An LCL Implementation for Nonlinear Optimization

Michael FRIEDLANDER

Mathematics & Computer Science Division
Argonne National Laboratory
Argonne, IL 60439, USA
michael@mcs.anl.gov

Michael SAUNDERS

Systems Optimization Laboratory
Stanford University
Stanford, CA 94305, USA
saunders@stanford.edu

18th International Symposium on
Mathematical Programming
August 18–22, 2003

KNOSSOS

An F90 package for nonlinearly constrained optimization

Kernel: Linearly constrained solver

Nonlinear constraints

Applies LCL method; repeated calls to the LC solver

Linear constraints

Single call to the LC solver

What's in a Name?

- New "home" for reduced gradient (LC) solver in MINOS
- Other LC solvers may be used (eg, SNOPT, other 1st/2nd deriv)

Constrained Optimization

NP

$$\min_{x\in\mathbb{R}^n} \ f(x)$$

$$\text{s.t.} \quad \ell \le \begin{pmatrix} x \\ Ax \\ c(x) \end{pmatrix} \le u$$

- Smooth functions: $f(x)$, $c_i(x)$
- 1st derivatives available
- Sparse Jacobian $J(x)$
- Sparse A

Keeping It Simple

GNP

$$\min_{x\in\mathbb{R}^n} \quad f(x)$$
$$\text{s.t.} \quad c(x) = 0$$
$$x \geq 0$$

LINEARLY CONSTRAINED LAGRANGIAN

Linearly Constrained Lagrangian (LCL)

Robinson ('72), **Rosen & Kreuser** ('72) first proposed LCL

- Good $(x_0, y_0) \Rightarrow$ quadratic convergence (outer iterations)
- No global convergence property

MINOS (Murtagh and Saunders '83)

- Widely used LCL implementation (e.g., AMPL, GAMS, NEOS)
- Modified LCL to improve robustness
- Many heuristics

Minimizing the Augmented Lagrangian

$$\mathcal{L}(x, y, \rho) \stackrel{\text{def}}{=} f(x) - y^T c(x) + \tfrac{1}{2}\rho \|c(x)\|^2$$

BCL (e.g., LANCELOT)	**LCL** (e.g., MINOS)
$\min_x \; \mathcal{L}(x, y_k, \rho_k)$	$\min_x \; \mathcal{L}(x, y_k, \rho_k)$
s.t.	s.t. Linearized c
$x \geq 0$	$x \geq 0$
Globally convergent	Quadratically convergent

A Stabilized LCL Algorithm

Properties

- Globally convergent
- Fast local convergence (using 1st derivs only)

Important

- Handles infeasible subproblems
- Solves LC subproblems inexactly

LCL (Robinson '72)

Linearization: $\bar{c}_k(x) = c(x_k) + J(x_k)(x - x_k)$

Minimizing the Lagrangian subject to linearized c:

---

1. Find $(x_k^*, \Delta y_k^*)$ that solves

$$\begin{array}{ll} \min\limits_{x} & f(x) - y_k^T c(x) \\ \text{s.t.} & \bar{c}_k(x) = 0 \\ & \qquad x \geq 0 \end{array}$$

2. $x_{k+1} \leftarrow x_k^*$
   $y_{k+1} \leftarrow y_k + \Delta y_k^*$

- Good $(x_0, y_0) \Rightarrow$ converges quadratically
- Not globally convergent (in general)

MINOS LCL (Murtagh & Saunders '83)

LC Obj $=$ augmented Lagrangian

---

1. Find $(x_k^*, \Delta y_k^*)$ that locally solves

$$\begin{array}{ll} \min\limits_x & f(x) - y_k^T c(x) + \frac{1}{2}\rho_k \|c(x)\|^2 \\ \text{s.t.} & \bar{c}_k(x) = 0 \\ & \qquad x \geq 0 \end{array}$$

2. Restrict $\|\Delta x\|$, $\|\Delta y\|$
3. Adjust penalty parameter ρ_k

Drawbacks

- Heuristic step length restriction
- Penalty ρ_k improves robustness, but no guarantees
- Subproblems may be infeasible

Toward Stabilized LCL

When x_k is far from x_*,

$$\bar{c}_k(x) = 0 \quad \text{may be poor approximation to} \quad c(x) = 0$$

- Current Jacobian may be ill-conditioned
  $\Rightarrow$ $\|\Delta x\|$ can be large

