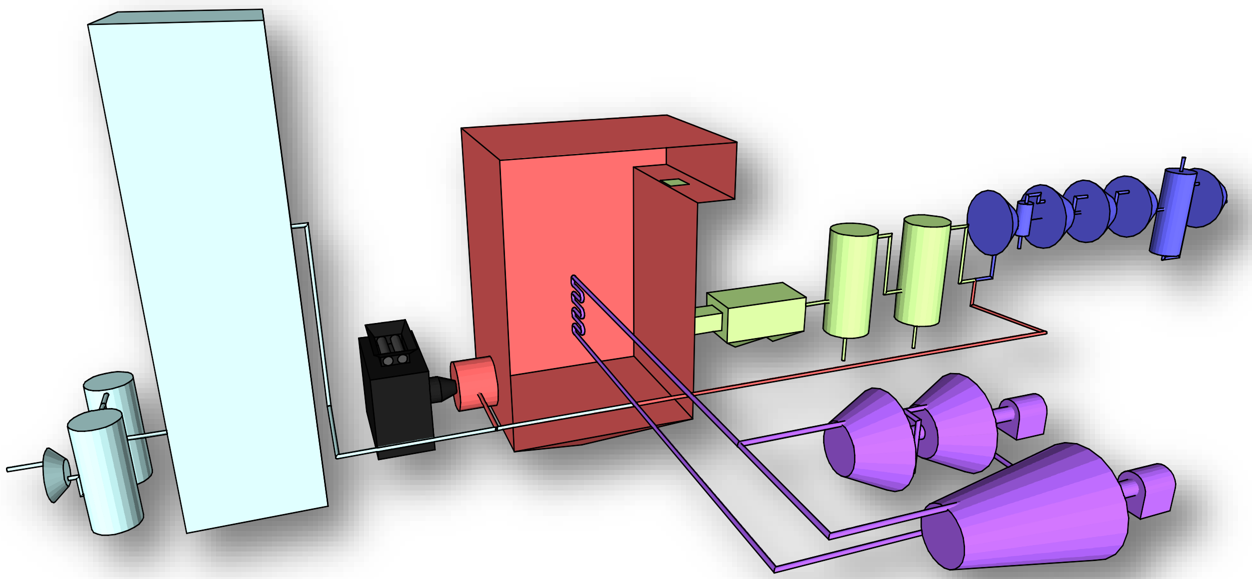
Trust region methods for optimization with reduced order models embedded in chemical process models

John P. Eason and Lorenz T. Biegler
Carnegie Mellon University

ISMP 2015
17 July 2015

Chemical process optimization



When developing new chemical processes,
most operations are well understood (accurate, open models)

New technologies may require simulation (e.g. CFD)

This leads to grey-box constraints in optimization

Introduction

- We will consider problems of the following form:

$$\begin{aligned} \min_{x} \quad & f(x, d(x)) \\ s.t. \quad & c(x, d(x)) = 0 \qquad\qquad x \in \mathbb{R}^n \\ & g(x, d(x)) \leq 0 \end{aligned}$$

where $d : \mathbb{R}^n \to \mathbb{R}^p$ is a "simulation function," i.e. expensive, possibly with derivatives unavailable

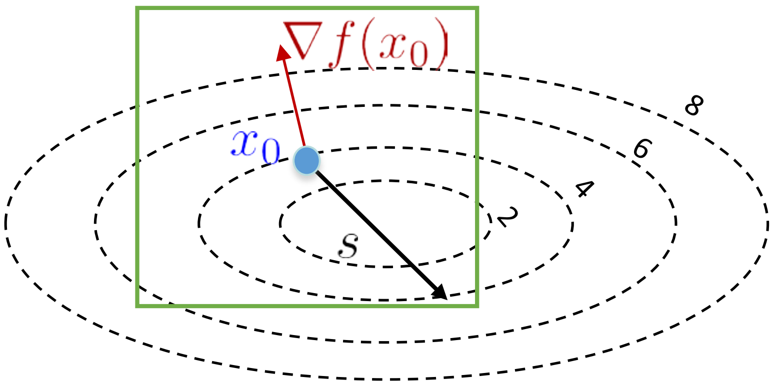
Note: In our applications of interest, $n \gg p$

Methods

- In the engineering literature, common approach is to build a reduced model (RM)
  - Neural networks, Kriging interpolation, Polynomial regression
- Not guaranteed to find optima of true problem
- To get convergence properties, use ideas from derivative free optimization
- RM ~ model function from DFO literature
  - Model-based methods allow us to use the derivative information from the open model

Trust region methods

- Build RMs that we "trust" in a local region
  - Satisfy certain condition on accuracy
- Use RM, $m_0(s)$, to generate a step s
- Adaptively adjust based on accuracy of the step
  - Guaranteed convergence



