Interior-Point Methods for Nonlinear Dynamic Optimization

Alexandre Caboussat

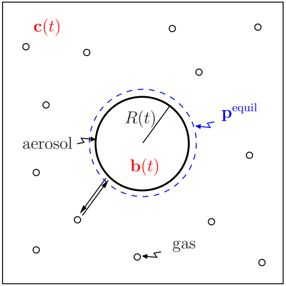
Department of Mathematics, University of Houston
Houston, TX 77204-3008, USA

caboussat@math.uh.edu
http://math.uh.edu/∼caboussat

INFORMS Annual Meeting 2007, Seattle, WA
November 6, 2007

Motivations

- *"The chemical and physical properties of aerosols are needed to estimate and predict direct and indirect climate forcing"*, (IPCC, 2001).
- Modeling and computation of the physical state and chemical composition of atmospheric aerosol particles.
- Constrained optimization problem models the thermodynamic equilibrium.
- Differential equations model the mass transfer between the particle and the gas.



Outline

- Introduction and Motivations
- Modeling of Dynamic Optimization
  - Dynamic Optimization
  - Event Location
- Application
  - Phase equilibrium
  - Mass Transfer
- Numerical Results
  - Trajectories and Warm-Starting
  - Tracking Techniques
- Current Work and Perspectives

Model Problem

- Coupling between an optimization problem and a differential equation.

$$\begin{aligned} \frac{d}{dt}\mathbf{b}(t) &= f(t,\mathbf{b}(t),\mathbf{z}(t)), \quad \mathbf{b}(0)=\mathbf{b}_0, \\ \mathbf{z}(t) = \arg\min_{\mathbf{z}^\star} & \quad g(\mathbf{z}^\star) \\ \text{s. t.} & \quad A\mathbf{z}^\star = \mathbf{b}(t), \quad \mathbf{z}^\star \geq 0. \end{aligned}$$

  i.e. $\mathbf{z}(t)$ is the **global minimum** of the optimization problem at time $t \in (0, T)$.
- Minimization under equality and inequality constraints.
- The inequality constraints imply that the variables $\mathbf{z}(t)$ are truncated and not smooth.

Time Discretization

- Consider $h > 0$, $t^n = nh$ and $\mathbf{b}^n \simeq \mathbf{b}(t^n)$, $\mathbf{z}^n \simeq \mathbf{z}(t^n)$.
- An implicit first order discretization of the differential equation reads, for all $n \geq 0$:

$$
\begin{aligned}
\frac{\mathbf{b}^{n+1} - \mathbf{b}^n}{h} &= f(t^{n+1}, \mathbf{b}^{n+1}, \mathbf{z}^{n+1}), \quad \mathbf{b}^0 = \mathbf{b}_0, \\
\mathbf{z}^{n+1} = \arg\min_{\mathbf{z}^\star} \quad & g(\mathbf{z}^\star), \\
\text{s. t.} \quad & A\mathbf{z}^\star = \mathbf{b}^{n+1}, \\
& \mathbf{z}^\star \geq 0.
\end{aligned}
$$

- Coupling of the optimization problem with one additional algebraic relation.
- The optimization problem is replaced with the relaxed first order (Karush-Kuhn-Tucker) conditions.

KKT conditions

$$
\begin{aligned}
&\frac{\mathbf{b}^{n+1}-\mathbf{b}^{n}}{h} = f(t^{n+1}, \mathbf{b}^{n+1}, \mathbf{z}^{n+1}),\\
&\nabla g(\mathbf{z}^{n+1}) + A^T \boldsymbol{\lambda}^{n+1} - \boldsymbol{\theta}^{n+1} = 0,\\
&A\mathbf{z}^{n+1} = \mathbf{b}^{n+1},\\
&\boldsymbol{\theta}^{n+1}\mathbf{z}^{n+1} = 0, \quad \boldsymbol{\theta}^{n+1} \geq 0, \quad \mathbf{z}^{n+1} \geq 0.
\end{aligned}
$$

- $\boldsymbol{\lambda}^{n+1}$ is the Lagrange multiplier related to the equality constraint.
- $\boldsymbol{\theta}^{n+1}$ is the Kuhn-Tucker multiplier related to the inequality constraint.
- Equivalence for convex optimization, but not necessarily for the non-convex case.