- Linearization may be infeasible
  $\Rightarrow$ subproblem can be ill-defined

Suggests the need for an elastic subproblem

An Elastic LC Subproblem

min aug Lagrangian

subj to bounds

$$\min_x \quad f(x) - y_k^T c(x) + \tfrac{1}{2}\rho_k \|c(x)\|^2$$

$$\text{s.t.} \qquad x \geq 0$$

An Elastic LC Subproblem

min	aug Lagrangian	
subj to	Linearized c	bounds

$$\min_{x} \quad f(x) - y_k^T c(x) + \tfrac{1}{2}\rho_k \|c(x)\|^2$$
$$\text{s.t.} \quad \bar{c}_k(x) = 0, \qquad x \geq 0$$

An Elastic LC Subproblem

min aug Lagrangian + ℓ_1 penalty function

subj to Linearized c + elastic vars, bounds

$$\min_{x,v,w} \quad f(x) - y_k^T c(x) + \tfrac{1}{2}\rho_k \|c(x)\|^2 + \sigma_k e^T (v + w)$$

$$\text{s.t.} \quad \bar{c}_k(x) + v - w = 0, \qquad v, w, x \geq 0$$

An Elastic LC Subproblem

min	aug Lagrangian + ℓ_1 penalty function
subj to	Linearized c + elastic vars, bounds

$$\begin{array}{ll} \min\limits_{x,v,w} & f(x) - y_k^T c(x) + \frac{1}{2}\rho_k \|c(x)\|^2 + \sigma_k e^T(v+w) \\ \text{s.t.} & \bar{c}_k(x) + v - w = 0, \qquad v, w, x \geq 0 \end{array}$$

$$\Updownarrow$$

$$\begin{array}{ll} \min\limits_{x} & f(x) - y_k^T c(x) + \frac{1}{2}\rho_k \|c(x)\|^2 + \sigma_k \|\bar{c}_k(x)\|_1 \\ \text{s.t.} & x \geq 0 \end{array}$$

An Elastic LC Subproblem

min	aug Lagrangian + ℓ_1 penalty function
subj to	Linearized c + elastic vars, bounds

$$\min_{x,v,w} \quad f(x) - y_k^T c(x) + \tfrac{1}{2}\rho_k \|c(x)\|^2 + \sigma_k e^T(v+w)$$

$$\text{s.t.} \quad \bar{c}_k(x) + v - w = 0, \qquad v, w, x \geq 0$$

- Subproblems are always feasible
- Parameter σ_k used to control the multiplier estimates because

$$\|\Delta y_k^*\|_\infty \leq \sigma_k$$

- For σ_k sufficiently large, solution is same as MINOS subprob
- v, w introduced linearly

The Stabilized LCL (sLCL) Method

1. Solve "to within" $\texttt{optTol}_k \searrow 0$:

$$(x_k^*, \Delta y_k^*) \leftarrow \text{Elastic LC subproblem } (x_k, y_k, \rho_k, \sigma_k)$$

2. If $\|c(x_k^*)\| \le \eta_k \searrow$

Decrease ρ_k (finitely often)

Reset σ_k (sufficiently large)

Update iterates:

$$x_{k+1} \leftarrow x_k^*,$$

$$y_{k+1} \leftarrow \boxed{y_k - \rho_k c(x_k^*)} + \Delta y_k^*$$

else

Increase ρ_k, decrease σ_k

Reject x_k^*

- ρ_k and σ_k work together to ensure global convergence
- sLCL dynamically adjusts accuracy of subproblem solves

CONVERGENCE PROPERTIES

sLCL — Main Results (Friedlander & Saunders '02)

Global Convergence (closely related to BCL theory)

- All limit points are 1st-order stationary
  Can be 2nd-order stationary (with 2nd deriv LC solver)
- Inexact subproblem solutions ($\texttt{OptTol}_k \searrow 0$)
- ρ_k remains bounded

Local Convergence

- Quadratic rate (outer iterations, $\texttt{OptTol}_k \searrow$ fast enough)
- Eventually $v_k = w_k = 0$ for k large enough
  (ie, asymptotic equivalence to MINOS)

Infeasible Problems

$\rho_k \to \infty$

$\|y_k\|/\rho_k$ bounded

and

$$\begin{array}{ll} \min\limits_{x} & \frac{1}{2}\|c(x)\|_2^2 \\ \text{s.t.} & x \ge 0 \end{array}$$

