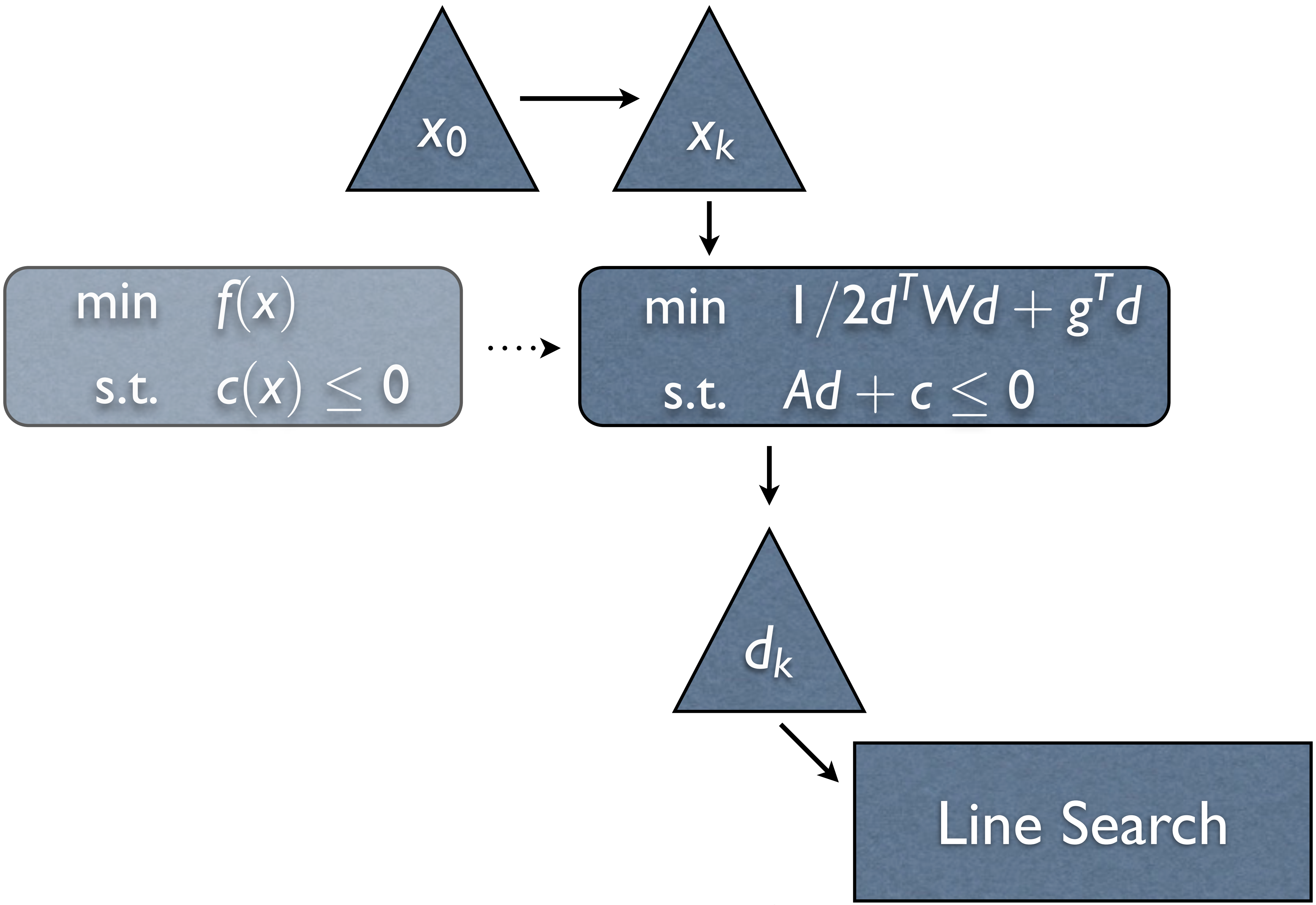
$$x_0 \longrightarrow x_k$$

$$\begin{array}{ll} \min & f(x) \\ \text{s.t.} & c(x) \leq 0 \end{array} \quad \dashrightarrow \quad \begin{array}{ll} \min & 1/2 d^T W d + g^T d \\ \text{s.t.} & Ad + c \leq 0 \end{array}$$