Interior-Point Method

- **At each time step**, the IPM provides a sequence $\mathbf{z}_\nu^{n+1} > \mathbf{0}$ that converges ultimately to $\mathbf{z}^{n+1}$ when $\nu \to 0$.

$$
\begin{aligned}
&\frac{\mathbf{b}^{n+1} - \mathbf{b}^n}{h} = f(t^{n+1}, \mathbf{b}^{n+1}, \mathbf{z}_\nu^{n+1}),\\
&\nabla g(\mathbf{z}_\nu^{n+1}) + A^T \boldsymbol{\lambda}^{n+1} - \boldsymbol{\theta}^{n+1} = 0,\\
&A\mathbf{z}_\nu^{n+1} = \mathbf{b}^{n+1},\\
&\boldsymbol{\theta}^{n+1}\mathbf{z}_\nu^{n+1} = \nu, \quad \boldsymbol{\theta}^{n+1} > 0, \quad \mathbf{z}_\nu^{n+1} > 0.
\end{aligned}
$$

- **Result:** The solution of the relaxed system of nonlinear equations tends to the solution of the constrained optimization problem when $\nu \to 0$.

Newton System

- At each time step and for a given value of ν, a Newton method is used to solve the system of nonlinear equations.

$$\begin{pmatrix} \frac{1}{h}\mathbf{I} - \nabla_{\mathbf{b}^{n+1}} f & -\nabla_{\mathbf{z}^{n+1}} f & -\nabla_{\boldsymbol{\lambda}^{n+1}} f \\ 0 & \nabla^2 g + \frac{\nu}{(\mathbf{z}^{n+1})^2} & A^T \\ -\mathbf{I} & A & 0 \end{pmatrix} \begin{pmatrix} \mathbf{p_b} \\ \mathbf{p_z} \\ \mathbf{p_\lambda} \end{pmatrix} = \begin{pmatrix} \mathbf{r_b} \\ \mathbf{r_z} \\ \mathbf{r_\lambda} \end{pmatrix}$$

- Newton method incorporated in the interior-point iterations (decreasing values of the parameter ν).
- Resolution of the linear system with direct decomposition techniques, with control of the inertia (Sequential quadratic programming or Schur complement techniques).

Cold-Start vs. Warm-Start

When the objective function is non-convex:

- **Cold-start techniques**.
  - Convergence to the global minimum.
  - Accurate detection of active/inactive constraints.
- **Warm-start techniques**
  - Quadratic convergence in a neighborhood of a KKT point.
  - Possible convergence to a local minimum.
  - Activation/deactivation of inequality constraints can be missed!

Bifurcation Problem

- Bifurcation to branches of local minima when inequality constraints are activated/deactivated.
- Local minima are **local attractors** (metastable states).
- KKT system does not contain information about the global minimum.
- Tracking techniques for the time of activation/deactivation of the constraints $\mathbf{z} \geq 0$.
  - Automatic detection of activation/deactivation to allow large time steps.
  - Computation of the exact time of activation/deactivation.
  - Computation of unknowns at this precise time to restart.

Event Location Techniques

- Model problem for an event in (t^n, t^{n+1}):

$$\mathbf{b}^{n+1} = \mathbf{b}^n + hf(t^{n+1}, \mathbf{b}^{n+1}, \mathbf{z}^{n+1}),$$
$$0 = g(t, \mathbf{b}^{n+1}, \mathbf{z}^{n+1}).$$

- One additional equation describes the *event*:

$$w(t^n + h^\star, \mathbf{b}^{n+1}(t^n + h^\star), \mathbf{z}^{n+1}(t^n + h^\star)) = 0.$$

- Determine the fraction of time step $h^\star \in (0, h)$ such that $t^\star = t^n + h^\star$ is the time of discontinuity.
- Compute the variables at the time of impact $t^\star$ with higher order methods, in order to **preserve accuracy of the algorithm**, and **compute exact time of activation or deactivation of constraints.**

Event Location Techniques

- Model problem for an event in (t^n, t^{n+1}):

$$\begin{aligned} \mathbf{b}^{n+1} &= \mathbf{b}^n + hf(t^{n+1}, \mathbf{b}^{n+1}, \mathbf{z}^{n+1}), \\ 0 &= g(t, \mathbf{b}^{n+1}, \mathbf{z}^{n+1}). \end{aligned}$$