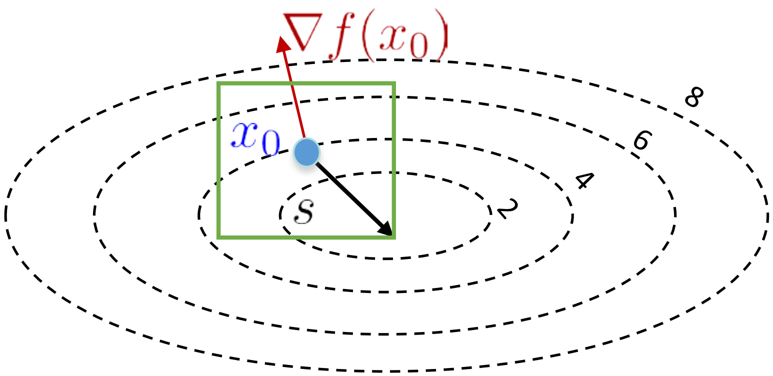
$$\min_s \; m_0(s)$$
$$s.t. \; \|s\| \le \Delta_0$$

Evaluate $f(x_0+s)$ and check if sufficiently reduced from $f(x_0)$

If no improvement! Shrink trust region Δ_0

Trust region methods

- Build RMs that we "trust" in a local region
  - Satisfy certain condition on accuracy
- Use RM to generate a step *s*
- Adaptively adjust based on accuracy of the step
  - Guaranteed convergence



New step *s* within smaller trust region

Evaluate $f(x_0+s)$

$$\frac{f(x_0+s)-f(x_0)}{m_0(s)-m_0(0)} \geq \eta \in (0,1)$$

Sufficiently decreased objective

Conditions on reduced models

- The key to convergence is the fully linear property:

  there exist finite κ_f and κ_g such that for all iterations k,

$$\|d(x) - r(x)\| \leq \kappa_f \Delta_k^2, \quad \|\nabla d(x) - \nabla r(x)\| \leq \kappa_g \Delta_k$$

- As trust region vanishes, function values and gradients approach original model
- Any type of RM may be used satisfying this property

Handling constraints

- We will use a trust region filter method

  Fletcher, R., Gould, N. I., Leyffer, S., Toint, P. L., & Wächter, A. (2002). Global convergence of a trust-region SQP-filter algorithm for general nonlinear programming. *SIAM Journal on Optimization*, *13*(3), 635-659.

  with extension to derivative free optimization:

  Conn, A. R., Scheinberg, K., & Vicente, L. N. (2009). *Introduction to derivative-free optimization*. SIAM.

- General NLP subproblems rather than QP

A quick reformulation

- Introduce new variables y to isolate the complicating constraints

$$
\begin{aligned}
\min_{x,y} \quad & f(x,y) \\
s.t. \quad & c(x,y) = 0 \\
& g(x,y) \leq 0 \\
& \boxed{y = d(x)}
\end{aligned}
$$

$d(x)$ will be approximated by RM $r(x)$

- Infeasibility measures:

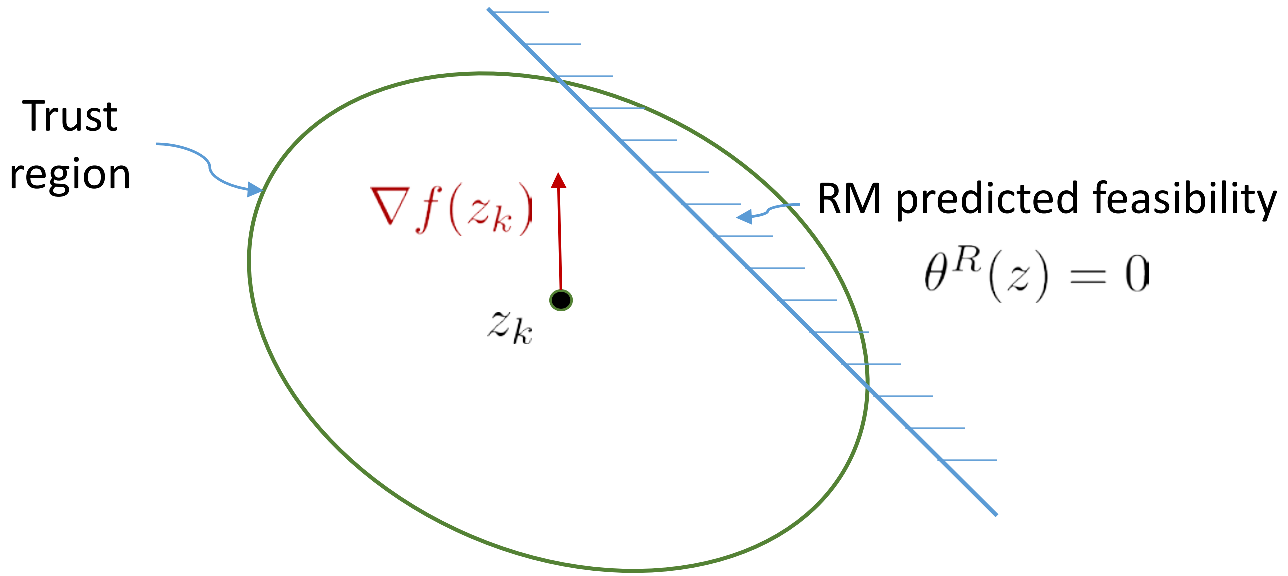
$$\theta(z) = \|d(x) - y\|$$

$$\theta^R(z) = \|r(x) - y\|$$

$$z^T = [x^T \; y^T]$$

Generating a trial step

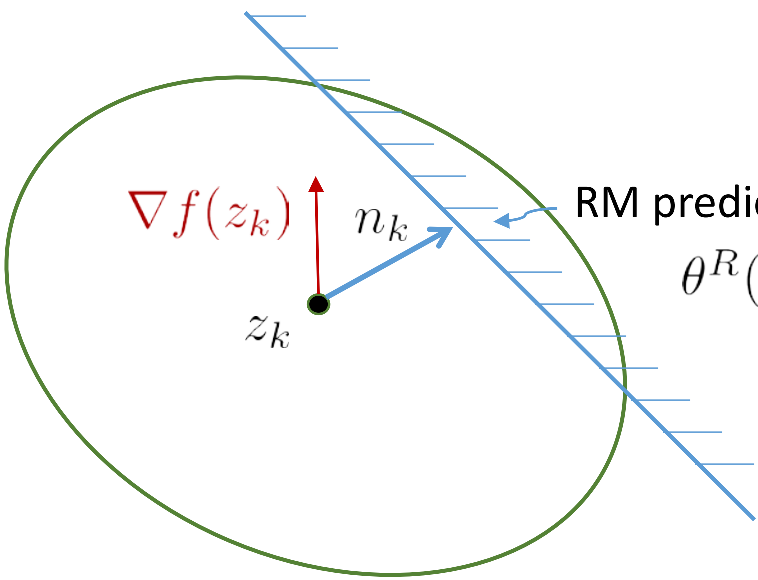
- Need to improve both feasibility and objective
- Separate into normal and tangential subproblems



Normal step

$$\min_{n_k} \quad \theta^R(z_k + n_k)$$
$$s.t. \quad c(z_k + n_k) = 0$$
$$g(z_k + n_k) \leq 0$$
$$\|n_k\| \leq \Delta_k$$

- Find a nearby feasible point
- If predicted feasibility is too far, go to restoration phase



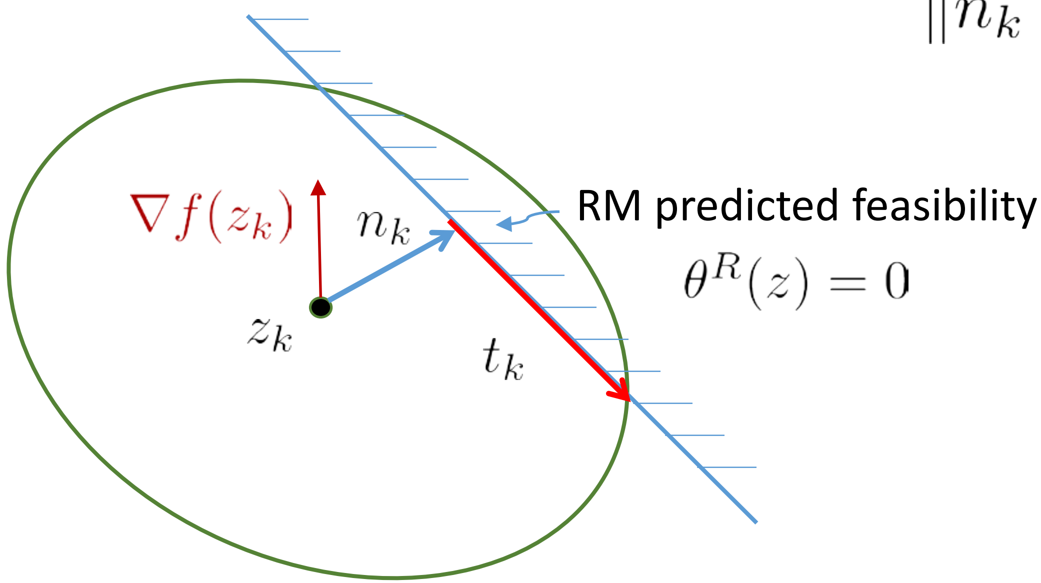
We require:

$$\exists \kappa_{usc}, \delta_n \text{ such that } \theta_k \leq \delta_n \implies \|n_k\| \leq \kappa_{usc}\theta_k.$$