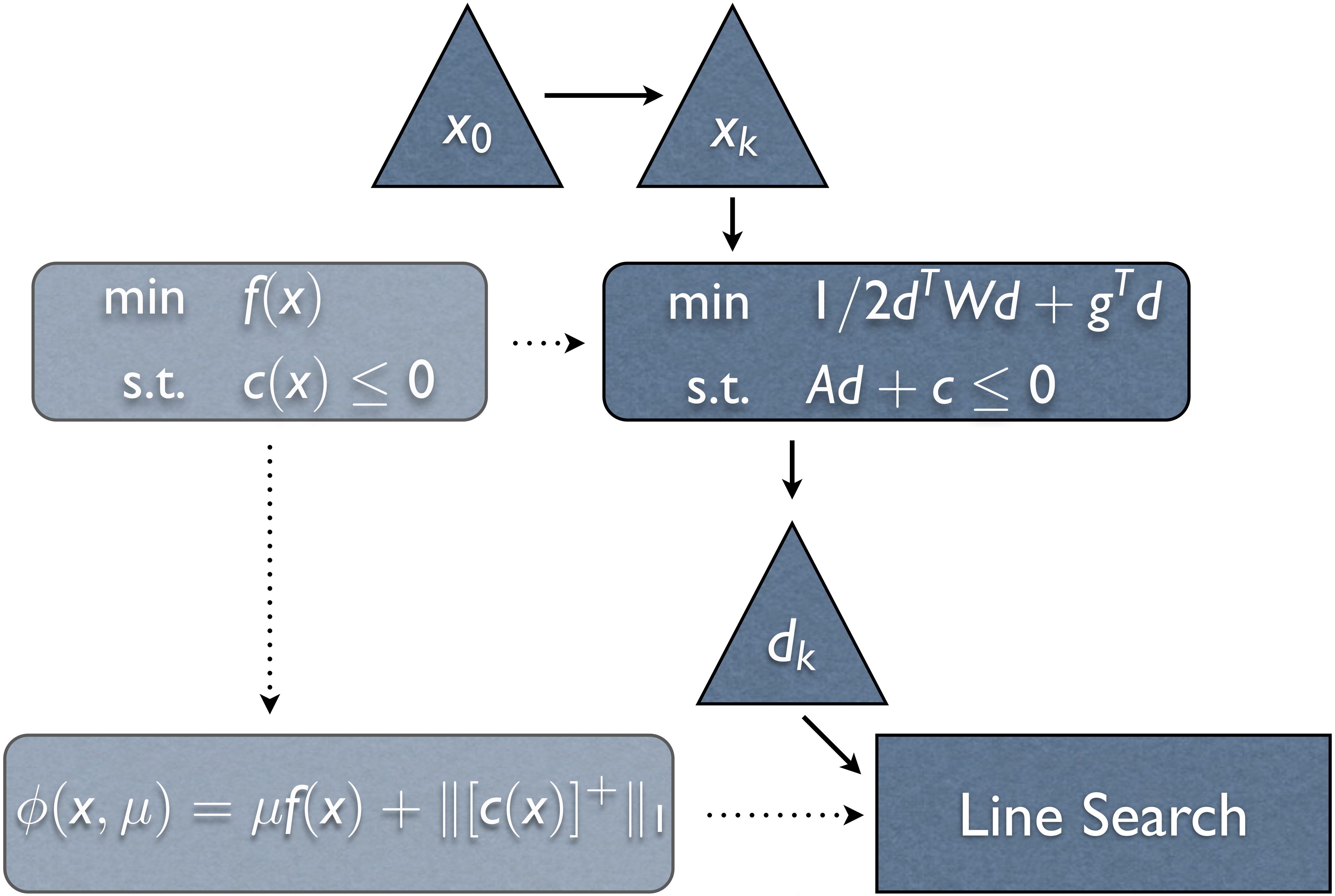
SQP & Line Search



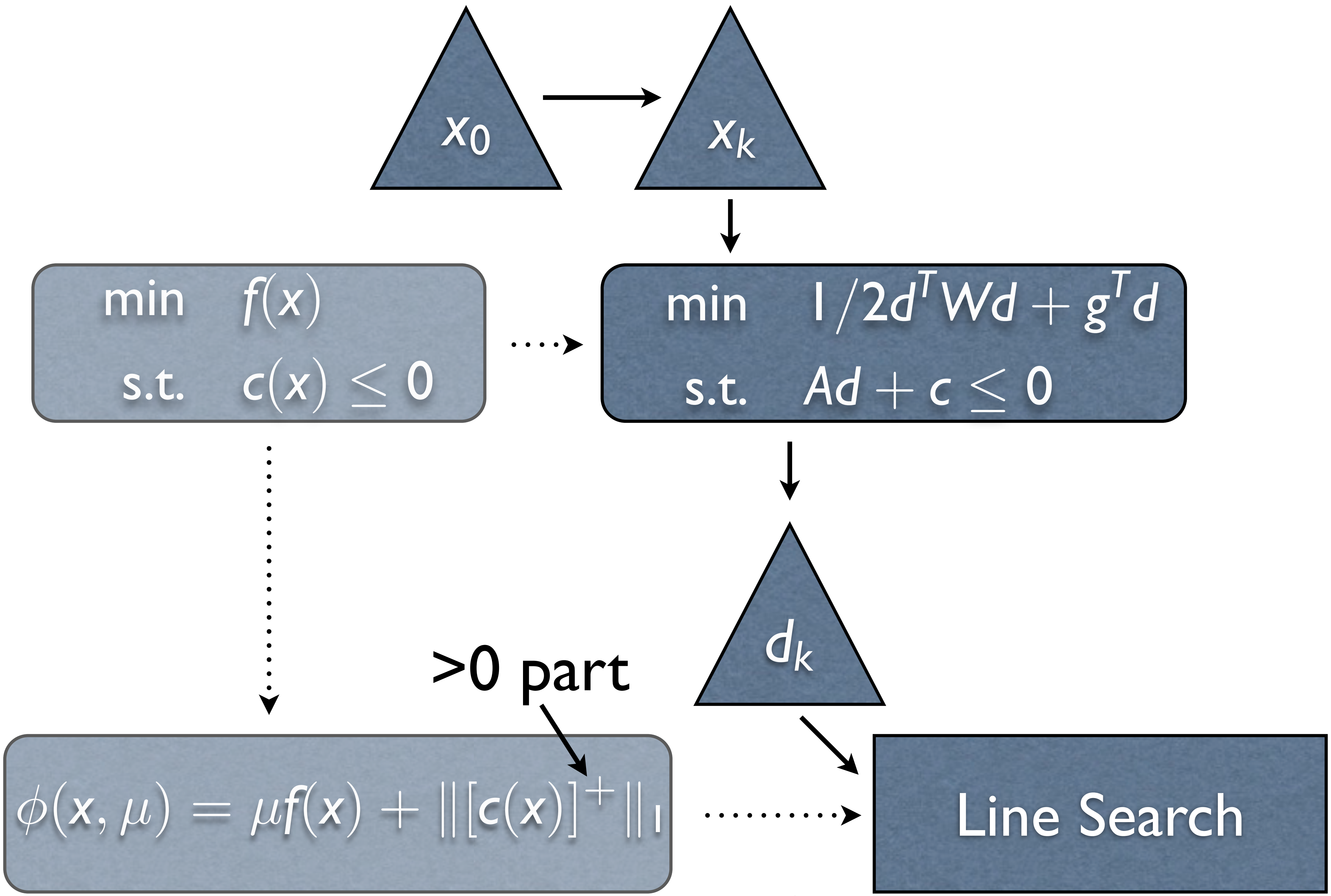
SQP & Line Search



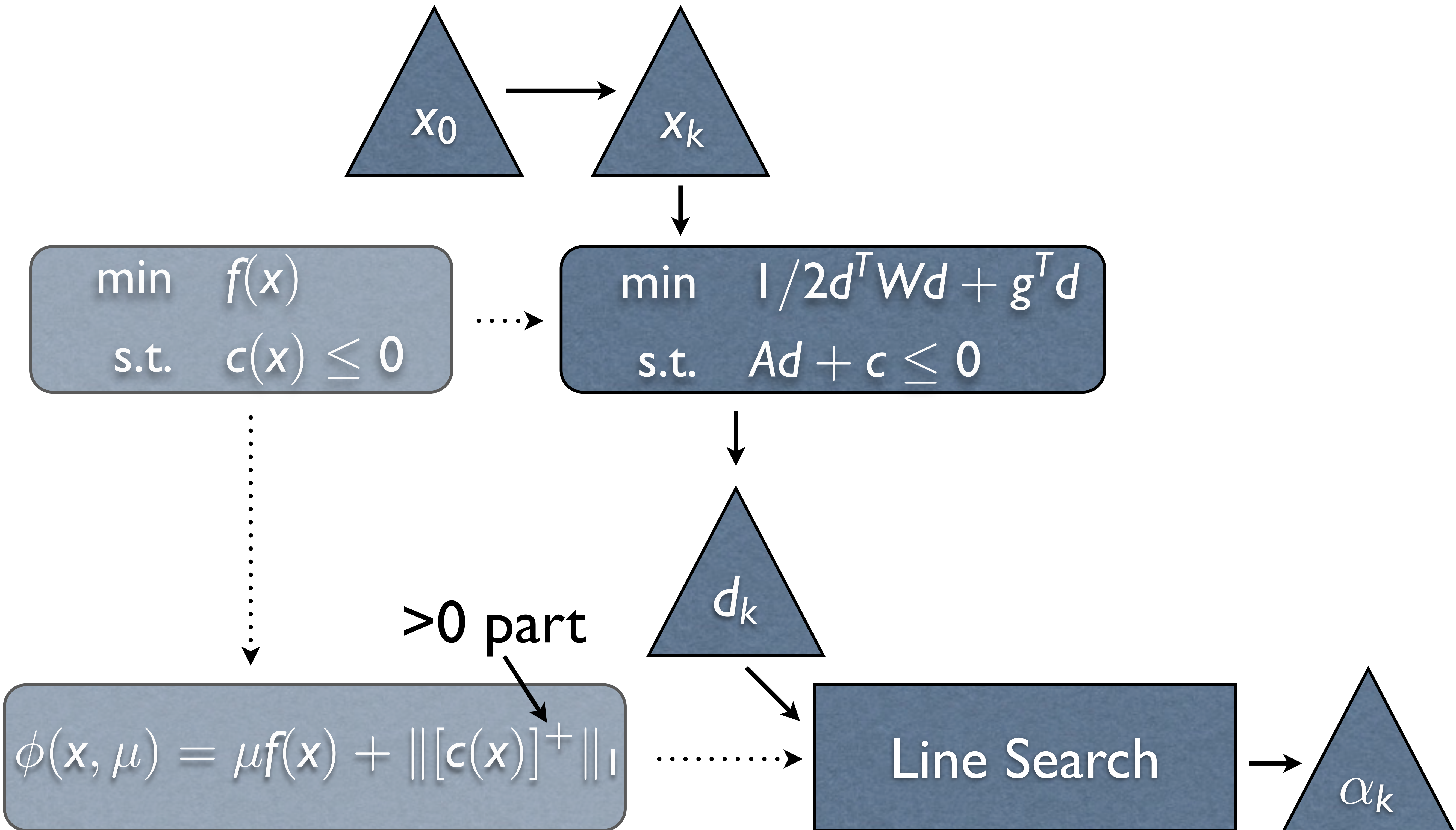
SQP & Line Search



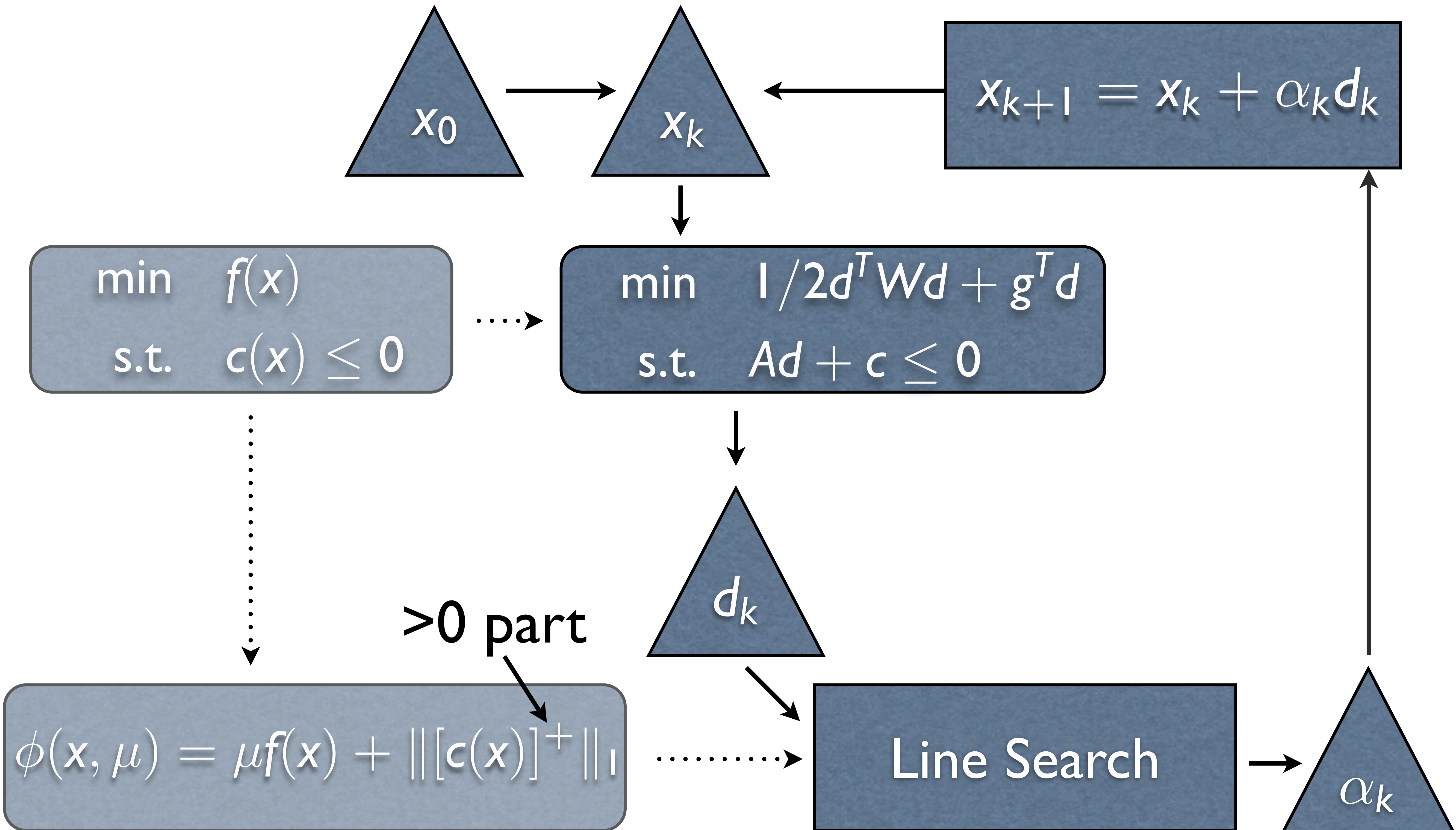
SQP & Line Search



SQP & Line Search



SQP & Line Search



Sℓ_1QP IDEA

$S\ell_1QP$ IDEA

$$\begin{array}{ll} \min & f(x) \\ \text{s.t.} & c(x) \leq 0 \end{array}$$

Sℓ_1QP IDEA

$$\begin{array}{ll} \min & f(x) \\ \text{s.t.} & c(x) \leq 0 \end{array}$$

$$\phi(x, \mu) = \mu f(x) + \|[c(x)]^+\|_1$$

Sℓ_1QP IDEA

$$\begin{aligned} \min \quad & f(x) \\ \text{s.t.} \quad & c(x) \leq 0 \end{aligned}$$

$$\phi(x, \mu) = \mu f(x) + \|[c(x)]^+\|_1$$

$$\min \quad \mu f(x) + \|[c(x)]^+\|_1$$

Sℓ_1QP IDEA

$$\begin{array}{ll} \min & f(x) \\ \text{s.t.} & c(x) \leq 0 \end{array}$$

$$\phi(x, \mu) = \mu f(x) + \|[c(x)]^+\|_1$$

$$\min \quad \mu f(x) + \|[c(x)]^+\|_1$$

$$\begin{array}{ll} \min & \mu f(x) + e^T p \\ \text{s.t.} & c(x) \leq p, p \geq 0 \end{array}$$

Sℓ_1QP IDEA

$$\min \quad f(x) \qquad \text{s.t.} \quad c(x) \le 0$$

$$\phi(x,\mu) = \mu f(x) + \|[c(x)]^+\|_1$$

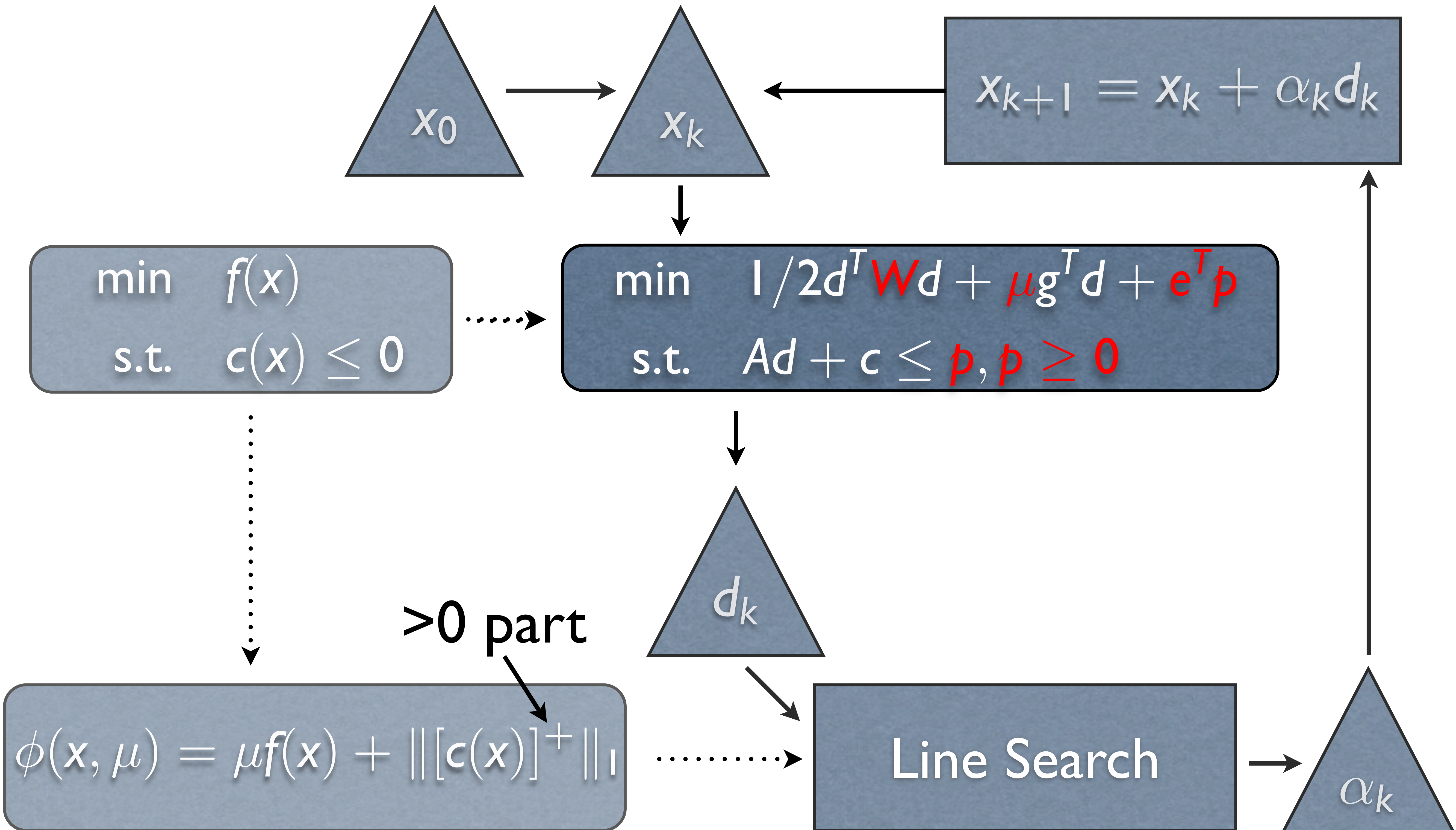
$$\min \quad \mu f(x) + \|[c(x)]^+\|_1$$

$$\min \quad \mu f(x) + e^T p \qquad \text{s.t.} \quad c(x) \le p, p \ge 0$$

ℓ_1**QP:**

$$\min \quad 1/2 d^T W d + \mu g^T d + e^T p \qquad \text{s.t.} \quad Ad + c \le p, p \ge 0$$

$S\ell_1$QP & Line Search



Two Questions

- How can we deal with inexactness?
- How can we deal with penalty parameter?

Two Questions

- **How can we deal with inexactness?**
- How can we deal with penalty parameter?