IMPLEMENTATION

KNOSSOS

A Useful Reformulation

Departure from linearity: $d_k(x) = c(x) - \bar{c}_k(x)$

$$\mathcal{L}_k(x) = f(x) - y_k^T c(x) + \tfrac{1}{2}\rho_k \|c(x)\|^2$$

$$\Uparrow$$

$$\bar{c}_k(x) + v - w = 0 \qquad \begin{pmatrix}\text{Subproblem's} \\ \text{linear} \\ \text{constraints}\end{pmatrix}$$

$$\Downarrow$$

$$\mathcal{M}_k(x, v, w) = f(x) - y_k^T \big(d_k(x) - v + w\big) + \tfrac{1}{2}\rho_k \|d_k(x) - v + w\|^2$$

The Reformulated LC Subproblem

$$\boxed{\begin{array}{ll} \min\limits_{x,v,w} & \mathcal{M}_k(x,v,w) + \sigma_k e^T(v+w) \\ \text{s.t.} & \bar{c}_k(x) + v - w = 0, \qquad x,v,w \ge 0 \end{array}} \longrightarrow \begin{pmatrix} x_k^* \\ y_k^* \end{pmatrix}$$

σ_k no longer bounds subproblem multipliers! But, set

$$\Delta y_k^* = y_k^* - \big(y_k - \rho_k c(x_k^*)\big)$$

1. $\begin{pmatrix} x_k^* \\ \Delta y_k^* \end{pmatrix}$ is a KKT point of $\begin{cases} \min\limits_{x,v,w} \mathcal{L}_k(x) + \sigma_k e^T(v+w) \\ \text{s.t. } \bar{c}_k(x) + v - w = 0, \quad x,v,w \ge 0 \end{cases}$

2. $\|\Delta y_k^*\|_\infty \le \sigma_k$

3. sLCL multiplier update: $\qquad y_{k+1} \leftarrow y_k - \rho_k c(x_k^*) + \Delta y_k^* = y_k^*$

Stabilized LCL — Implementation

Fortran 90

- Threadsafe design (LC solver permitting)
- LC solver "isolated"
- Dynamic memory

LC subproblems

- Solve with MINOS or SNOPT
- Hot start on major 2, 3, 4,. . .
  (keeping quasi-Newton Hessian approximation)
- Early termination (via `Iteration Limit` or `Major Opt Tol`)

Proximal Point

- Solve
  $$\boxed{\begin{array}{l} \min_x \frac{1}{2}\|x - x_0\|_2^2 \\ \text{s.t. } \ell_A \le Ax \le u_A, \quad \ell_x \le x \le u_x \end{array}}$$
  with loose tols
- $\ell_A \le Ax \le u_A, \quad \ell_x \le x \le u_x$ feasible thereafter