Tangential step

$$\min_{t_k} \quad f(z_k + n_k + t_k)$$
$$s.t. \quad c(z_k + n_k + t_k) = 0$$
$$g(z_k + n_k + t_k) \leq 0$$
$$\theta^R(z_k + n_k + t_k) = 0$$
$$\|n_k + t_k\| \leq \Delta_k$$

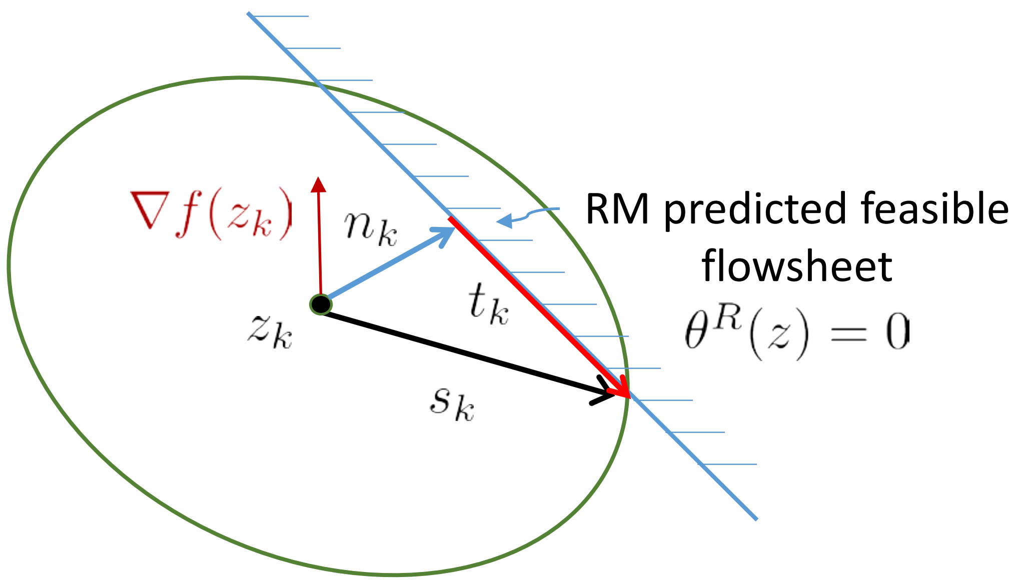
- Minimize objective, do not increase θ^R



- Satisfy fraction of Cauchy decrease condition

Total trial step

- The proposed step for iteration k: $s_k = n_k + t_k$



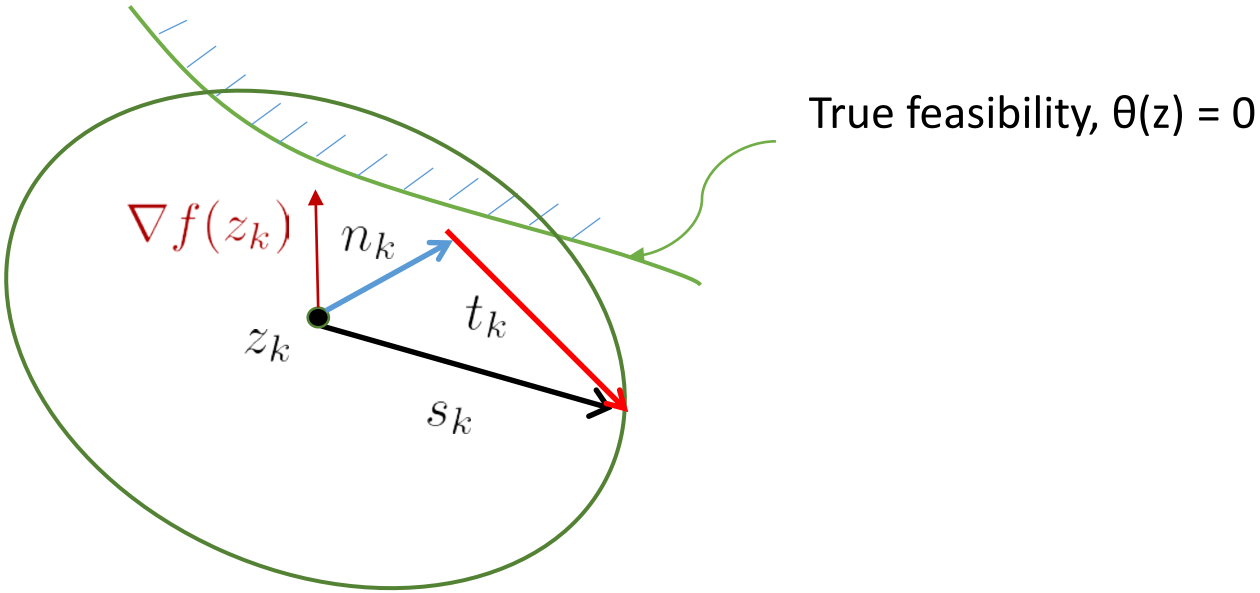
- We've generated a step where we've reduced