- One additional equation describes the *event*:

$$w(t^n + h^\star, \mathbf{b}^{n+1}(t^n + h^\star), \mathbf{z}^{n+1}(t^n + h^\star)) = 0.$$

- Determine the fraction of time step $h^\star \in (0, h)$ such that $t^\star = t^n + h^\star$ is the time of discontinuity.
- Compute the variables at the time of impact $t^\star$ with higher order methods, in order to **preserve accuracy of the algorithm**, and **compute exact time of activation or deactivation of constraints.**

Event Location Techniques

- Model problem for an event in (t^n, t^{n+1}):

$$\begin{aligned} \mathbf{b}^{n+1} &= \mathbf{b}^n + hf(t^{n+1}, \mathbf{b}^{n+1}, \mathbf{z}^{n+1}), \\ 0 &= g(t, \mathbf{b}^{n+1}, \mathbf{z}^{n+1}). \end{aligned}$$

- One additional equation describes the *event*:

$$w(t^n + h^\star, \mathbf{b}^{n+1}(t^n + h^\star), \mathbf{z}^{n+1}(t^n + h^\star)) = 0.$$

- Determine the fraction of time step $h^\star \in (0, h)$ such that $t^\star = t^n + h^\star$ is the time of discontinuity.
- Compute the variables at the time of impact $t^\star$ with higher order methods, in order to **preserve accuracy of the algorithm**, and **compute exact time of activation or deactivation of constraints.**

Event Location Techniques

- Model problem for an event in (t^n, t^{n+1}):

$$\mathbf{b}^{n+1} = \mathbf{b}^n + hf(t^{n+1}, \mathbf{b}^{n+1}, \mathbf{z}^{n+1}),$$
$$0 = g(t, \mathbf{b}^{n+1}, \mathbf{z}^{n+1}).$$

- One additional equation describes the *event*:

$$w(t^n + h^\star, \mathbf{b}^{n+1}(t^n + h^\star), \mathbf{z}^{n+1}(t^n + h^\star)) = 0.$$

- Determine the fraction of time step $h^\star \in (0, h)$ such that $t^\star = t^n + h^\star$ is the time of discontinuity.
- Compute the variables at the time of impact $t^\star$ with higher order methods, in order to **preserve accuracy of the algorithm**, and **compute exact time of activation or deactivation of constraints.**

Tracking of Activations

- Activation/deactivation of an inequality constraint:
  - Activation: $\mathbf{z}_i(t) > 0 \rightarrow \mathbf{z}_i(t) = 0$.
  - Deactivation: $\mathbf{z}_i(t) = 0 \rightarrow \mathbf{z}_i(t) > 0$.
- In the activation case, the event detection function w is implicitly given by $\mathbf{z}_i = 0$ and is the result of the optimization problem.
- Taylor expansion

$$0 = \mathbf{z}_i(t^n + h^\star) = \mathbf{z}_i(t^n) + h^\star \frac{d\mathbf{z}_i}{dt}(t^n) + \mathcal{O}((h^\star)^2).$$

- Truncation of the Taylor expansion:

$$h^\star \simeq -\frac{\mathbf{z}_i(t^n)}{\frac{d\mathbf{z}_i}{dt}(t^n)}.$$

Sensitivity Analysis

- Approximation of $\dfrac{d\mathbf{z}_i}{dt}(t^n)$ by differentiation of the KKT system (for $\nu = 0$):

$$\begin{bmatrix} \nabla^2 g & A^T \\ A & 0 \end{bmatrix} \begin{bmatrix} \dfrac{d\mathbf{z}}{dt}(t^n) \\ \dfrac{d\boldsymbol{\lambda}}{dt}(t^n) \end{bmatrix} = \begin{bmatrix} 0 \\ \dfrac{d\mathbf{b}}{dt}(t^n) \end{bmatrix}$$

- The right-hand side is approximated by the numerical fluxes

$$\frac{d\mathbf{b}}{dt}(t^n) \simeq f(t^n, \mathbf{b}^n, \mathbf{z}^n).$$

- *Remark:* Second order sensitivity analysis necessary to approximate the second term in the Taylor expansion.