CUTEr Problems

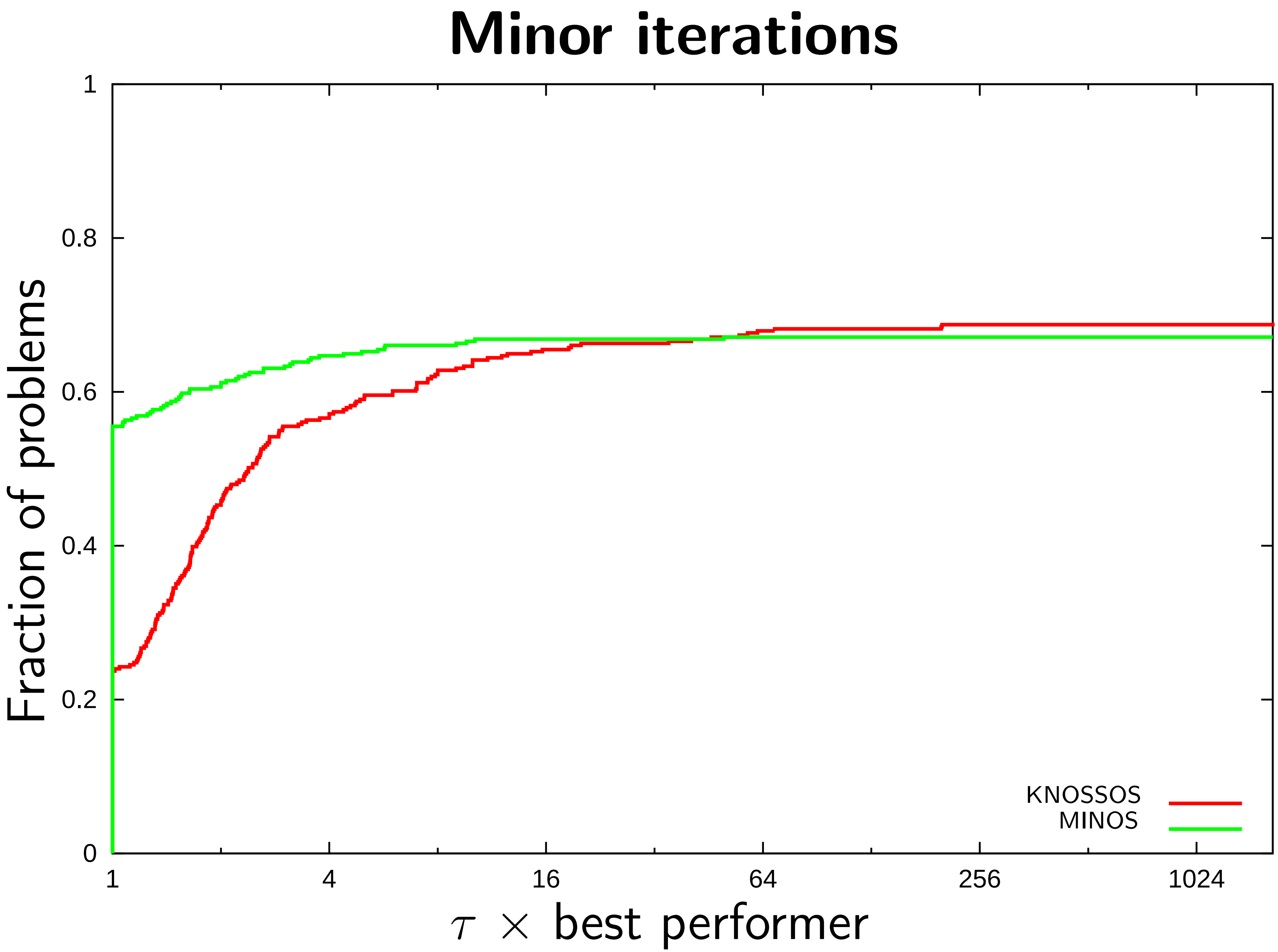
Nonlinearly constrained subset from CUTEr collection

Size $(n+m)$	# of problems
1 — 100	204
100 — 1000	40
1000 — 10000	83
10000 — 27608	44
Total	371

Comparison

- MINOS LCL v. KNOSSOS (with MINOS reduced gradient)
- Minor iterations
  - $\Rightarrow$ Roughly one (f, c) evaluation; or
  - $\Rightarrow$ Phase 1 simplex for feasibility

Performance Profiles I



Observations

Failures often related to elastic variables:

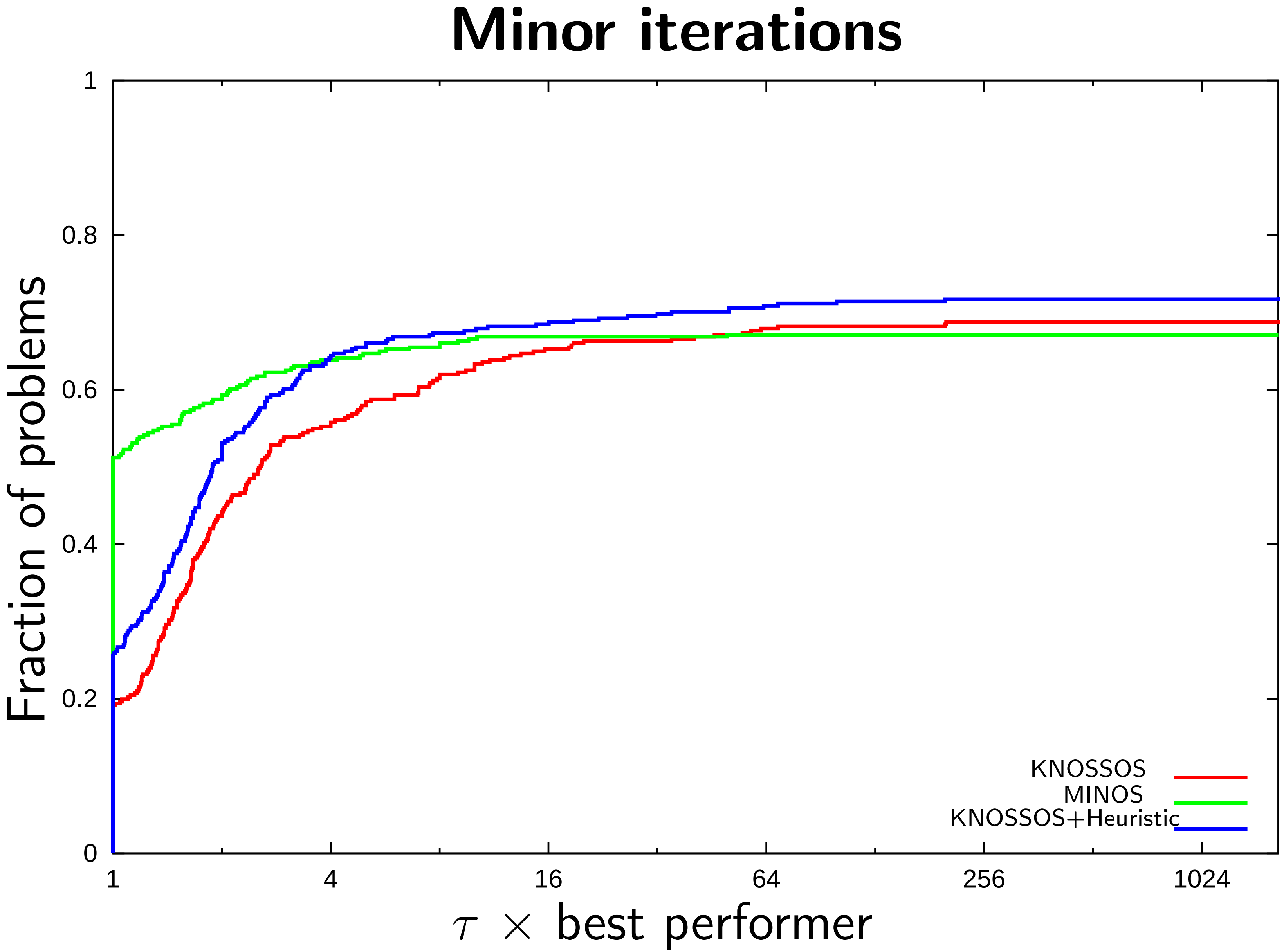
- $v, w \not\to 0$ and converge to spurious points
- Unbounded subproblems

