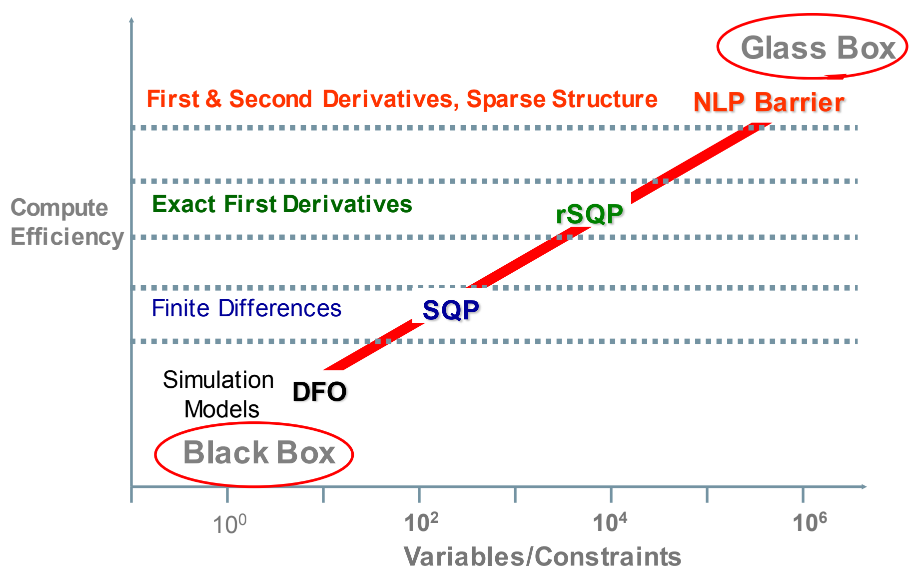
Nonlinear Programming Strategies for Multi-scale Chemical Process Optimization

L. T. Biegler
John Eason, Alex Dowling

January, 2016

How Process Optimization? Environments and NLP Solvers



Process Optimization: Equation-Oriented (Glass Box) Formulation

Min Overall Objective
s.t. Conservation Laws
Performance Equations
Constitutive Equations
Process/Product Specifications

Minimize Gibbs Free Energy (Vapor Liquid Equilibrium)

Minimize Gibbs Free Energy (Reactor Model)

Minimize Utilities Via Heat Integration (LP)

→ Consider MPCCs derived from Bi-level Optimization

Distillation Optimization (MESH Model)

Minimize Reboiler Duty

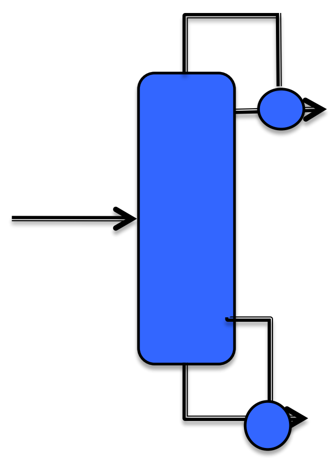
s.t. Top/ Bottom Product Specifications

Mass

Equilibrium

Heat

Summation



- Three component Separation (nC6, nC7, nC9)
- Fixed: 20 trays, Feed = 10
- Reboil and Reflux as decisions

Distillation Optimization (MESH Model)

Minimize Reboiler Duty

s.t. Top/ Bottom Product Specifications

$$(L_i + DL + rd)x_{ij} + DVy_{ij} = V_{i-1}y_{i-1,j} \quad i \in CON$$

$$L_i x_{ij} + V_i y_{ij} = L_{i+1}x_{i+1,j} + V_{i-1}y_{i-1,j} + \sum_k f_{ik} Fd_k xf_{dj} + g_i \cdot rd \cdot x_{ij} \quad i \in COL$$

$$Bx_{ij} + V_i y_{ij} = L_{i+1}x_j + \sum_k f_{ik} Fd_k xf_{dj} \quad i \in REB$$