Quadratic Programming Residuals

	Nonlinear Equations	Nonlinear Optimization
Problem		
Subproblem		

Quadratic Programming Residuals

	Nonlinear Equations	Nonlinear Optimization
Problem	$\lVert F(z) \rVert$	
Subproblem		

Quadratic Programming Residuals

	Nonlinear Equations	Nonlinear Optimization
Problem	$\|F(z)\|$	
Subproblem	$\|F(z) + F'(z)d\|$	

Quadratic Programming Residuals

	Nonlinear Equations	Nonlinear Optimization
Problem	$\|F(z)\|$	NLP KKT Residual
Subproblem	$\|F(z) + F'(z)d\|$	

Quadratic Programming Residuals

	Nonlinear Equations	Nonlinear Optimization
Problem	$\|F(z)\|$	NLP KKT Residual
Subproblem	$\|F(z)+F'(z)d\|$	QP KKT Residual

Quadratic Programming Residuals

	Nonlinear Equations	Nonlinear Optimization
Problem	$\|F(z)\|$	$\left\| \begin{bmatrix} \mu\nabla f + \nabla c\lambda \\ \min([c]^+, e - \lambda) \\ \min([c]^-, \lambda) \end{bmatrix} \right\|$
Subproblem	$\|F(z) + F'(z)d\|$	QP KKT Residual

Quadratic Programming Residuals

	Nonlinear Equations	Nonlinear Optimization
Problem	$\|F(z)\|$	$\left\| \begin{bmatrix} \mu\nabla f + \nabla c\lambda \\ \min([c]^+, e - \lambda) \\ \min([c]^-, \lambda) \end{bmatrix} \right\|$
Subproblem	$\|F(z) + F'(z)d\|$	$\left\| \begin{bmatrix} \mu g + Hd + J\lambda \\ \min([c + J^T d]^+, e - \lambda) \\ \min([c + J^T d]^-, \lambda) \end{bmatrix} \right\|$

Quadratic Programming Termination Tests

Quadratic Programming Termination Tests

Nonlinear Equations

Quadratic Programming Termination Tests

Nonlinear Equations

$$\|F(z) + F'(z)d\| \leq \kappa \|F(z)\|$$

Quadratic Programming Termination Tests

Nonlinear Equations

$$\|F(z) + F'(z)d\| \leq \kappa \|F(z)\|$$

Quadratic Programming

Quadratic Programming Termination Tests

Nonlinear Equations

$$\|F(z) + F'(z)d\| \leq \kappa \|F(z)\|$$

Quadratic Programming

$$\left\| \begin{bmatrix} \mu g + Hd + J\lambda \\ \min([c + J^T d]^+, e - \lambda) \\ \min([c + J^T d]^-, \lambda) \end{bmatrix} \right\| \leq \kappa \left\| \begin{bmatrix} \mu \nabla f + \nabla c \lambda \\ \min([c]^+, e - \lambda) \\ \min([c]^-, \lambda) \end{bmatrix} \right\|$$

Quadratic Programming Termination Tests

Nonlinear Equations

$$\|F(z) + F'(z)d\| \leq \kappa \|F(z)\|$$

Quadratic Programming

$$\left\| \begin{bmatrix} \mu g + Hd + J\lambda \\ \min([c + J^T d]^+, e - \lambda) \\ \min([c + J^T d]^-, \lambda) \end{bmatrix} \right\| \leq \kappa \left\| \begin{bmatrix} \mu \nabla f + \nabla c \lambda \\ \min([c]^+, e - \lambda) \\ \min([c]^-, \lambda) \end{bmatrix} \right\|$$

+ other conditions (on linear model reduction, multipliers, etc)

Two Questions

- How can we deal with inexactness?
- **How can we deal with penalty parameter?**

Sℓ_1QP

is exact if

Sℓ_1QP

- Infeasible subproblems...

is exact if

Sℓ_1QP

- Infeasible subproblems...
- and infeasible problems!

is exact if

$S\ell_1 QP$

- Infeasible subproblems...
- and infeasible problems!
- can be EXACT, but

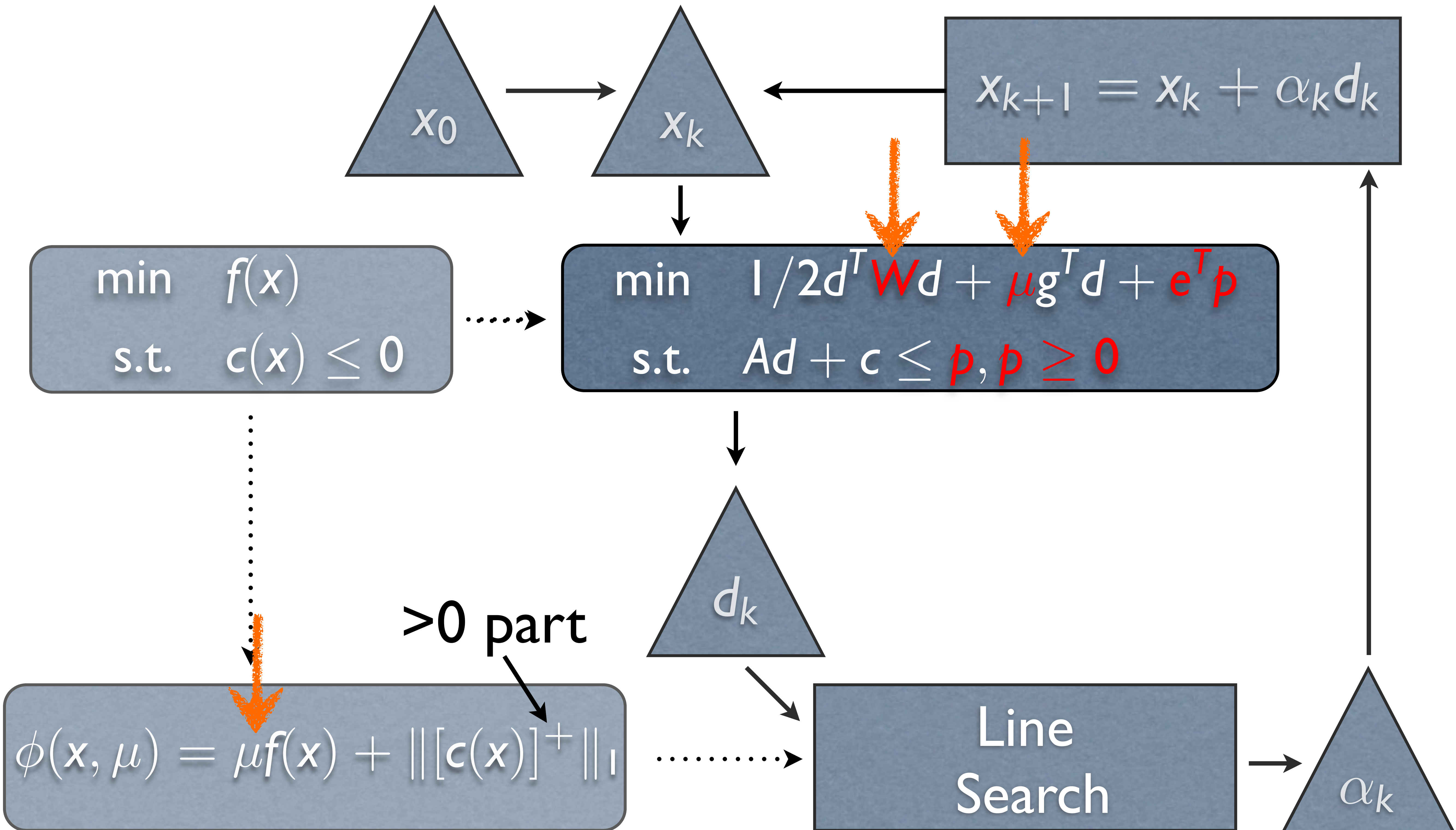
is exact if

Sℓ_1QP

- Infeasible subproblems...
- and infeasible problems!
- can be EXACT, but Need $\mu \leq \frac{1}{\|\lambda^*\|_\infty}$

is exact if

Sℓ_1QP



x_0
x_k
$x_{k+1} = x_k + \alpha_k d_k$
min $f(x)$
s.t. $c(x) \leq 0$
min $1/2 d^T W d + \mu g^T d + e^T p$
s.t. $Ad + c \leq p, p \geq 0$
d_k
>0 part
$\phi(x, \mu) = \mu f(x) + \|[c(x)]^+\|_1$
Line Search
α_k

Sℓ_1QP: μ strategy

REWORK THIS SLIDE.

Sℓ_1QP: μ strategy

REWORK THIS SLIDE.

μ too large: Not exact yet? Infeasible??

Sℓ_1QP: μ strategy

REWORK THIS SLIDE.

μ too large: Not exact yet? Infeasible??

μ too small: Stuck at feasible point???

IDEA: Penalty Steering

$$\begin{array}{ll} \min & f(x) \\ \text{s.t.} & c(x) \leq 0 \end{array}$$

IDEA: Penalty Steering

$$\begin{aligned} \min \quad & f(x) \\ \text{s.t.} \quad & c(x) \leq 0 \end{aligned}$$

$$\mu = \mu_k$$

$$\min \quad \mu f(x) + \|[c(x)]^+\|_1$$

IDEA: Penalty Steering

$$\min \quad f(x) \qquad \text{s.t.} \quad c(x) \le 0$$

$$\mu = \mu_k$$

$$\min \quad \mu f(x) + \|[c(x)]^+\|_1$$

$$\mu = 0$$

$$\min \quad \|[c(x)]^+\|_1$$

IDEA: Penalty Steering

$$\begin{aligned} \min \quad & f(x) \\ \text{s.t.} \quad & c(x) \leq 0 \end{aligned}$$