Robinson LCL $\left(+\frac{1}{2}\rho_k\|c(x)\|^2\right)$ seems robust to singular Jacobians

Heuristic

1. Start with $0 \le v, w \le 0$
2. Juggle penalty params as usual: $\begin{cases} \rho_k & \text{for } \|c(x)\|^2 \\ \sigma_k & \text{for } \|\bar{c}_k(x)\|_1 \end{cases}$
3. If $\left\{\begin{array}{c} \text{Subprob infeasible} \\ \text{or} \\ \|\Delta y_k^*\|_\infty > \sigma_k \end{array}\right.$ then $0 \le v, w \le +\infty$
4. Carry on with sLCL

Performance Profiles II



EXTENSION

An SQP Variant

QP approximation to kth elastic LCL subproblem:

$$
\begin{array}{ll}
\min\limits_{\Delta x, v, w} & g_k^T \Delta x + \Delta x^T H_k \Delta x + \sigma_k e^T (v + w) \\
\text{s.t.} & c_k + J_k \Delta x + v - w = 0 \\
 & x_k + \Delta x, v, w \geq 0
\end{array}
\longrightarrow \begin{pmatrix} \Delta x_k^* \\ \Delta y_k^* \end{pmatrix}
$$

sLCL theory might guide update rules for:

- ρ_k, σ_k (penalty/elastic params)
- Inexact QP subproblem solves (Murray & Prieto '94)

Some similarities to SNOPT (Gill et al. '02)

	Elastic vars	Quasi-Newton Hessian
sLCL/SQP:	$\\|\Delta y_k^*\\|_\infty \leq \sigma_k$	$H_L + \rho J^T J$
SNOPT:	$\\|y_k + \Delta y_k^*\\|_\infty \leq \sigma_k$	H_L or $H_L + \rho J^T J$

CONCLUSIONS

Work in Progress

- How to best control elastic variables?
- How to recover from/prevent unbounded subproblems?
- Can the heuristic be integrated into the theory?
- What is the best multiplier update (e.g., trimming)?
- Need a 2nd derivative linearly constrained solver

Thanks!

- The CUTEr team: Gould, Orban, & Toint
- Daniel Friedlander

— EXTRA SLIDES —

By the Numbers

Size $(n+m)$	No. of problems	Number Optimal		
		MINOS	Knossos	Knossos+
1 — 100	204	176	185	182
100 — 1000	40	28	28	27
1000 — 10000	83	31	25	36
10000 — 27608	44	17	21	23
Total	371	252	259	268