$$\theta^R(z) = \|y - r(x)\|$$

and also improved the objective function *f(z)*

- Now evaluate *d(x)* and determine whether the RM solution actually made progress in reducing

$$\theta(z) = \|y - d(x)\|$$



Filter method

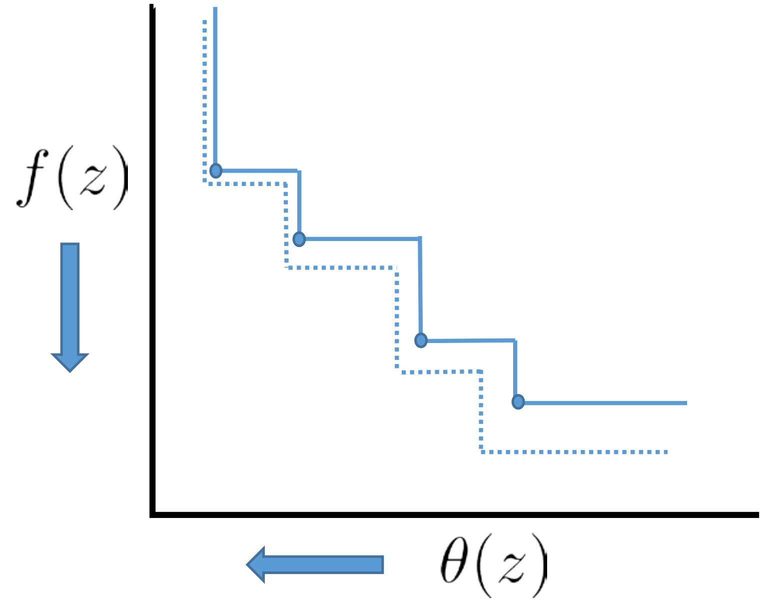
- Store (θ_k, f_k) at allowed iterates
- If unacceptable to filter, decrease trust region
- If switching condition

$$f(z_k) - f(z_k + s_k) \geq \kappa_\theta \theta(z_k)^{\gamma_s}$$

$$\kappa_\theta \in (0,1),\ \gamma_s > \frac{1}{2}$$

is satisfied, possibly increase trust region

- Else, adjust trust region by ratio test on θ



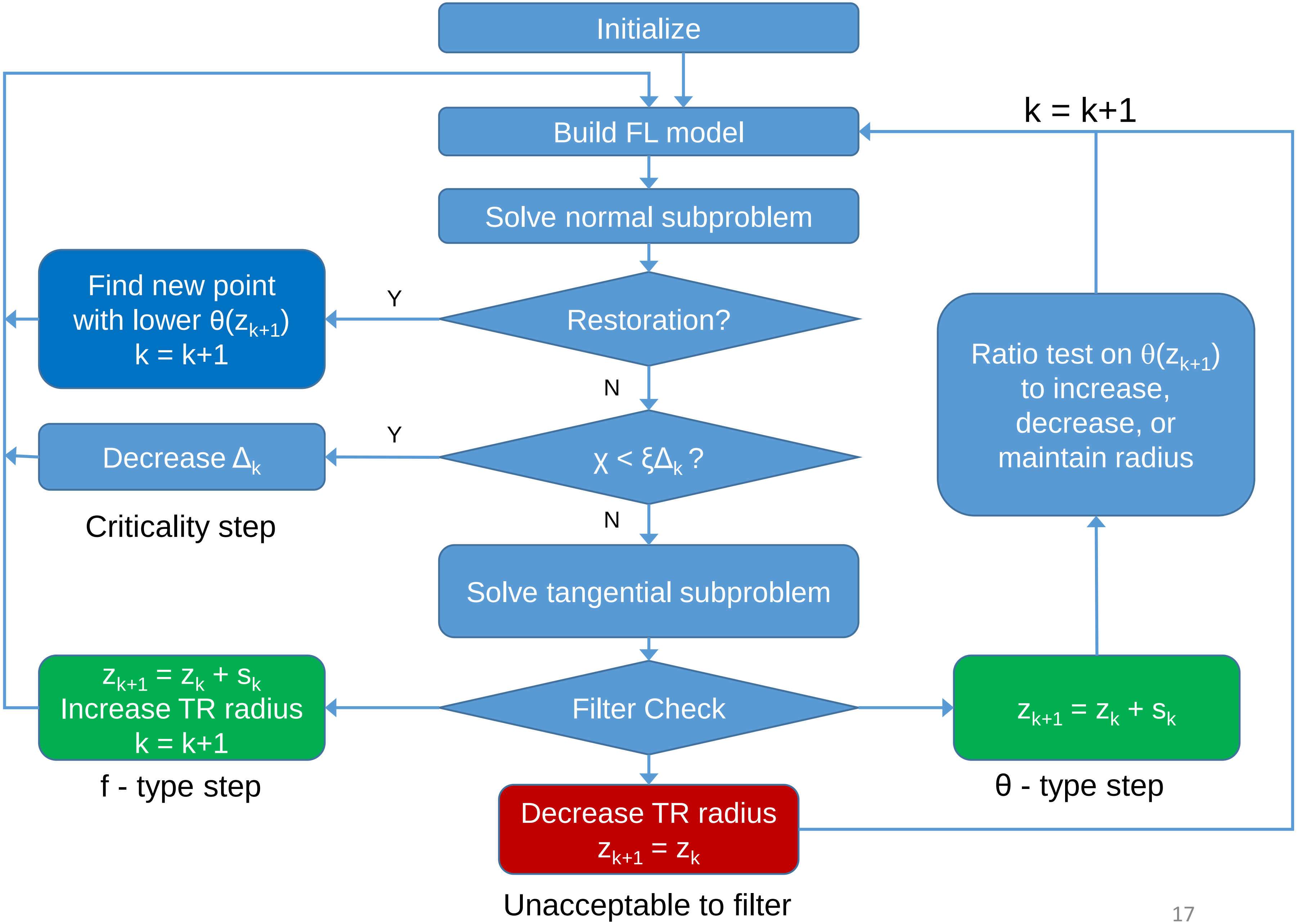
$$\left\{ \begin{array}{c} \theta(z_k + s_k) \leq (1 - \gamma_\theta)\theta_j \\ \text{OR} \\ f(z_k + s_k) \leq f_j - \gamma_f \theta_j \end{array} \right\}$$