$\mu = \mu_k$

$$\min \quad \mu f(x) + \|[c(x)]^+\|_1$$

$\mu = 0$

$$\min \quad \|[c(x)]^+\|_1$$

$S\ell_1QP$ Framework

IDEA: Penalty Steering

$$\min \quad f(x)$$
$$\text{s.t.} \quad c(x) \le 0$$

$\mu = \mu_k$

$$\min \quad \mu f(x) + \|[c(x)]^+\|_1$$

$\mu = 0$

$$\min \quad \|[c(x)]^+\|_1$$

Sℓ_1QP Framework

Solve

IDEA: Penalty Steering

$$\min \quad f(x)$$
$$\text{s.t.} \quad c(x) \le 0$$

$\mu = \mu_k$

$$\min \quad \mu f(x) + \|[c(x)]^+\|_1$$

$\mu = 0$

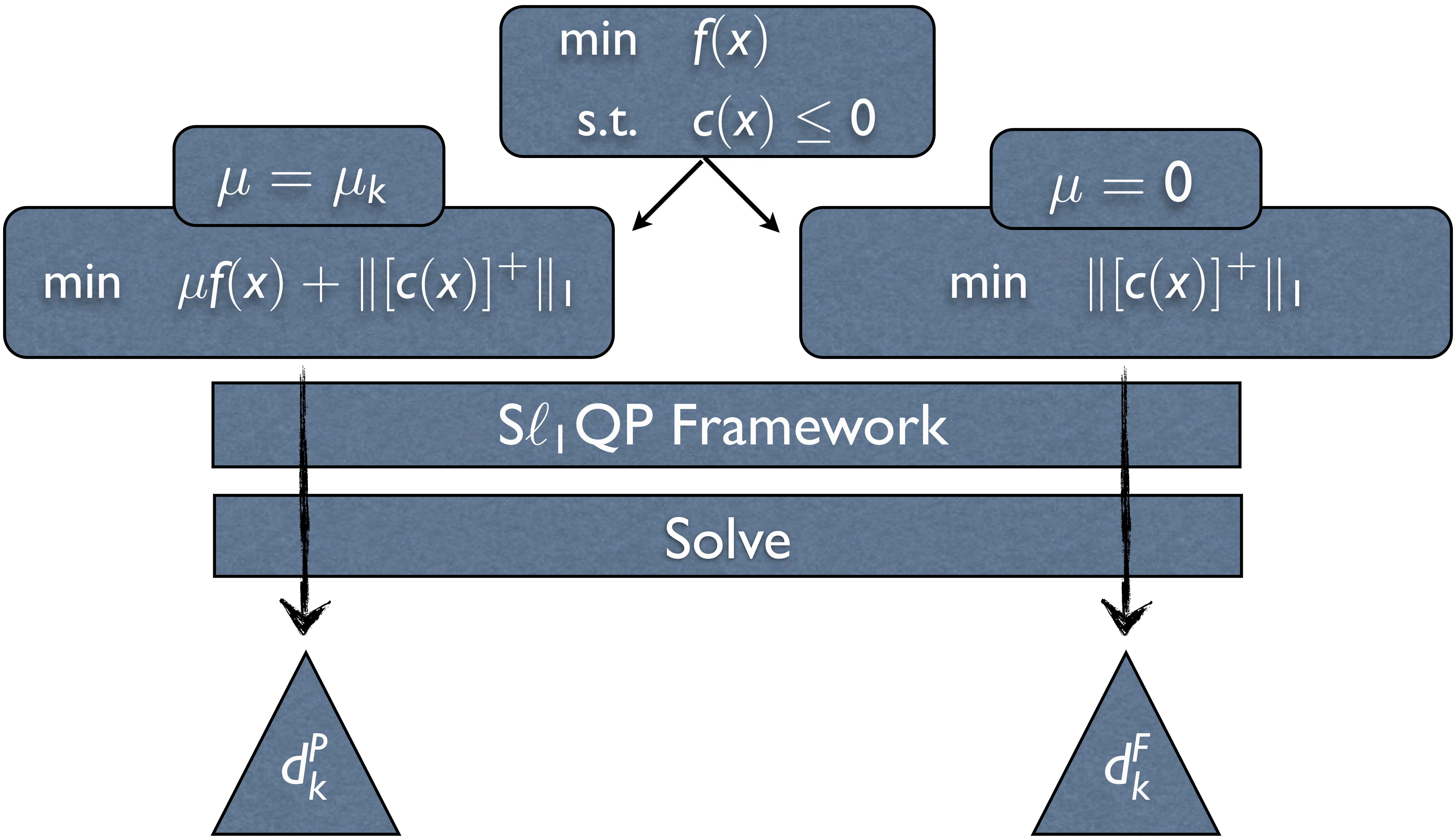
$$\min \quad \|[c(x)]^+\|_1$$

Sℓ_1QP Framework

Solve

d_k^P

IDEA: Penalty Steering



IDEA: Penalty Steering

$$\min \quad f(x)$$
$$\text{s.t.} \quad c(x) \leq 0$$

$\mu = \mu_k$

$$\min \quad \mu f(x) + \|[c(x)]^+\|_1$$

$\mu = 0$

$$\min \quad \|[c(x)]^+\|_1$$

Sℓ_1QP Framework

Solve

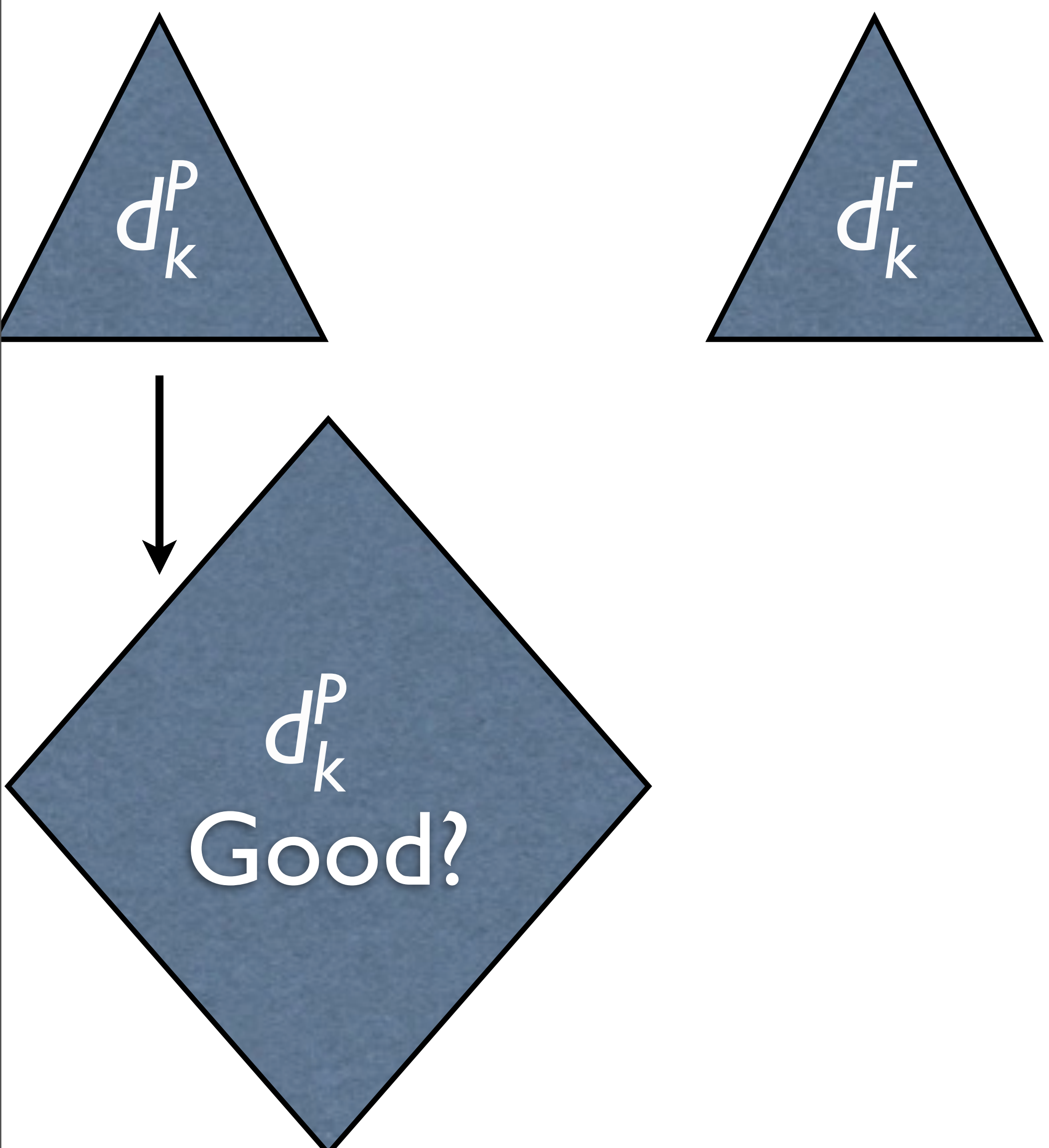
d_k^P

Now what??