$$y_{ij} = \beta_i K_{ij} x_{ij}$$

$$-s_i^V \le \beta_i - 1 \le s_i^L$$

$$0 \le L_i \perp s_i^L \ge 0$$

$$0 \le V_i \perp s_i^V \ge 0$$

$$(L_i + DL + rd)hl_i + DV \cdot hv_i = V_{i-1}hv_{i-1} + Q_c \quad i \in CON$$

$$L_i hl_i + V_i hv_i = L_{i+1}hl_{i+1} + V_{i-1}hv_{i-1} + \sum_k f_{ik} Fd_k shf_{dj} + g_i \cdot rd \cdot hl_i \quad i \in COL$$

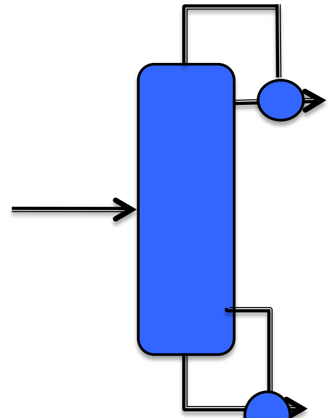
$$B \cdot hl_i + V_i hv_i = L_{i+1}hl_{i+1} + \sum_k f_{ik} Fd_k shf_{dj} + Q_H \quad i \in REB$$

$$\sum_J x_{ij} - \sum_J y_{ij} = 1$$

$$R_{total} = R \cdot D$$

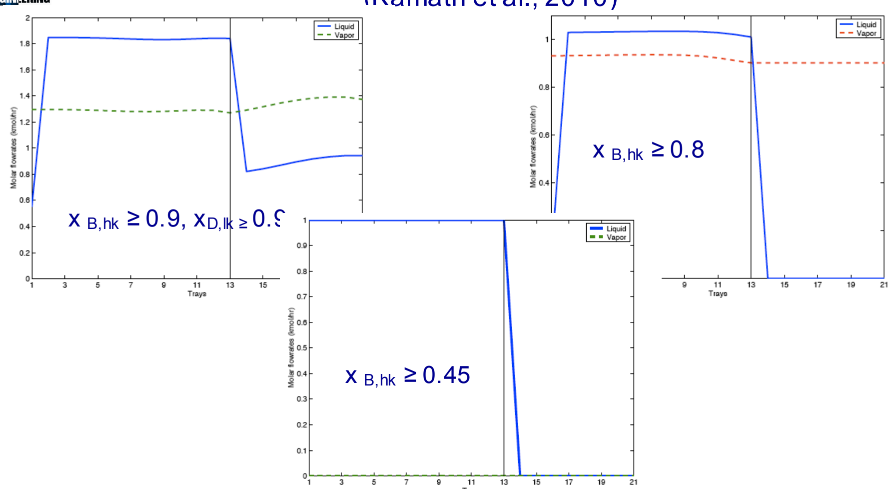
$$L_{N_{max}} = (1 - rdf)R_{total}$$

$$rd = rdf \cdot R_{total}$$



- Three component Separation (nC6, nC7, nC9)
- Fixed: 20 trays, Feed = 10
- Reboil and Reflux as decisions

Distillation Results – Min Heat Duty
(Kamath et al., 2010)



Case	Product purity constraints	Optimal reboiler duty (MW)
1	$x_{\text{top},n\text{-hexane}} \ge 0.9$, $x_{\text{bottom},n\text{-nonane}} \ge 0.9$	28.14
2	$x_{\text{bottom},n\text{-nonane}} \ge 0.8$	19.337
3	$x_{\text{bottom},n\text{-nonane}} \ge 0.45$	0.0

Equation-Oriented Optimization: Air Separation Units (Dowling, B., 2014)