$$\forall (\theta_j, f_j) \in \mathcal{F} \cup (\theta_k, f_k)$$

Fletcher et al., SIAM J. Optim, 13, 3, (2002)

Restoration

- If either of the following hold, then call restoration
  a) $\theta^R(z_k + n_k) > 0$
  b) $\|n_k\| \geq \kappa_\Delta \Delta_k \min[1, \kappa_\mu \Delta_k^\mu]$
- Restoration must return a new point z_{k+1} such that
  a) Restoration is not called at iterate k+1
  b) z_{k+1} is acceptable to $\mathcal{F} \cup (\theta_k, f_k)$
- Improving feasibility will satisfy these conditions
- We use tailored algorithms for chemical process simulations to converge constraints in restoration.


Initialize
Build FL model
k = k+1
Solve normal subproblem
Find new point with lower θ(z_{k+1}) k = k+1
Y
Restoration?
N
Ratio test on θ(z_{k+1}) to increase, decrease, or maintain radius
Decrease Δ_k
Y
χ < ξΔ_k ?
N
Criticality step
Solve tangential subproblem
z_{k+1} = z_k + s_k Increase TR radius k = k+1
Filter Check
z_{k+1} = z_k + s_k
f - type step
θ - type step
Decrease TR radius z_{k+1} = z_k
Unacceptable to filter