d_k^F

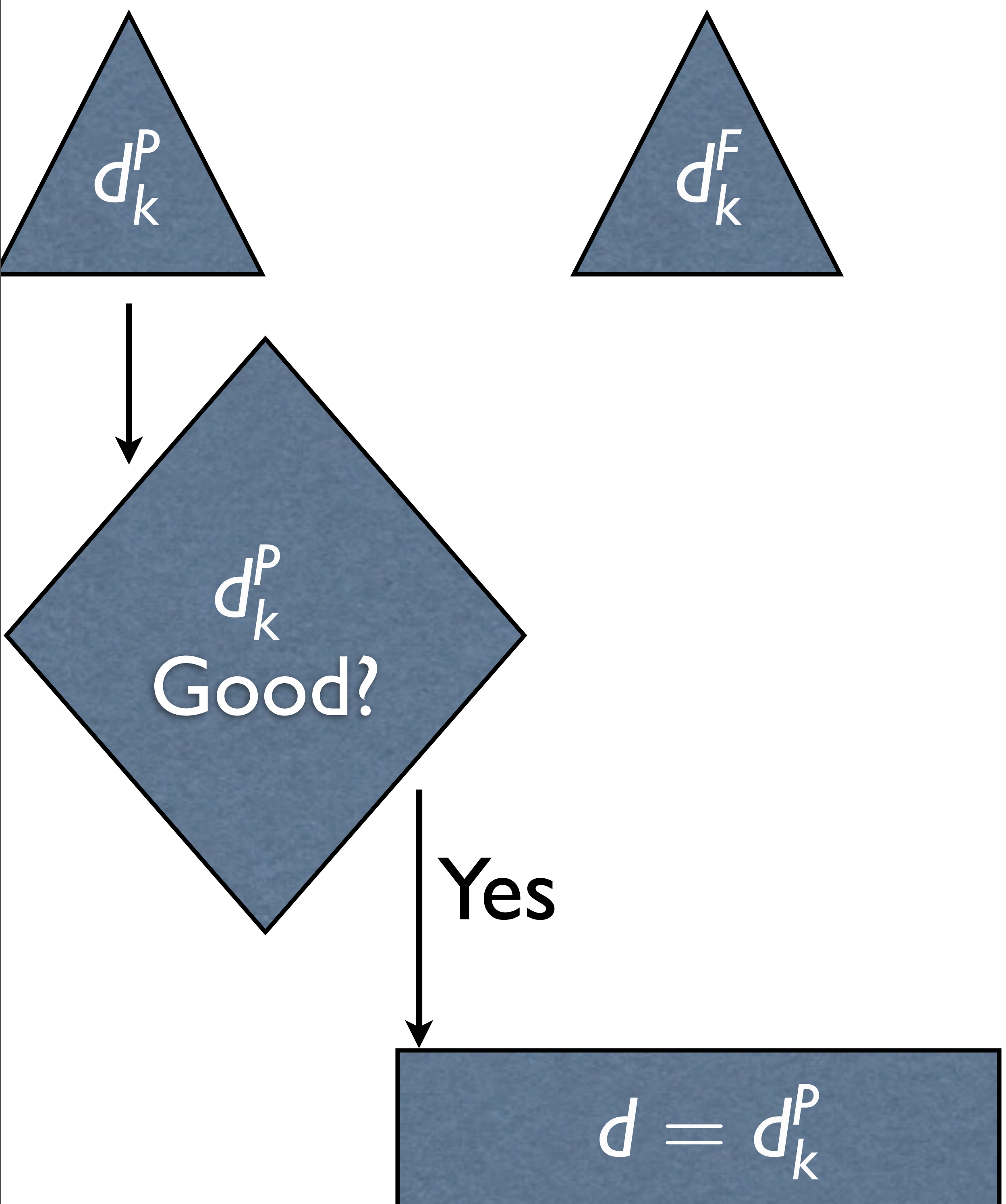
IDEA: Penalty Steering

IDEA: Penalty Steering

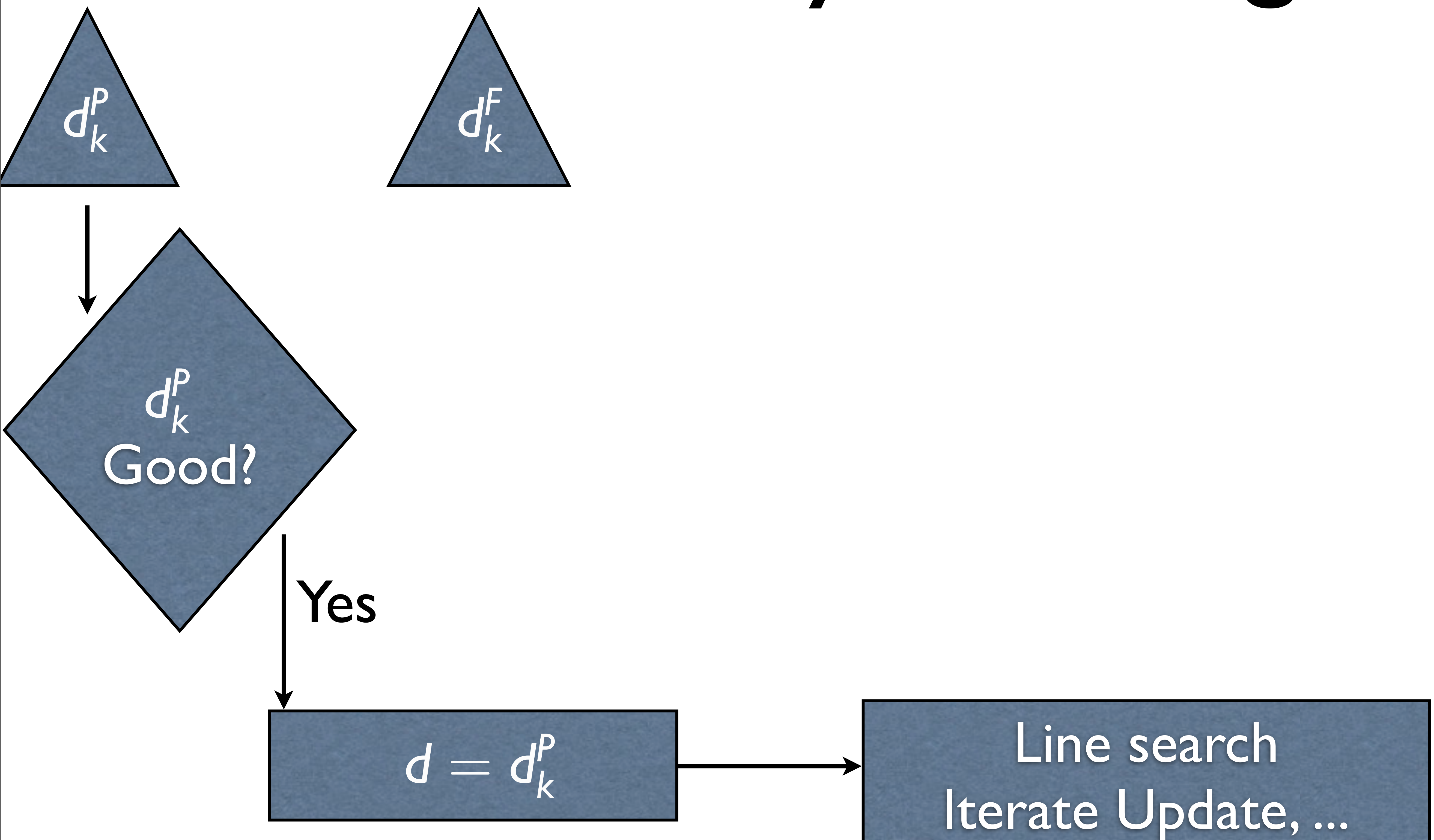


Wednesday, October 23, 13

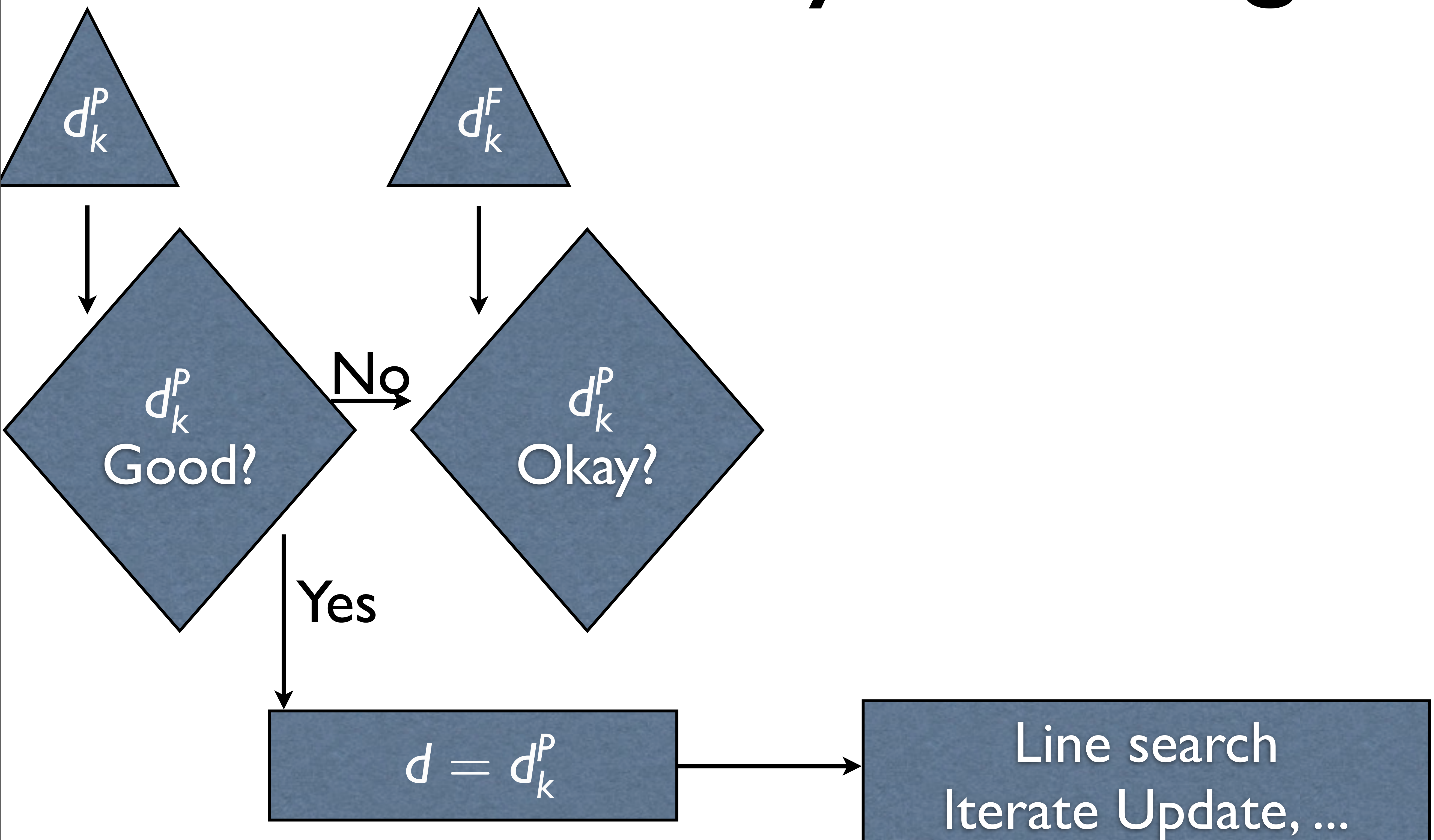
IDEA: Penalty Steering



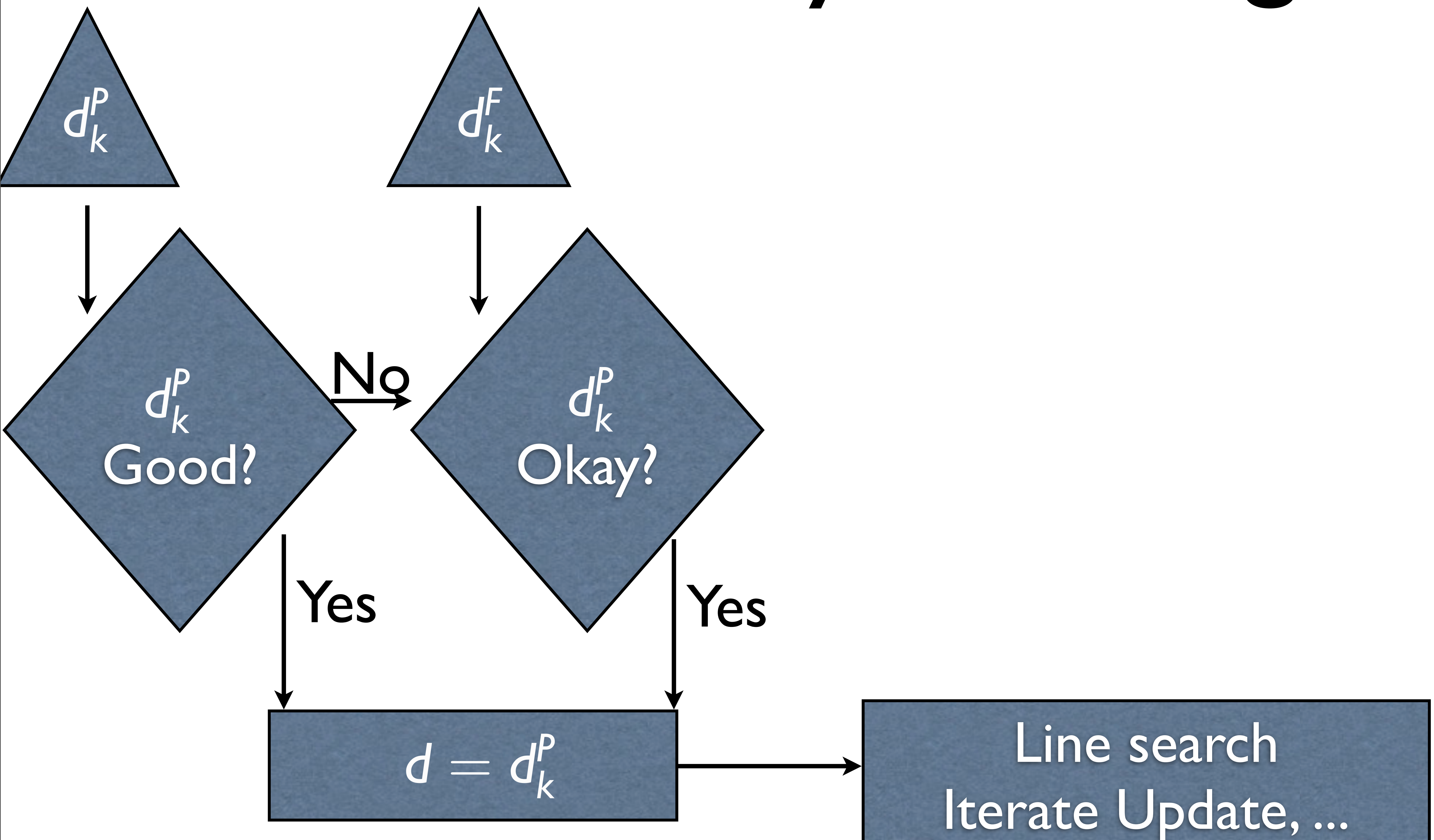
IDEA: Penalty Steering



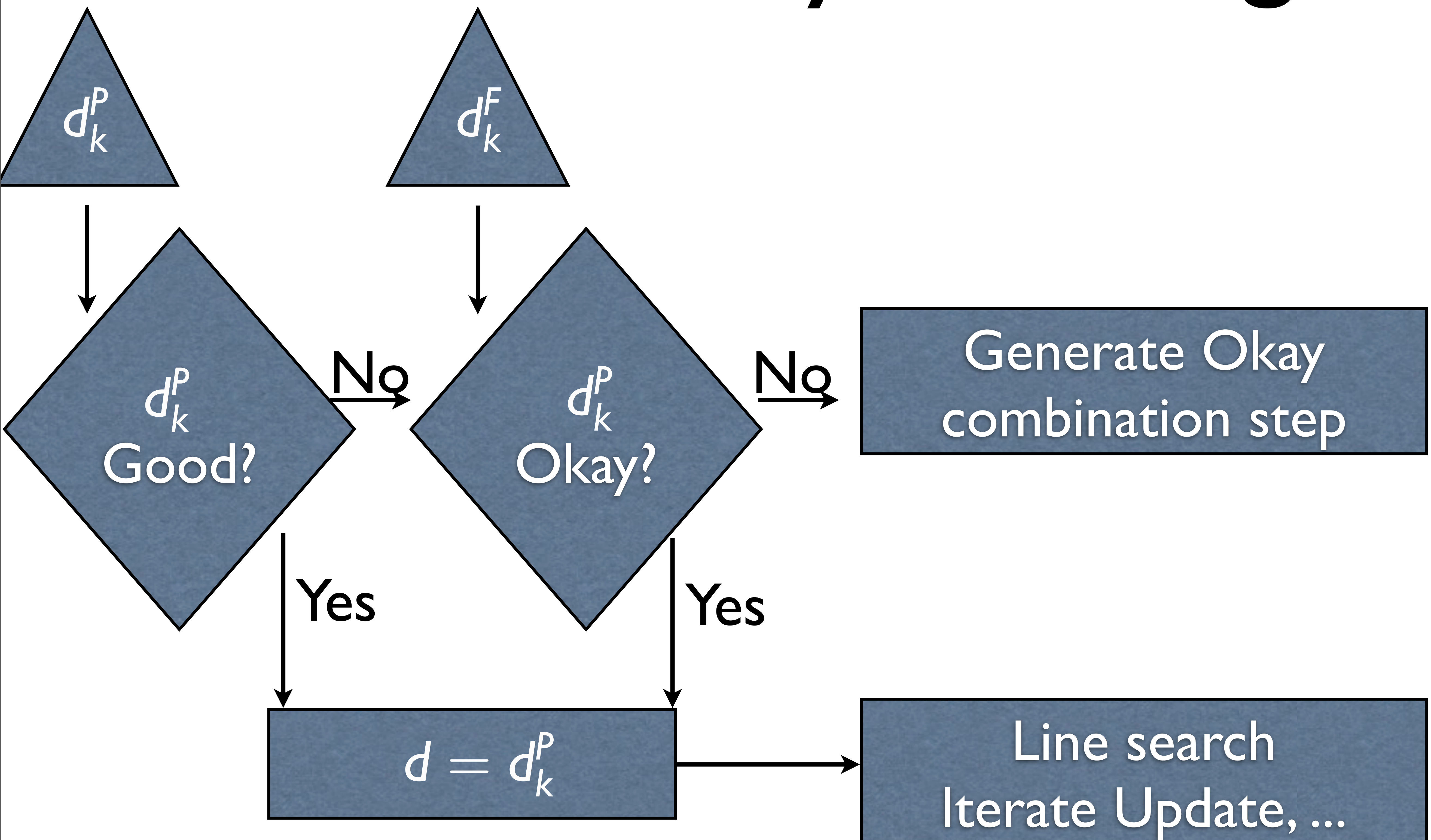
IDEA: Penalty Steering



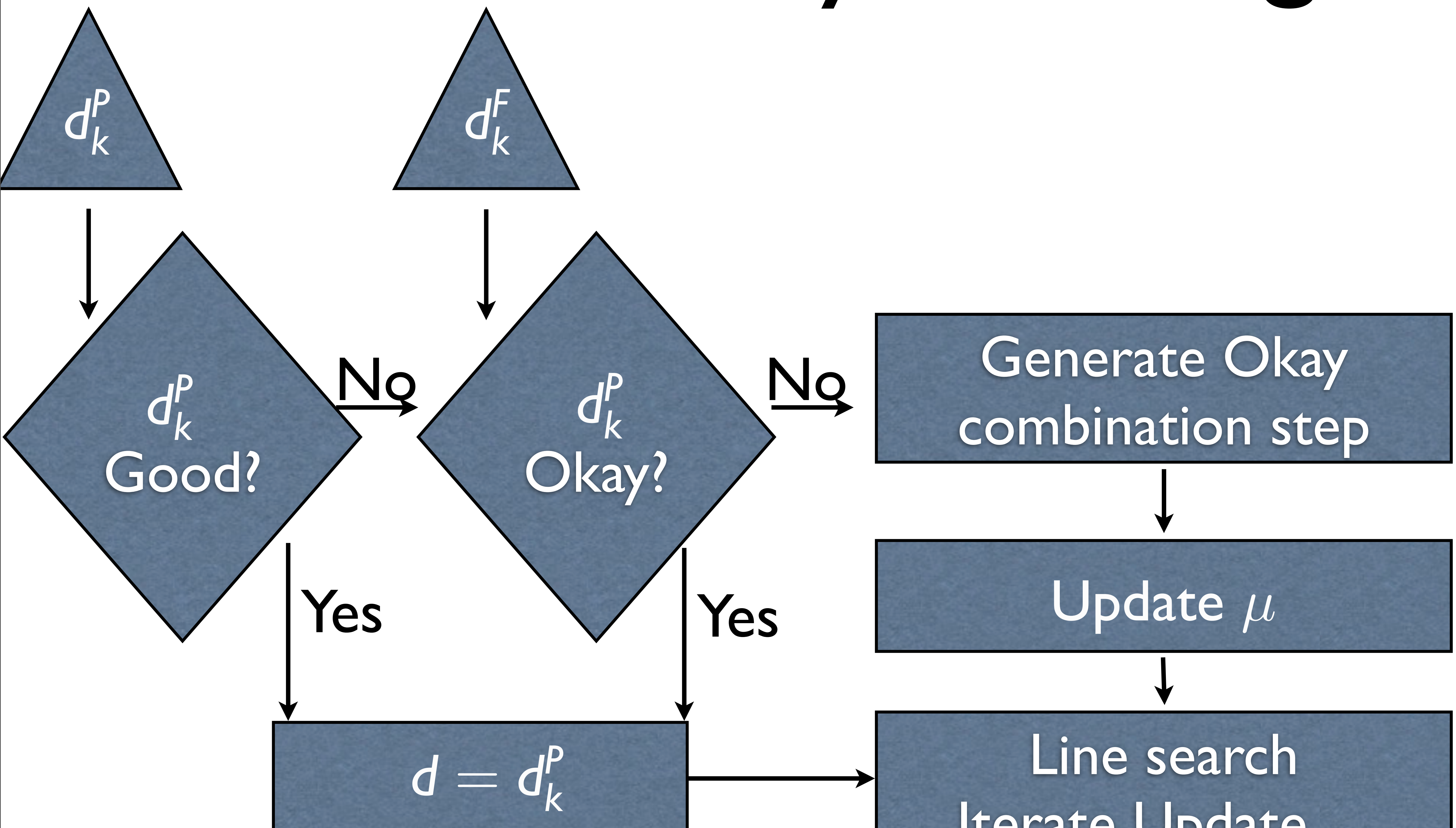
IDEA: Penalty Steering



IDEA: Penalty Steering



IDEA: Penalty Steering



Penalty Steering Details

move non-convexity strategy to inexact subsection?

27 minutes

- Good? Okay?
- What combination?
- Need to solve two QPs?
- What about non-convexity?

Penalty Steering Details

- **Good? Okay?**
- What combination?
- Need to solve two QPs?
- What about non-convexity?

Good or Okay? Intuition

d_k^P
Good?