Boiling pts (1 atm.)
- Oxygen: 90 K
- Argon: 87.5 K
- Nitrogen: 77.4 K

Feedstock (air) is free: dominant cost is compression energy

Multicomponent distillation with tight heat integration

Nonideal Phase Equilibrium: Cubic Equations of State

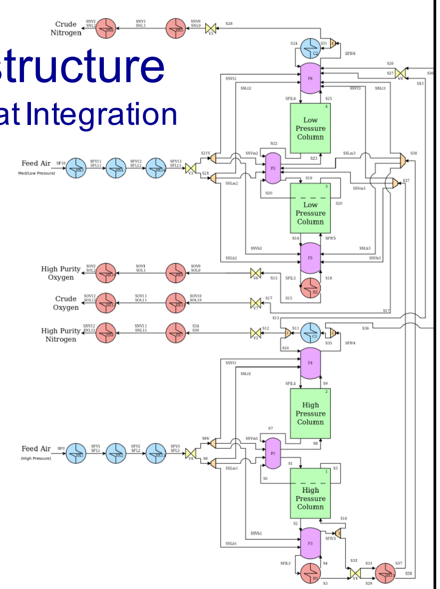
Phase conditions not known *a priori*

ASU NLP Superstructure

Phase Equilibrium and Heat Integration

$$\min \quad \begin{array}{c}\text{ASU Compression Energy} \\ (\text{kWh / kg } O_2 \text{ product})\end{array}$$

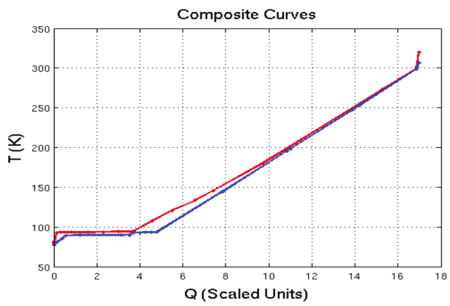
$$\text{s.t.} \quad \begin{array}{c}\text{Thermodynamics Module} \\ \text{Unit Operation Models} \\ \text{Cascade Model} \\ \text{Flowsheet Connectivity} \\ \text{Heat Integration} \\ O_2 \text{ product purity} \geq 95 \text{ mol\%}\end{array}$$

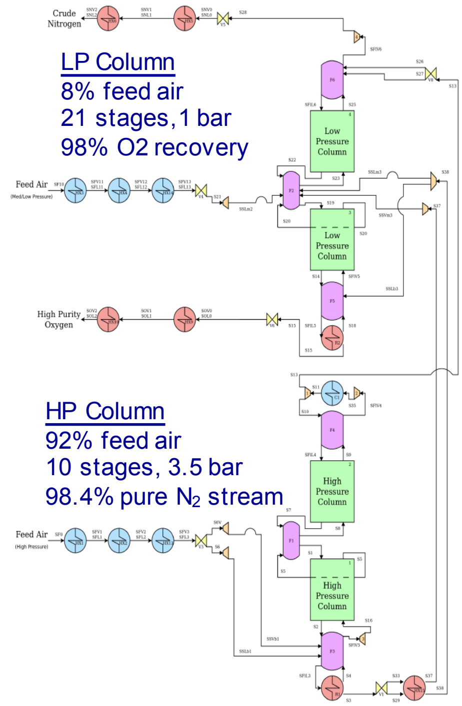


ASU Optimization

ΔT_{min} = 1.5 K, 95% O_2 purity

- Balanced Reboiler/Condenser
- No heating and cooling, only power
- Typical NLP: 15534 variables, 261 degrees of freedom
- MPCC bootstrapping "work process" ➔ 15 CPU min (CONOPT/ GAMS)
- 0.196 kWh/kg (86% comp efficiency)
- Compares well with industrial designs