Convergence Properties

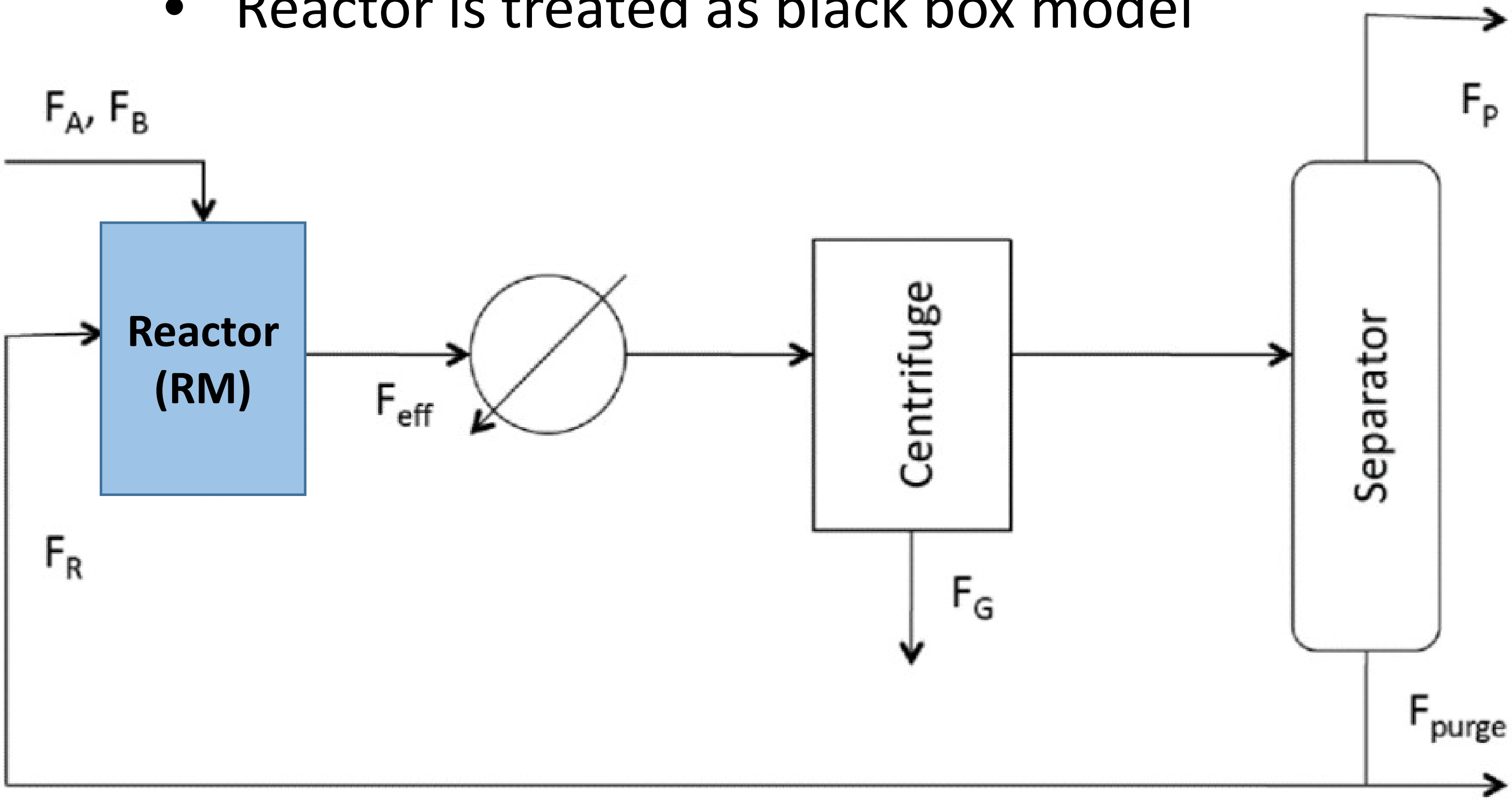
- Standard assumptions (Lipschitz functions etc.)
- Exact derivatives: Proof based on Fletcher et al (2002).
- Derivative free: extended using analysis from Conn, Scheinberg and Vicente (2010).
- A couple key differences:
    - Trust region must go to zero
    - Criticality measure

$$\begin{aligned} \chi(x_k) = \quad & \left| \min_{d} \quad \nabla f(x_k + n_k)^T d \right| \\ \text{s.t.} \quad & A_k^r (x_k + n_k) d = 0 \\ & x_l \leq x_k + n_k + d \leq x_u \\ & \|d\| \leq 1 \end{aligned}$$