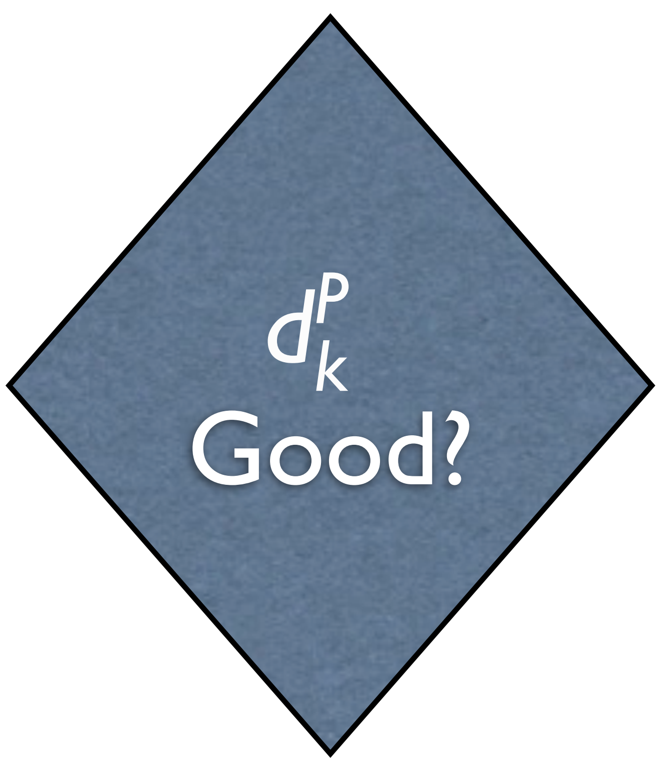
d_k^P
Okay?

Distance to minimizer proportional to infeasibility

Distance to minimizer proportional to distance to minimizer of infeasibility

Good or Okay? Rigorous

- $\phi(x;\mu) = \mu f(x) + \|[c(x)]^+\|_1$
- $l(d;\mu) = \mu(f_k + g_k^T d) + \|[c_k + J_k^T d]^+\|_1$
- $\Delta l(d;\mu) = l(0;\mu) - l(d;\mu) = -\mu(g_k^T d) + \|[c_k]^+\|_1 - \|[c_k + J_k^T d]^+\|_1$



Good or Okay? Rigorous

- $\phi(x; \mu) = \mu f(x) + \|[c(x)]^+\|_1$
- $l(d; \mu) = \mu(f_k + g_k^T d) + \|[c_k + J_k^T d]^+\|_1$
- $\Delta l(d; \mu) = l(0; \mu) - l(d; \mu) = -\mu(g_k^T d) + \|[c_k]^+\|_1 - \|[c_k + J_k^T d]^+\|_1$

$$\Delta l(d_k^P, \mu) \geq \epsilon \|[c(x_k)]^+\|_1$$

Good or Okay? Rigorous

- $\phi(x;\mu) = \mu f(x) + \|[c(x)]^+\|_1$
- $l(d;\mu) = \mu(f_k + g_k^T d) + \|[c_k + J_k^T d]^+\|_1$
- $\Delta l(d;\mu) = l(0;\mu) - l(d;\mu) = -\mu(g_k^T d) + \|[c_k]^+\|_1 - \|[c_k + J_k^T d]^+\|_1$

d_k^P
Good?

d_k^P
Okay?

$$\Delta l(d_k^P, \mu) \geq \epsilon \|[c(x_k)]^+\|_1$$

$$\Delta l(d_k^P, \mu) \geq \epsilon \Delta l(d_k^F, 0)$$

Penalty Steering Details

- Good? Okay?
- **What combination?**
- Need to solve two QPs?
- What about non-convexity?

Combining Steps

- $\phi(x;\mu) = \mu f(x) + \|[c(x)]^+\|_1$
- $l(d;\mu) = \mu(f_k + g_k^T d) + \|[c_k + J_k^T d]^+\|_1$
- $\Delta l(d;\mu) = l(0;\mu) - l(d;\mu) = -\mu(g_k^T d) + \|[c_k]^+\|_1 - \|[c_k + J_k^T d]^+\|_1$

GOAL: $\tau d_k^P + (1-\tau) d_k^F$ Okay

swap?

Combining Steps

- $\phi(x; \mu) = \mu f(x) + \|[c(x)]^+\|_1$
- $l(d; \mu) = \mu(f_k + g_k^T d) + \|[c_k + J_k^T d]^+\|_1$
- $\Delta l(d; \mu) = l(0; \mu) - l(d; \mu) = -\mu(g_k^T d) + \|[c_k]^+\|_1 - \|[c_k + J_k^T d]^+\|_1$

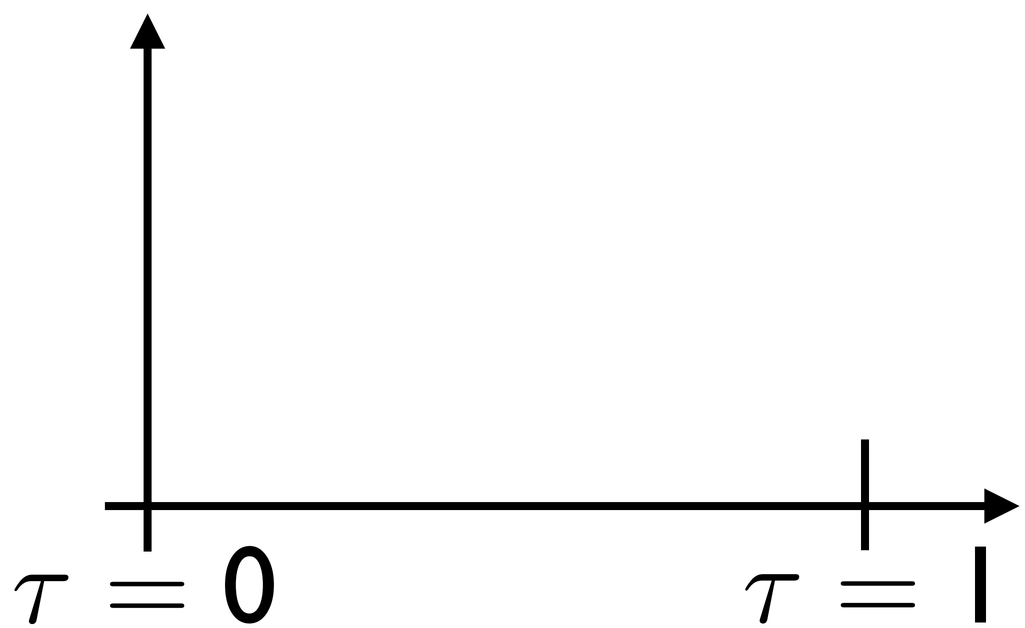
GOAL: $\Delta l(\tau d_k^P + (1 - \tau) d_k^F, \mu) \geq \epsilon \Delta l(d_k^F, 0)$

swap?

Combining Steps

- $\phi(x;\mu) = \mu f(x) + \|[c(x)]^+\|_1$
- $l(d;\mu) = \mu(f_k + g_k^T d) + \|[c_k + J_k^T d]^+\|_1$
- $\Delta l(d;\mu) = l(0;\mu) - l(d;\mu) = -\mu(g_k^T d) + \|[c_k]^+\|_1 - \|[c_k + J_k^T d]^+\|_1$

GOAL: $\Delta l(\tau d_k^P + (1-\tau)d_k^F, \mu) \geq \epsilon \Delta l(d_k^F, 0)$

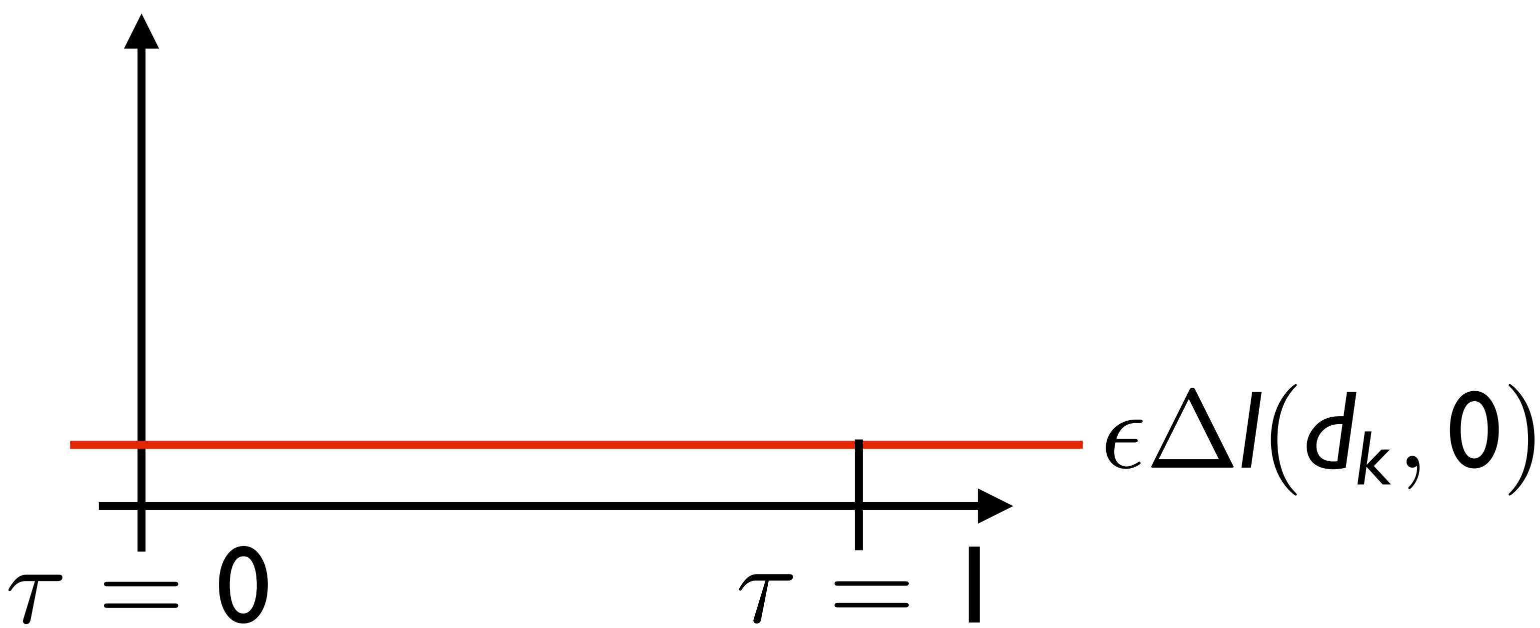


swap?

Combining Steps

- $\phi(x;\mu) = \mu f(x) + \|[c(x)]^+\|_1$
- $l(d;\mu) = \mu(f_k + g_k^T d) + \|[c_k + J_k^T d]^+\|_1$
- $\Delta l(d;\mu) = l(0;\mu) - l(d;\mu) = -\mu(g_k^T d) + \|[c_k]^+\|_1 - \|[c_k + J_k^T d]^+\|_1$

GOAL: $\Delta l(\tau d_k^P + (1-\tau) d_k^F, \mu) \geq \epsilon \Delta l(d_k^F, 0)$

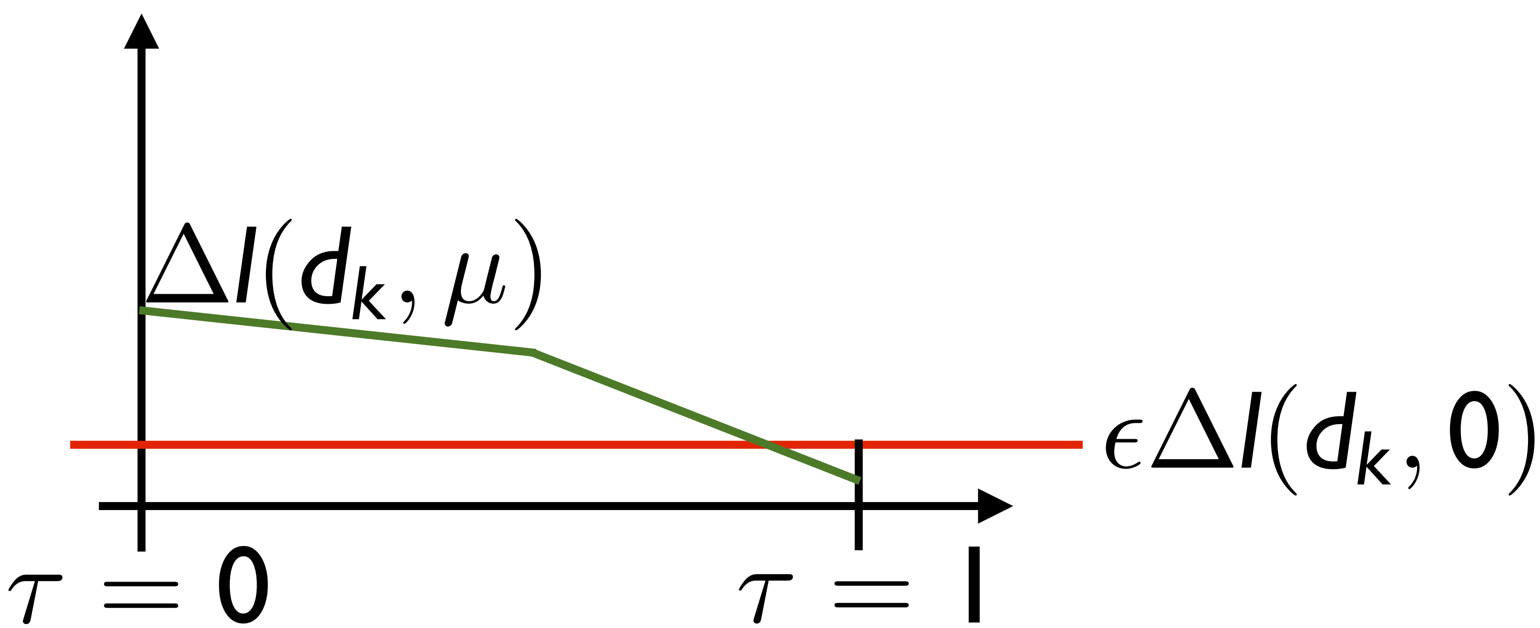


swap?