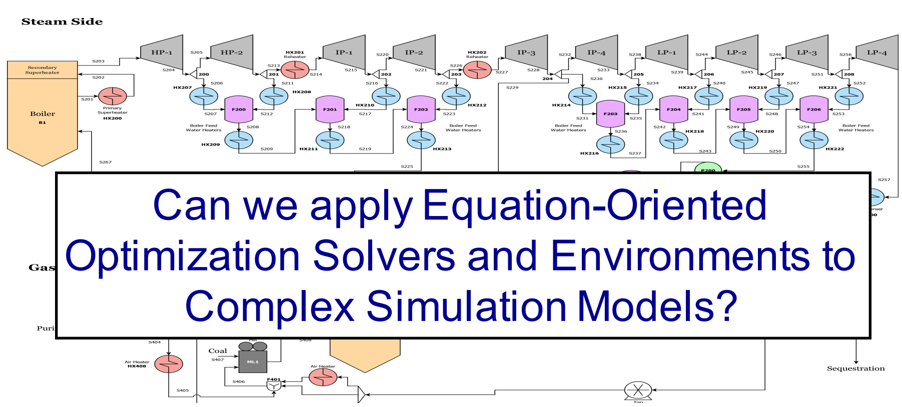




LP Column
8% feed air
21 stages, 1 bar
98% O2 recovery

HP Column
92% feed air
10 stages, 3.5 bar
98.4% pure N_2 stream

Heterogeneous Models for Process Optimization



Can we apply Equation-Oriented Optimization Solvers and Environments to Complex Simulation Models?

Incorporate Equation-Oriented Reduced (Surrogate) Models

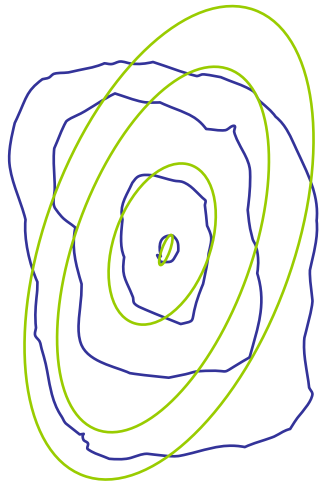
Original Detailed Models (ODMs)	Reduced Models (RMs)
Reactor PDEs	POD, α model …
Distillation	Shortcuts…
Physical Props.	Polynomials, Kriging

How Simple Should the RM be for Optimization?

Consistency
- ODM and Reduced model (RM) must match (be feasible)
- ODM and RM must recognize same optimum point

  => satisfy same KKT conditions (gradient-based)

Stability
- Sequence of objective functions remains bounded
- Provide sufficient improvement toward ODM optimum



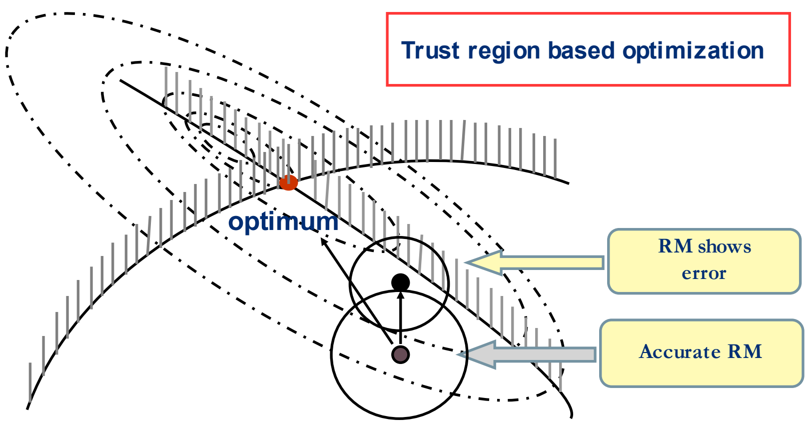
Roberts (1979), B. et al., (1985); Forbes and Marlin (1998); Engell (2007)