Multistep Techniques

- $h^\star$ given, multisteps methods are used for the approximation of the differential variables $\mathbf{b}$ at time $t^n + h^\star$.
- Predictor (Adams-Bashforth 2-steps):

$$\tilde{\mathbf{b}}^{n+1}(h^\star) = \mathbf{b}^n + h^\star \left[\left(1 + \frac{h^\star}{2h}\right) f(t^n, \mathbf{b}^n, \mathbf{z}^n) - \frac{h^\star}{2h} f(t^{n-1}, \mathbf{b}^{n-1}, \mathbf{z}^{n-1}) \right].$$

- Computation of $\tilde{\mathbf{z}}^{n+1}(h^\star)$ corresponding to $\tilde{\mathbf{b}}^{n+1}(h^\star)$.
- Corrector (Adams-Moulton 2-steps):

$$\mathbf{b}^{n+1} = \mathbf{b}^n + \frac{1}{2} h^\star \left(f(t^n + h^\star, \tilde{\mathbf{b}}^{n+1}(h^\star), \tilde{\mathbf{z}}^{n+1}(h^\star)) + f(t^n, \mathbf{b}^n, \mathbf{z}^n) \right).$$

Multistep Techniques

- $h^\star$ given, multisteps methods are used for the approximation of the differential variables $\mathbf{b}$ at time $t^n + h^\star$.
- Predictor (Adams-Bashforth 2-steps):

$$\tilde{\mathbf{b}}^{n+1}(h^\star) = \mathbf{b}^n + h^\star \left[\left(1 + \frac{h^\star}{2h}\right) f(t^n, \mathbf{b}^n, \mathbf{z}^n) - \frac{h^\star}{2h} f(t^{n-1}, \mathbf{b}^{n-1}, \mathbf{z}^{n-1}) \right].$$

- Computation of $\tilde{\mathbf{z}}^{n+1}(h^\star)$ corresponding to $\tilde{\mathbf{b}}^{n+1}(h^\star)$.
- Corrector (Adams-Moulton 2-steps):

$$\mathbf{b}^{n+1} = \mathbf{b}^n + \frac{1}{2} h^\star \left(f(t^n + h^\star, \tilde{\mathbf{b}}^{n+1}(h^\star), \tilde{\mathbf{z}}^{n+1}(h^\star)) + f(t^n, \mathbf{b}^n, \mathbf{z}^n) \right).$$

Multistep Techniques

- $h^\star$ given, multisteps methods are used for the approximation of the differential variables $\mathbf{b}$ at time $t^n + h^\star$.
- Predictor (Adams-Bashforth 2-steps):

$$\tilde{\mathbf{b}}^{n+1}(h^\star) = \mathbf{b}^n + h^\star \left[\left(1 + \frac{h^\star}{2h}\right) f(t^n, \mathbf{b}^n, \mathbf{z}^n) - \frac{h^\star}{2h} f(t^{n-1}, \mathbf{b}^{n-1}, \mathbf{z}^{n-1}) \right].$$

- Computation of $\tilde{\mathbf{z}}^{n+1}(h^\star)$ corresponding to $\tilde{\mathbf{b}}^{n+1}(h^\star)$.
- Corrector (Adams-Moulton 2-steps):

$$\mathbf{b}^{n+1} = \mathbf{b}^n + \frac{1}{2} h^\star \left(f(t^n + h^\star, \tilde{\mathbf{b}}^{n+1}(h^\star), \tilde{\mathbf{z}}^{n+1}(h^\star)) + f(t^n, \mathbf{b}^n, \mathbf{z}^n) \right).$$

Multistep Techniques

- $h^\star$ given, multisteps methods are used for the approximation of the differential variables $\mathbf{b}$ at time $t^n + h^\star$.
- Predictor (Adams-Bashforth 2-steps):

$$\tilde{\mathbf{b}}^{n+1}(h^\star) = \mathbf{b}^n + h^\star \left[\left(1 + \frac{h^\star}{2h}\right) f(t^n, \mathbf{b}^n, \mathbf{z}^n) - \frac{h^\star}{2h} f(t^{n-1}, \mathbf{b}^{n-1}, \mathbf{z}^{n-1}) \right].$$

- Computation of $\tilde{\mathbf{z}}^{n+1}(h^\star)$ corresponding to $\tilde{\mathbf{b}}^{n+1}(h^\star)$.
- Corrector (Adams-Moulton 2-steps):

$$\mathbf{b}^{n+1} = \mathbf{b}^n + \frac{1}{2} h^\star \left(f(t^n + h^\star, \tilde{\mathbf{b}}^{n+1}(h^\star), \tilde{\mathbf{z}}^{n+1}(h^\star)) + f(t^n, \mathbf{b}^n, \mathbf{z}^n) \right).$$