Combining Steps

- $\phi(x;\mu) = \mu f(x) + \|[c(x)]^+\|_1$
- $l(d;\mu) = \mu(f_k + g_k^T d) + \|[c_k + J_k^T d]^+\|_1$
- $\Delta l(d;\mu) = l(0;\mu) - l(d;\mu) = -\mu(g_k^T d) + \|[c_k]^+\|_1 - \|[c_k + J_k^T d]^+\|_1$

GOAL: $\Delta l(\tau d_k^P + (1-\tau)d_k^F, \mu) \geq \epsilon \Delta l(d_k^F, 0)$

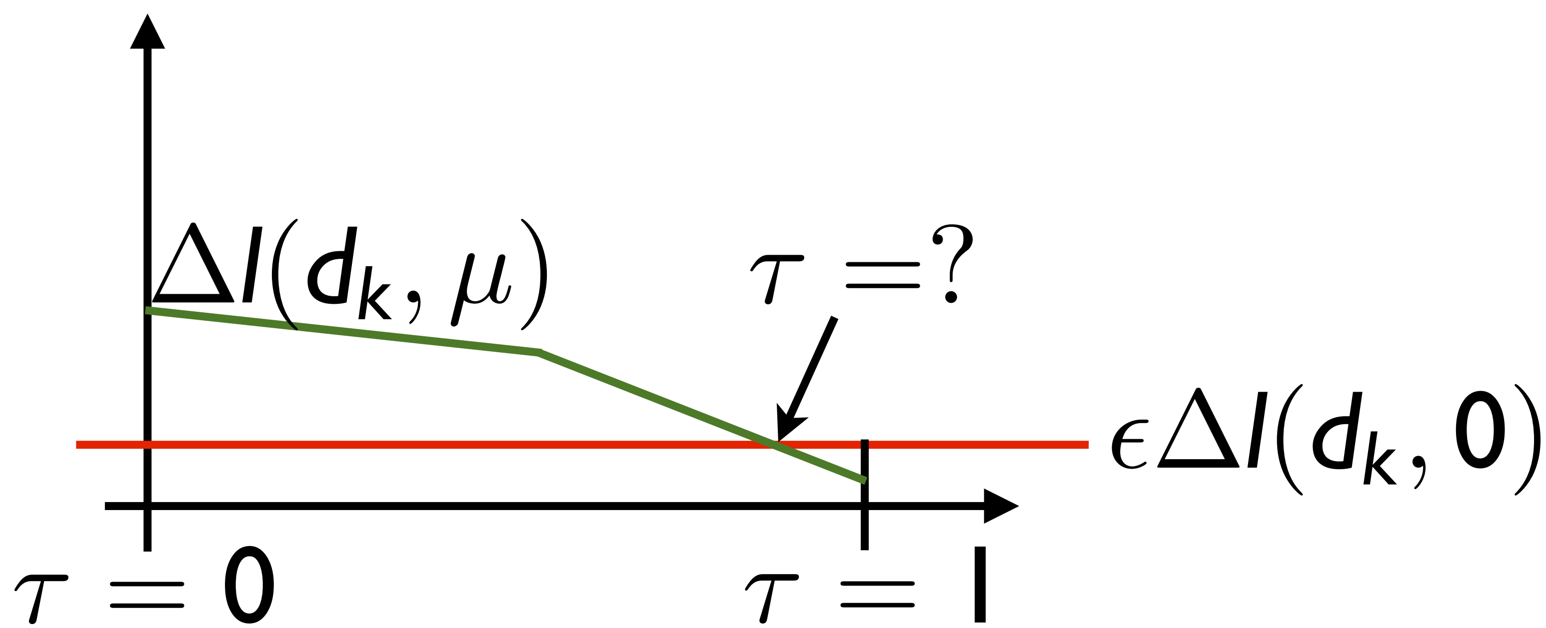


swan?

Combining Steps

- $\phi(x;\mu) = \mu f(x) + \|[c(x)]^+\|_1$
- $l(d;\mu) = \mu(f_k + g_k^T d) + \|[c_k + J_k^T d]^+\|_1$
- $\Delta l(d;\mu) = l(0;\mu) - l(d;\mu) = -\mu(g_k^T d) + \|[c_k]^+\|_1 - \|[c_k + J_k^T d]^+\|_1$

GOAL: $\Delta l(\tau d_k^P + (1-\tau)d_k^F, \mu) \geq \epsilon \Delta l(d_k^F, 0)$

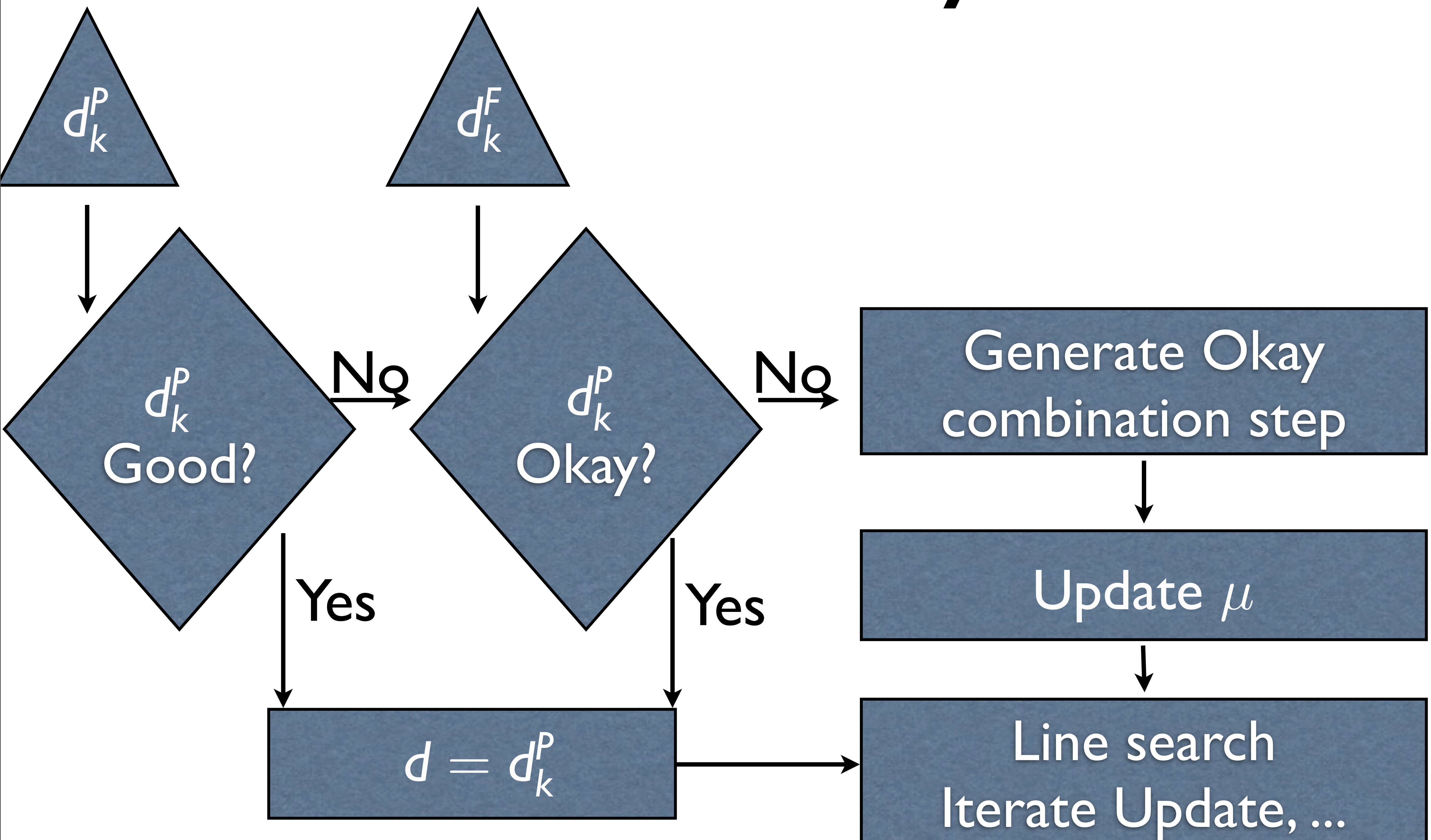


swap?

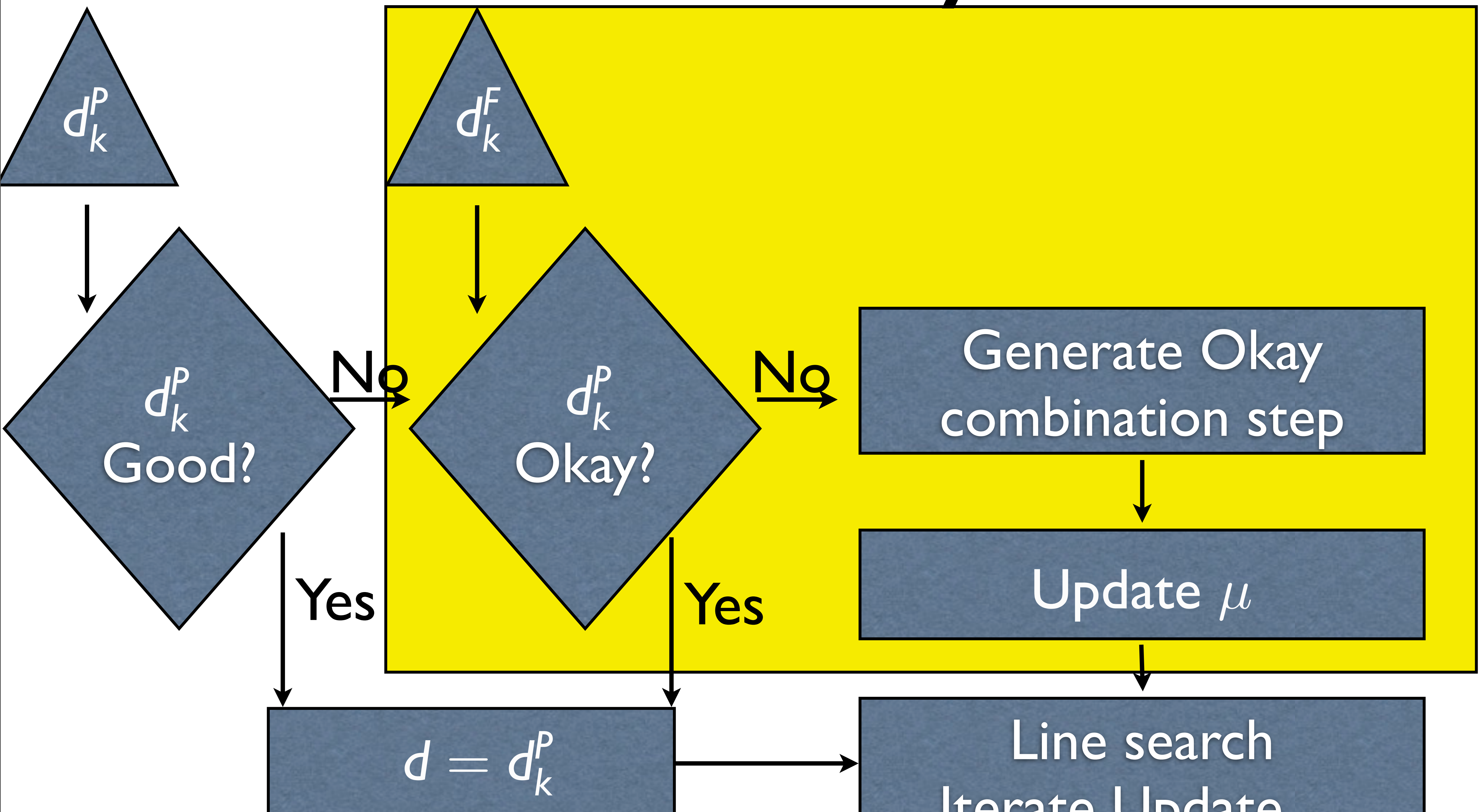
Penalty Steering Details

- Good? Okay?
- What combination?
- **Need to solve two QPs?**
- What about non-convexity?