Reduced Model Optimization Strategy

RM depends on ODM information at current parameter values

ODM gradients - often not available

Trust region based optimization



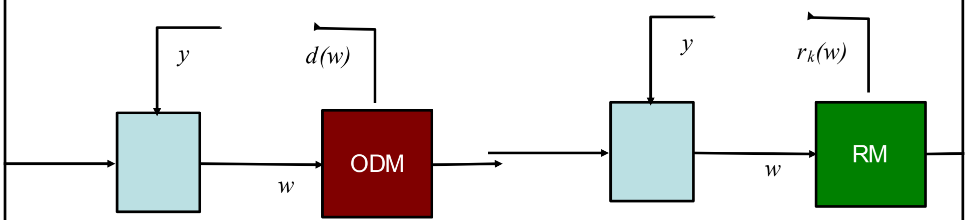
Extend RM-based trust region strategy for objective and constraint functions

Lifting to Separate Glass and Black Box Models

$$\begin{aligned} \min_x \quad & f(x) \\ \text{s.t.} \quad & h(x) = 0 \\ & g(x) \le 0 \\ & y = d(w) \end{aligned}$$

$$x = \begin{bmatrix} w \\ y \\ z \end{bmatrix}$$

$$\begin{aligned} \min \quad & f(x) \\ \text{s.t.} \quad & h(x) = 0 \\ & g(x) \le 0 \\ & y = r_k(w) \end{aligned}$$



Elements of TR Subproblem

(Fletcher, Gould, Leyffer, Toint, Wächter, 2002)

$$\begin{aligned} \min \quad & f(x) \\ \text{s.t.} \quad & h(x) = 0 \\ & g(x) \le 0 \\ & y = r_k(w) \\ & \|x - x_k\| \le \Delta_k. \end{aligned}$$

Trust Region Subproblem

$$\begin{aligned} \min_x \quad & \|y - r_k(w)\| \\ \text{s.t.} \quad & h(x) = 0 \\ & g(x) \le 0 \\ & \|x - x_k\| \le \kappa_\Delta \Delta_k \min[1, \kappa_\mu \Delta_k^\mu] \end{aligned}$$

Compatibility Check
(analogous to 'normal problem')

$$\theta(x) = \|y - d(w)\|.$$

$$\begin{aligned} \chi(x) \;=\; & \left|\min_v \;\; \nabla f(x)^T v\right| \\ \text{s.t.} \quad & \nabla h(x)^T v = 0 \\ & g(x) + \nabla g(x)^T v \le 0 \\ & v_y - \nabla r_k(w)^T v_w = 0 \\ & \|v\| \le 1 \end{aligned}$$

Criticality Measure

$$\rho_k = \frac{ared_k}{pred_k} = \frac{\theta(x_k) - \theta(x_k + s_k)}{\theta^r(x_k) - \theta^r(x_k + s_k)}$$

$$= \frac{\theta(x_k) - \theta(x_k + s_k)}{\|y_k - r_k(w_k)\|} = 1 - \frac{\theta(x_k + s_k)}{\theta^r(x_k)}$$

Trust Region Management

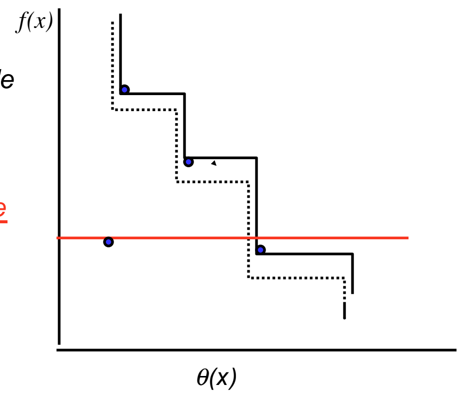
Trust Region Filter Method

(Fletcher, Gould, Leyffer, Toint, Wächter, 2002)

- *Store (f_j, θ_j) at allowed iterates → $\mathcal{F}$*
- *Allow TR step if trial point is acceptable to filter with θ_j margin* $\forall(\theta_j, f_j) \in \mathcal{F}$

$$\theta(x) \le (1-\gamma_\theta)\theta_j \quad \text{or} \quad f(x) \le f_j - \gamma_f \theta_j$$