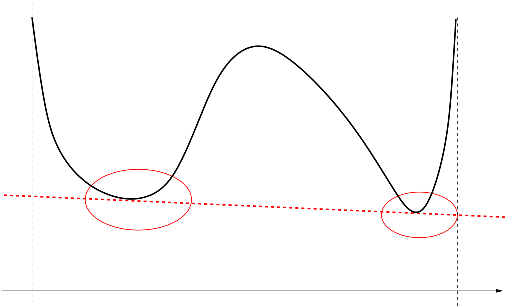
Application - Phase Equilibrium Problem

- Global optimization problem for the modeling of phase separation into different liquid phases.

$$
\begin{array}{ll}
\min\limits_{y_\alpha \mathbf{x}_\alpha} & \sum\limits_{\alpha=1}^{N+1} y_\alpha g(\mathbf{x}_\alpha) \\
\text{s. t.} & \sum\limits_{\alpha=1}^{N+1} y_\alpha \mathbf{x}_\alpha = \mathbf{b}, \\
& y_\alpha \geq 0, \quad \alpha = 1, \dots, N+1. \\
& \mathbf{e}^T \mathbf{x}_\alpha = 1, \quad \mathbf{x}_\alpha > 0, \quad \alpha = 1, \dots, N+1.
\end{array}
$$

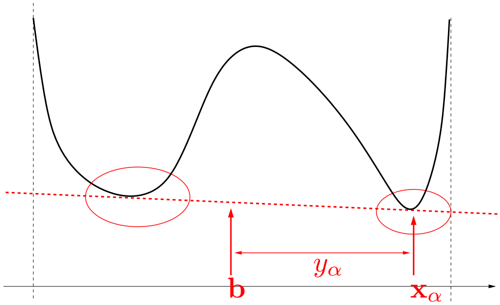
- y_α is the total number of moles in phase α and $\mathbf{x}_\alpha$ is the (normalized) mole-fraction.

Geometric Interpretation and Supporting Tangent Plane



- Global optimization corresponds to the determination of the convex envelope of the function g, or to the determination of the supporting tangent plane.
- **Result:** A KKT point is a global minimum if and only if $\nabla^2 g(\mathbf{x}_\alpha) > 0$ when $y_\alpha > 0$.

Geometric Interpretation and Supporting Tangent Plane



- Global optimization corresponds to the determination of the convex envelope of the function g, or to the determination of the supporting tangent plane.
- **Result:** A KKT point is a global minimum if and only if $\nabla^2 g(\mathbf{x}_\alpha) > 0$ when $y_\alpha > 0$.

Interior-Point Method

- Log/Barrier Penalty parameter for the treatment of the inequality constraints $y_\alpha \geq 0$

$$\begin{aligned}
\min_{y_\alpha \mathbf{x}_\alpha} \quad & \sum_{\alpha=1}^{N+1} y_\alpha g(\mathbf{x}_\alpha) \\
\text{s. t.} \quad & \sum_{\alpha=1}^{N+1} y_\alpha \mathbf{x}_\alpha = \mathbf{b}, \\
& \mathbf{e}^T \mathbf{x}_\alpha = 1, \quad \mathbf{x}_\alpha > 0, \quad \alpha = 1, \ldots, N+1. \\
& y_\alpha \geq 0, \quad \alpha = 1, \ldots, N+1.
\end{aligned}$$

- **Result:** For a "good" initial guess of the primal and dual variables and the relaxation parameter, the interior-point method converges to the global optimum.

Interior-Point Method

- Log/Barrier Penalty parameter for the treatment of the inequality constraints $y_\alpha \geq 0$

$$\min_{y_\alpha \mathbf{x}_\alpha} \quad \sum_{\alpha=1}^{N+1} y_\alpha g(\mathbf{x}_\alpha) - \nu \sum_{\alpha=1}^{N+1} \ln(y_\alpha)$$

$$\text{s. t.} \quad \sum_{\alpha=1}^{N+1} y_\alpha \mathbf{x}_\alpha = \mathbf{b},$$

$$\mathbf{e}^T \mathbf{x}_\alpha = 1, \quad \mathbf{x}_\alpha > 0, \quad \alpha = 1, \ldots, N+1.$$

- **Result:** For a "good" initial guess of the primal and dual variables and the relaxation parameter, the interior-point method converges to the global optimum.