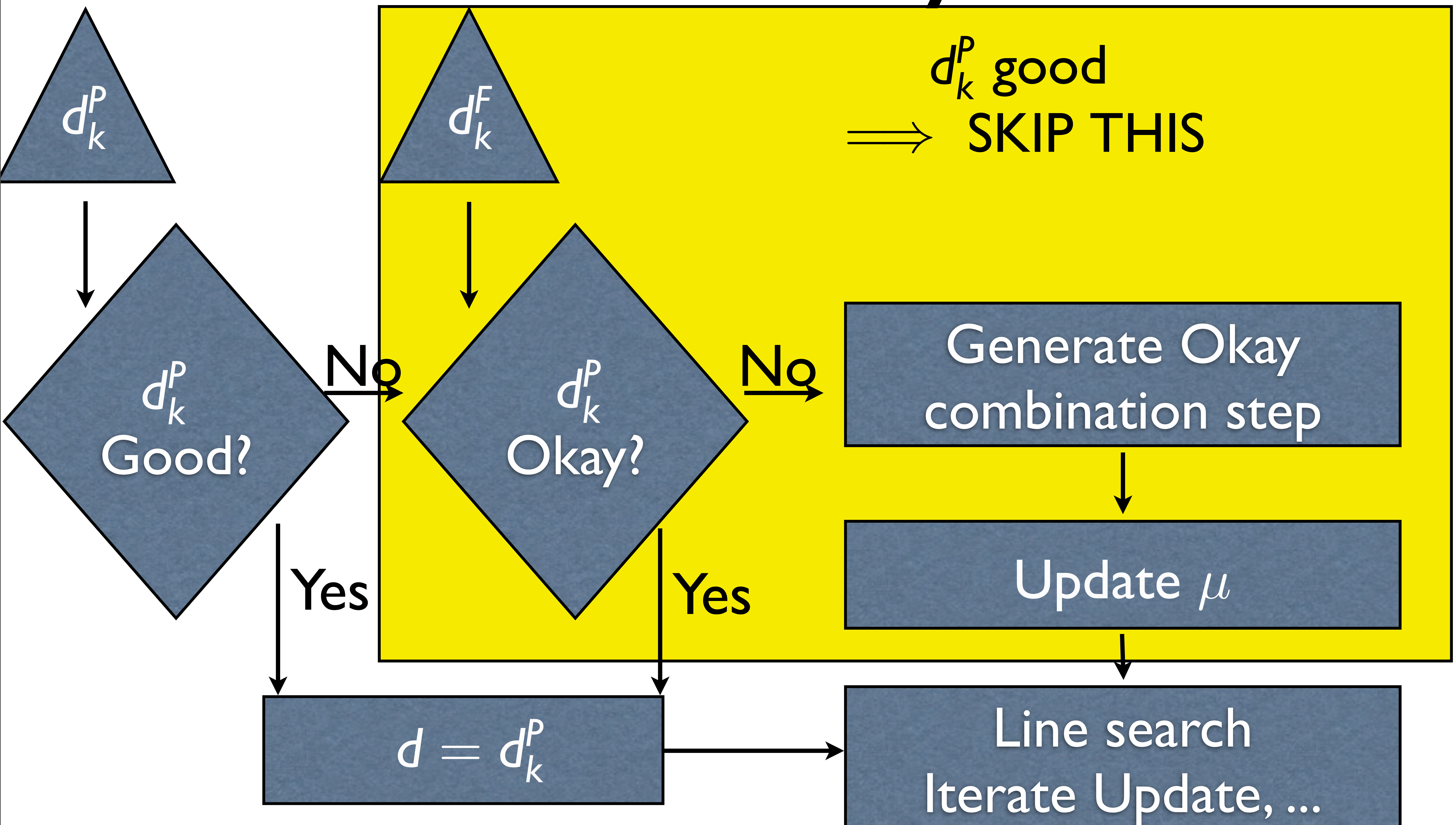
Avoid Feasibility Solve



Avoid Feasibility Solve



Avoid Feasibility Solve



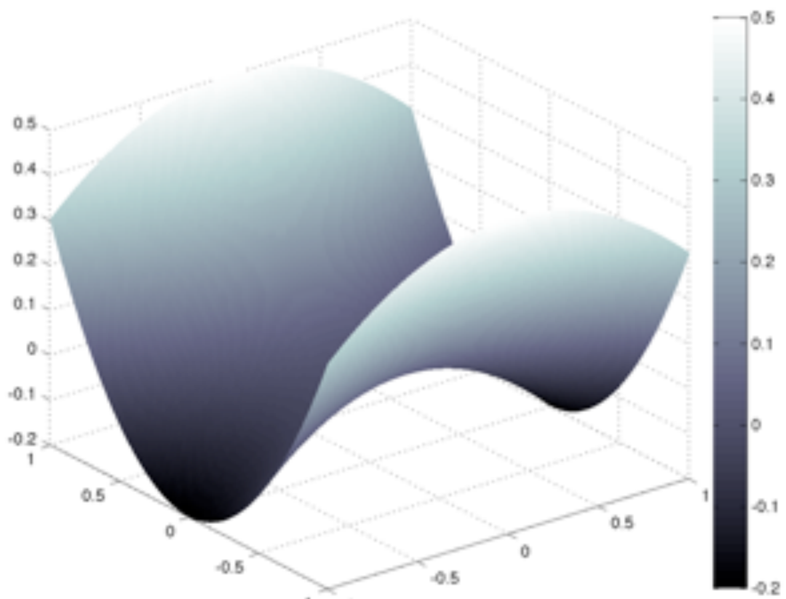
Penalty Steering Details

- Good? Okay?
- What combination?
- Need to solve two QPs?
- **What about non-convexity?**

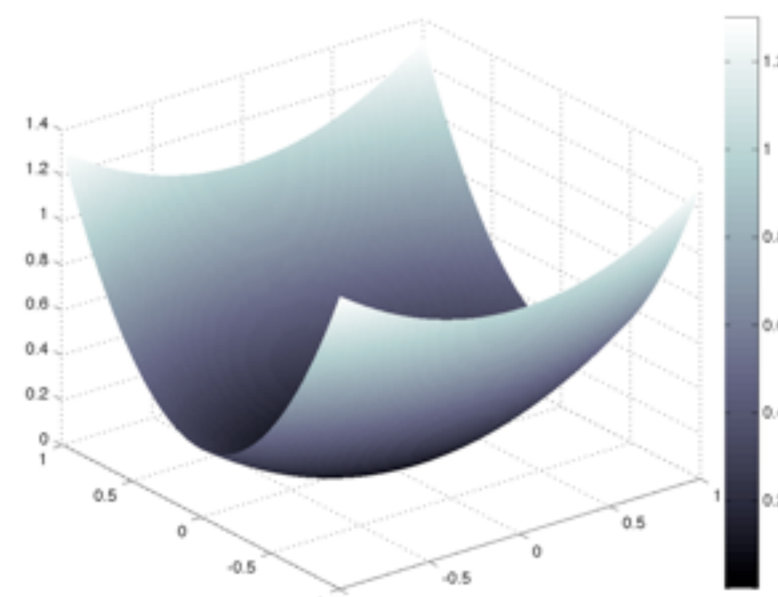
Non-Convexity Strategy

If non-positive curvature in step direction:

$$H \leftarrow H + \xi I$$



Unshifted



Shifted

GLOBAL CONVERGENCE THEORY

GLOBAL CONVERGENCE

Theorem: If

1. f, c **are continuously differentiable and** $f, c, \nabla f$, **and** ∇c **are bounded,**
2. x_k **and** μ_k **are generated by the algorithm,**

then one of the following holds:

(a) $\mu_k = \underline{\mu}$ **for some** $\underline{\mu} > 0$ **for all large** k **and either every limit point** x_* **of** $\{x_k\}$ **corresponds to a KKT point or is an infeasible stationary point;**

(b) $\mu_k \to 0$ **and every limit point** x_* **of** $\{x_k\}$ **infeasible and a stationary point for min**$\|[c(x)]^+\|_1$**.**

(c) $\mu_k \to 0$**, all limit points of** $\{x_k\}$ **are feasible for , and, with** $K_\mu := \{k : \mu_{k+1} < \mu_k\}$**, every limit point** x_* **of** $\{x_k\}_{k \in K_\mu}$ **corresponds to a Fritz-John point at which the MFCQ fails.**

NUMERICAL RESULTS

Numerical Questions

- How to generate inexact subproblem solutions?
- Practical success rate?
- Typical inexactness in accepted solutions?
- How many more iterations?

Implementation

- MATLAB code, C++ version in works
- Subproblem solver: **bqpd**
- Test set: 307 CUTEr/AMPL problems

Simulating Inexactness

Simulating Inexactness

- Ideally: inexact QP solver

Simulating Inexactness

- Ideally: inexact QP solver
- For now: Perturb exact solutions

Simulating Inexactness

- Ideally: inexact QP solver
- For now: Perturb exact solutions
- Reduction Imposed: Measured by κ

Simulating Inexactness

- Ideally: inexact QP solver
- For now: Perturb exact solutions
- Reduction Imposed: Measured by κ

$$\left\| \begin{bmatrix} \mu g + Hd + J\lambda \\ \min([c + J^T d]^+, e - \lambda) \\ \min([c + J^T d]^-, \lambda) \end{bmatrix} \right\| \leq \kappa \left\| \begin{bmatrix} \mu \nabla f + \nabla c \lambda \\ \min([c]^+, e - \lambda) \\ \min([c]^-, \lambda) \end{bmatrix} \right\|$$

Success Rate

	Exact	Inexact		
		$\kappa = 0.01$	$\kappa = 0.1$	$\kappa = 0.5$
Optimal solution found	271	269	272	275
Infeasible stationary point found	4	3	2	2
Iteration limit reached	12	10	11	9
Subproblem solver failure	18	23	20	19

Measuring realized (induced) inexactness

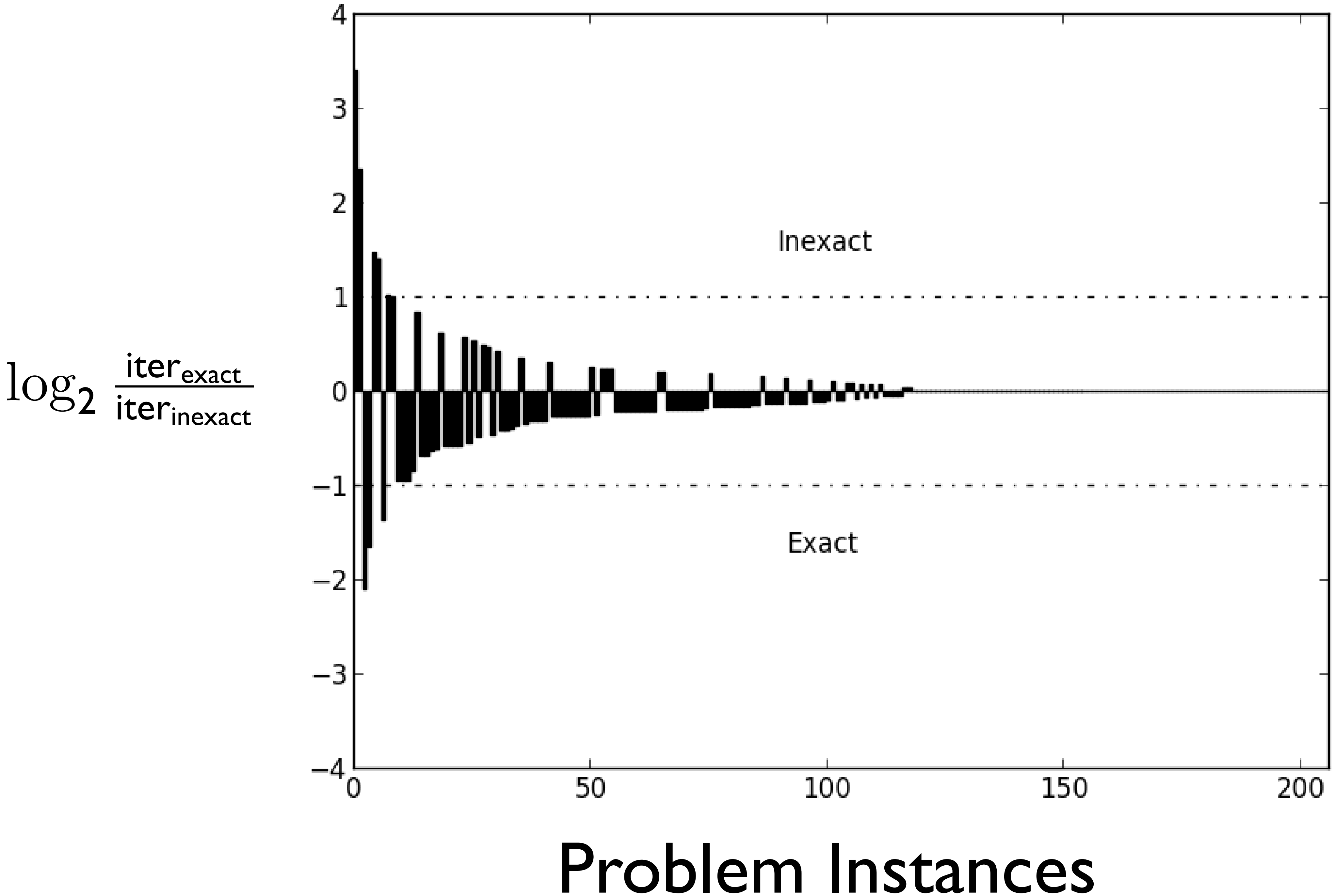
$$\kappa_I := \left\| \begin{bmatrix} \mu g + Hd + A\lambda \\ \min([c + J^T d]^+, e - \lambda) \\ \min([c + J^T d]^-, \lambda) \end{bmatrix} \right\| \Big/ \left\| \begin{bmatrix} \mu \nabla f + (\nabla^2 \mathcal{L}) d + \nabla c \lambda \\ \min([c]^+, e - \lambda) \\ \min([c]^-, \lambda) \end{bmatrix} \right\|$$

(QP residual) (NLP residual)

Typical Inexactness

min	κ	$\kappa_{I,\text{mean}}$	$[0, 10^{-8})$	$[10^{-8}, 10^{-6})$	$[10^{-6}, 10^{-4})$	$[10^{-4}, 10^{-3})$	$[10^{-3}, 0.01)$	$[0.01, 0.1)$	$[0.1, 0.5)$	$[0.5, 1)$	$[1, \infty)$
$\kappa_I(j)$	0.01	`3.5e-03`	0	2	10	7	253	0	0	0	0
	0.1	`2.8e-02`	0	0	2	10	30	232	0	0	0
	0.5	`8.8e-02`	0	0	2	4	23	69	179	0	0
mean	κ	$\bar{\kappa}_{I,\text{mean}}$									
$\bar{\kappa}_I(j)$	0.01	`7.3e-03`	0	0	0	0	254	18	0	0	0
	0.1	`6.9e-02`	0	0	0	0	0	261	13	0	0
	0.5	`3.5e-01`	0	0	0	0	0	1	264	12	0

Inexactness Slowdown?



Key Numerical Results

- Solver:
  - is reliable
  - is able to take large perturbations
  - uses more QP solves

Concluding Remarks

- Inexact methods: Lazily reduce residual
- Penalty steering methods: Update mu
- Both: Trade off more iterations to faster iterations
  - Hopefully!! Still need an active set, inexact QP solver (if it exists)

THANKS.