- *If switching condition is satisfied (f-type step), only f(x) should be sufficiently reduced in TR.*
- *If insufficient progress on in TR, evoke restoration phase to reduce θ.*



- Exact derivatives from ODM (Agarwal, B., 2012)
- Inexact Jacobians from ODM (Walther, B., 2015)
- Derivative Free Models from ODM (Eason, B., 2015 -)

  (assuming no noisy functions)

Trust Region Strategy w/o ODM Gradients

(Conn, Scheinberg, Vicente, 2010)

κ - Fully Linear Property (FLP): As trust region Δ shrinks, both model functions and gradients must match:

$$\|\nabla r_k(w) - \nabla d(w)\| \le \kappa_g \Delta_k \qquad and \qquad \|r_k(w) - d(w)\| \le \kappa_f \Delta_k^2$$

If κ - FLP does not hold in Δ, refine RM
(apply more ODM evaluations until property holds)

- *Requires Criticality Phase as in DFO Trust Region Algorithm*
- ***Convergence of TR method requires Δ_k → 0***

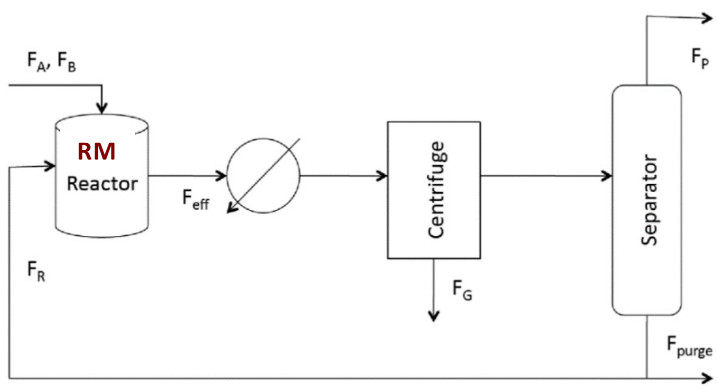
Trust Region Filter with Reduced Models

1. Initialize constants, starting guesses, trust region and filter.
2. Generate RM that is κ-fully linear on trust region.
3. Check for Compatibility. If infeasible, add (f_k, θ_k) to filter. Go to Step 9.
4. Check for convergence of TRSP. If satisfied, reduce Δ. Go to Step 2.
5. Solve TRSP to yield $x_k + s_k$. If infeasible, add (f_k, θ_k) to filter. Go to Step 9.
6. If $x_k + s_k$ is not acceptable to filter, reduce Δ. Go to step 7.
7. If s_k is f-type step, improve trust region Δ. Else, add (f_k, θ_k) to filter, update Δ.
8. Set k = k + 1, so to step 2.
9. Restoration: compute some x_{k+1} acceptable to filter (heuristics). Go to step 7.

Algorithm is globally convergent to KKT point with ODM. Proof modified from Fletcher et al. to deal with nonlinear TRSP (MFCQ needed) and Steps 2-4 (criticality phase from CSV)

Williams-Otto (Toy) Optimization Problem

$$Max\ ROI = 100(2207F_P + 50F_{purge} - 168F_A - 252F_B \\ -2.22F_{eff}^{sum} - 84F_G - 60V\rho)/(600V\rho)$$



RMs: Extent of reaction models as function of T, V, and component fraction

18

Williams-Otto: ODM→ RM

$$A + B \rightarrow C,$$
$$B + C \rightarrow P + E$$
$$P + C \rightarrow G.$$

Replaced by RMs

$$k_1 = a_1\, exp(-120/T),$$
$$k_2 = a_2\, exp(-150/T),$$
$$k_3 = a_3\, exp(-200/T),$$
$$r_1 = k_1 x_A x_B V \rho,$$
$$r_2 = k_2 x_B x_C V \rho,$$
$$r_3 = k_3 x_P x_C V \rho.$$