- Result: global liminf convergence to 1st order KKT point

Williams-Otto process

- Simple flowsheet optimization problem
- Reactor is treated as black box model

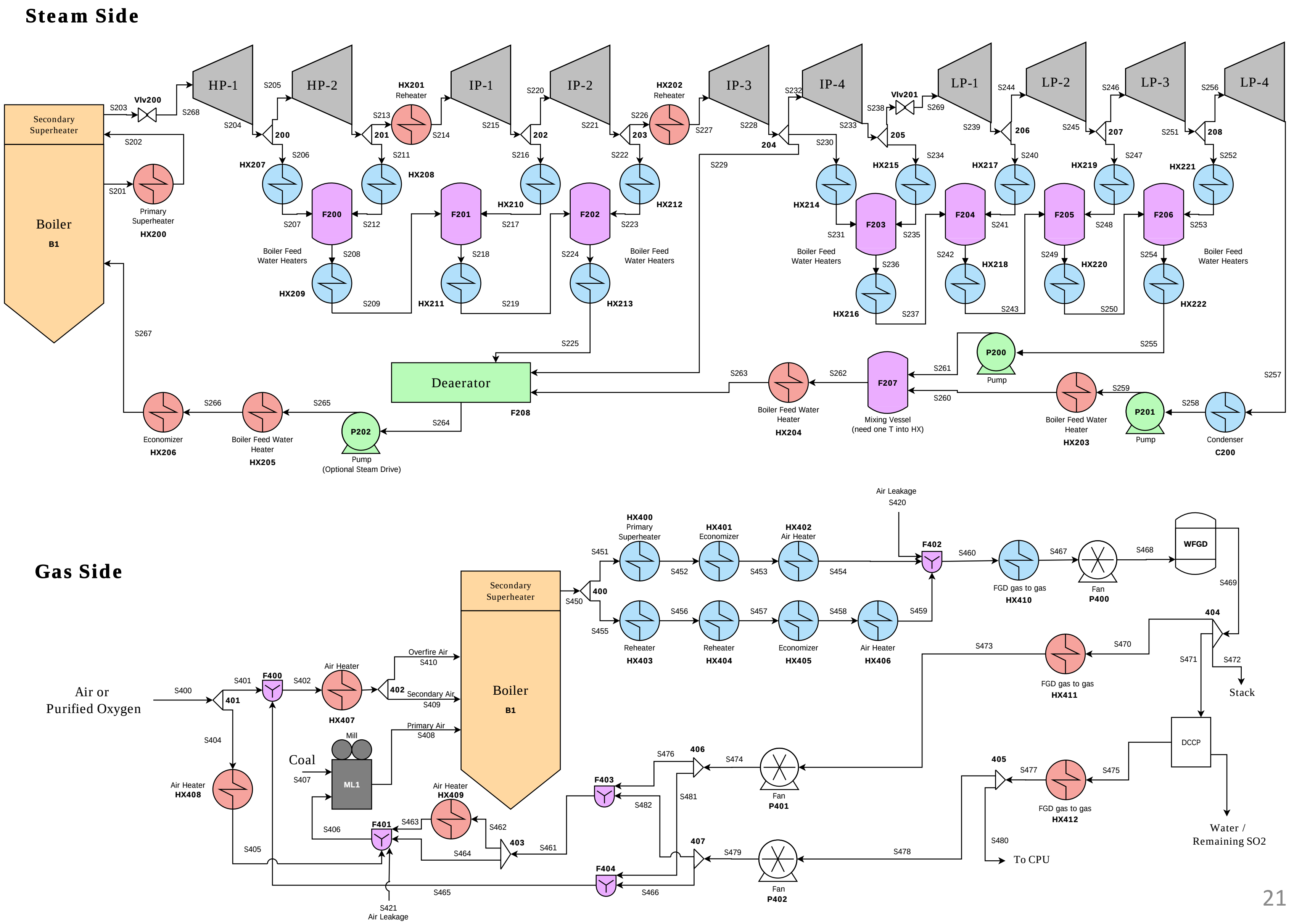


Results

ROM:	Kriging	linear	linear	SQP
note:	dace	interpolation	finite diff	finite diff
Iterations:	148	302	7	15
d(x) calls:	1621	2114	49	210

- Sometimes the derivatives of d(x) are available
- We can use a (slightly modified) version of the algorithm to reduce the simulation calls
  - Subproblems exploit the cheap derivatives of the closed-form portions of the model

Oxycombustion steam cycle



Case Studies: *Air-fired Steam Cycle*

maximize **Thermal Efficiency**

s.t. Steam cycle connectivity
Heat exchanger model
Pump model
Fixed isentropic efficiency turbine model
Hybrid boiler model with fixed fuel rate
Heat integration model
Steam thermodynamics

Using trust region method

Solved in GAMS 24.2.1 with CONOPT 3
Trust region algorithm in MATLAB R2013a

Case Studies: *Air-fired Steam Cycle*

- **Gross electrical efficiency: 46.04%** (HHV)
- Optimized steam extraction and feed water heating
- Ongoing work: <u>assumption refinement</u>

Solution time:	**167.1 minutes**
Total boiler simulations:	247 (run on 4 cores)
HP turbines work	126.1 MW
IP turbines work	309.7 MW
LP turbines work	347.4 MW
Fuel rate (HHV)	1325.5 MW
Steam exit temperature	863 K
Steam exit pressure	350 bar

Conclusions

- A trust region filter method is presented for integrating grey box constraints in larger NLPs
- Promising performance on some chemical process applications
- Future work:
  - Improved RM management – Can we do better than finite differences? (quadratic updates?)
  - Benchmark algorithm with more examples

Thanks to : Alex Dowling, Jinliang Ma

CCSI
Carbon Capture Simulation Initiative

Case Study: *Oxycombustion Steam Cycle*

maximize **Thermal Efficiency**

s.t. Steam cycle connectivity

Heat exchanger model

Pump model

Fixed isentropic efficiency turbine model

Hybrid boiler model with fixed fuel rate

Heat integration model

Steam thermodynamics

Pollution control models

Using trust region method

Solved in GAMS 24.2.1 with CONOPT 3

Trust region algorithm in MATLAB R2013a

Case Study: *Oxycombustion Steam Cycle*

- **Gross electrical efficiency: ????%** (HHV)
- Optimized steam extraction, recycle strategy

Solution time:	
HP turbines work	
IP turbines work	
LP turbines work	
Fuel rate (HHV)	
Steam exit temperature	
Steam exit pressure	
FEGT	