$$F_{eff}^A = F_A + F_R^A - r_1,$$
$$F_{eff}^B = F_B + F_R^B - (r_1 + r_2),$$
$$F_{eff}^C = F_R^C + 2r_1 - 2r_2 - r_3,$$
$$F_{eff}^E = F_R^E + 2r_2,$$
$$F_{eff}^P = 0.1F_R^E + r_2 - 0.5r_3,$$
$$F_{eff}^G = 1.5r_3,$$
$$F_{eff}^{sum} = \sum_j F_{eff}^j, \quad j \in \{A, B, C, E, P, G\}$$
$$F_{eff}^j = F_{eff}^{sum} x_j.$$

$$F_P = F_{eff}^P - 0.1F_{eff}^E,$$
$$F_{purge} = \eta(F_{eff}^A + F_{eff}^B + F_{eff}^C + 1.1F_{eff}^E)$$

19

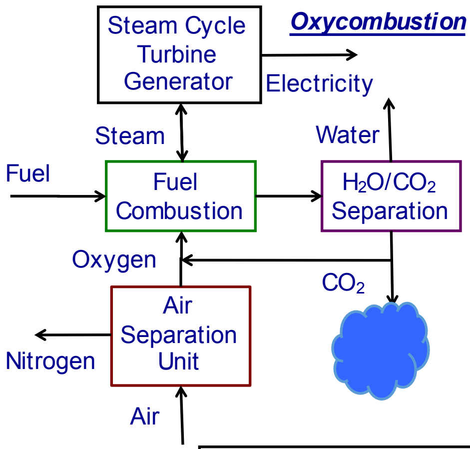
Williams-Otto Optimization Results

- Filter method with linear RM efficiently reduces simulation calls
- Kriging RM also converges but more expensive due to poorer approximations
- Linearization with finite differences in the filter method drastically reduces ODM calls over Full Process (SQP, exact gradients) by 57%
- Unconstrained DFO with Kriging RMs is 100x more expensive

	Linear RM 1.e-5	Linear RM Δ/2	Kriging on ODM (DACE)	Process Opt. (SQP)	Kriging on Process (Penalty)
Objective	-1.2111	-1.2111	-1.211	-1.2111	-1.2111
Iterations	13	92	123	15	53
RM Points	91	644	3141	-	573
ODM Calls	91	644	3141	210	11622

20

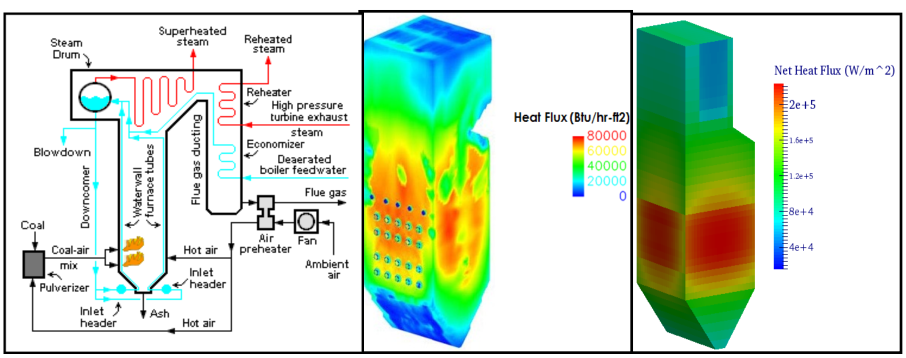
Oxycombustion Power Plants with CO_2 Capture

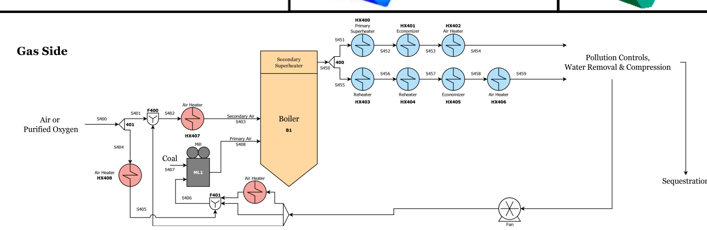


Schwarze Pumpe, 30MW Pilot (2008)
Feed: Lignite; Bituminous Coal
Brandenburg, Germany

Heterogeneous Process Optimization Models:
ASU – distillation (MPEC)
Combustion – PDAE/CFD Models
Steam Tables – Procedural Model
Gas separation – EO equipment, distillation

Air-fired Power Plant Optimization: Integrating Detailed Boiler Model





Air-fired Steam Cycle Optimization

(Dowling et al., 2015)

Max **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

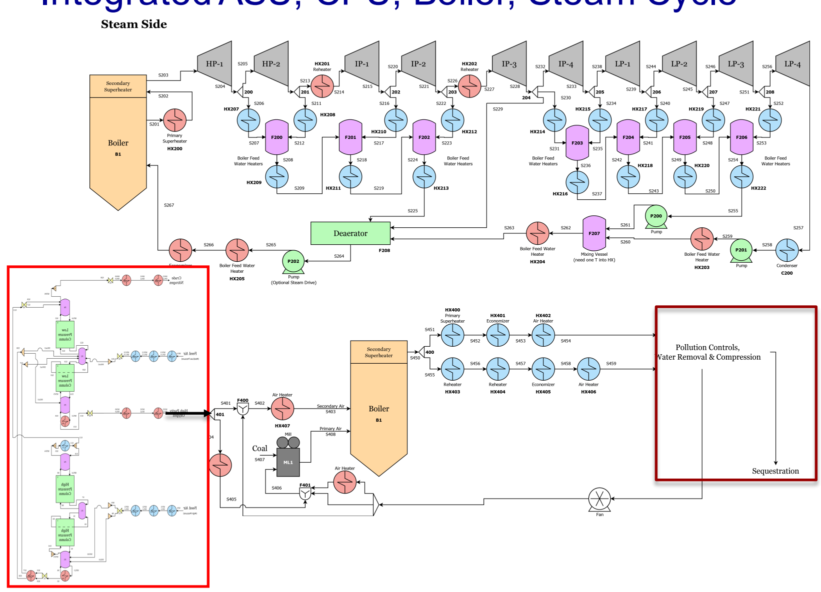
Using reduced models with trust region method ➔ rigorous optimum

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

23

Oxycombustion Optimization:

Integrated ASU, CPU, Boiler, Steam Cycle



Oxycombustion Optimization

(Dowling et al., 2015)

Max **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

Correlation models for ASU and CPU

Standard supercritical steam cycle, double reheat

Using reduced models with trust region method ➔ rigorous optimum

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

25

Oxycombustion Compared to Air-Fired Power Plants

	Air-fired	Oxy-fired
Solution time (hours)	9.8	8.6
Boiler simulations: (run on 4 cores)	759	598
Flue exit gas temperature (K)	1600	1600
Steam exit temperature (K)	835	835
Steam exit pressure (bar)	223	223
Fuel rate, HHV (MW)	1325.5	1325.5
ASU + CPU Power (MW)	N/A	114.3
Net Power (MWe)	515.5	437.4
Efficiency (HHV)	**38.9%**	**33.0%**

Only 5.9% penalty for oxy-fired configuration

26

Conclusions

- Need efficient optimization strategies for heterogeneous models (glass box vs. black box)
  - Use RMs to turn black box to glass box formulations
- Trust region algorithm embeds RMs into equation-based optimization problems
  - Guaranteed convergence to rigorous optimum
- Process Optimization Results
  - Validated on Toy (W-O) Problem
  - Demonstrated on Oxycombustion Power Plant Optimization
  - 38.9% efficient (optimal) air-fired configuration
  - Optimal CO_2 capture has less than 6% efficiency penalty over air-fired
  - Next steps: enhanced boiler model with geometric decisions and economics