Dynamic Optimization

- Coupling of the evolution of the gas concentrations with the optimization problem for the modeling of mass transfer.

$$
\begin{aligned}
\min_{y_\alpha(t),\mathbf{x}_\alpha(t)} \quad & \sum_{\alpha=1}^{N+1} y_\alpha(t) g(\mathbf{x}_\alpha(t)) - \nu \sum_{\alpha=1}^{N+1} \ln(y_\alpha(t)) \\
\text{s. t.} \quad & \sum_{\alpha=1}^{N+1} y_\alpha(t)\mathbf{x}_\alpha(t) = \mathbf{b}(t), \\
& \mathbf{e}^T \mathbf{x}_\alpha(t) - 1 = 0, \quad \mathbf{x}_\alpha(t) > 0, \quad \alpha = 1, \ldots, N+1.
\end{aligned}
$$

Dynamic Optimization

- Coupling of the evolution of the gas concentrations with the optimization problem for the modeling of mass transfer.

$$\frac{d}{dt}\mathbf{b}(t) = \mathbf{h}\left(\mathbf{b}^{\mathrm{tot}} - \mathbf{b}(t) - \mathbf{C}\exp\left(\nabla g(\mathbf{x}_\alpha(t))\right)\right)$$

$$\begin{aligned}
\min_{y_\alpha(t),\mathbf{x}_\alpha(t)} \quad & \sum_{\alpha=1}^{N+1} y_\alpha(t) g(\mathbf{x}_\alpha(t)) - \nu \sum_{\alpha=1}^{N+1} \ln(y_\alpha(t)) \\
\text{s. t.} \quad & \sum_{\alpha=1}^{N+1} y_\alpha(t)\mathbf{x}_\alpha(t) = \mathbf{b}(t), \\
& \mathbf{e}^T\mathbf{x}_\alpha(t) - 1 = 0, \quad \mathbf{x}_\alpha(t) > 0, \quad \alpha = 1, \dots, N+1.
\end{aligned}$$

Time Discretization and KKT conditions

- Discretization in time to obtain an extended optimization problem at each time step.

$$\frac{\mathbf{b}^{n+1} - \mathbf{b}^{n}}{h} = \mathbf{h}\left(\mathbf{b}^{tot} - \mathbf{b}^{n+1} - \mathbf{C}\exp\left(\nabla g(\mathbf{x}_{\alpha}^{n+1})\right)\right)$$

$$\begin{aligned} \min_{y_\alpha^{n+1},\mathbf{x}_\alpha^{n+1}} \quad & \sum_{\alpha=1}^{N+1} y_\alpha^{n+1} g(\mathbf{x}_\alpha^{n+1}) - \nu \sum_{\alpha=1}^{N+1} \ln(y_\alpha^{n+1}) \\ \text{s. t.} \quad & \sum_{\alpha=1}^{N+1} y_\alpha^{n+1}\mathbf{x}_\alpha^{n+1} = \mathbf{b}^{n+1}, \\ & \mathbf{e}^T\mathbf{x}_\alpha^{n+1} - 1 = 0, \quad \mathbf{x}_\alpha^{n+1} > 0, \quad \alpha = 1, \ldots, N+1. \end{aligned}$$

Time Discretization and KKT conditions

- Discretization in time to obtain an extended optimization problem at each time step.

$$\frac{\mathbf{b}^{n+1} - \mathbf{b}^{n}}{h} = \mathbf{h}\left(\mathbf{b}^{tot} - \mathbf{b}^{n+1} - \mathbf{C}\exp\left(\nabla g(\mathbf{x}_\alpha^{n+1})\right)\right)$$

$$y_\alpha^{n+1}\left(\nabla g(\mathbf{x}_\alpha^{n+1}) + \boldsymbol{\lambda}^{n+1}\right) + \zeta_\alpha^{n+1}\mathbf{e} = 0, \quad \alpha = 1, \ldots, N+1,$$

$$g(\mathbf{x}_\alpha^{n+1}) + (\boldsymbol{\lambda}^{n+1})^T \mathbf{x}_\alpha^{n+1} - \frac{\nu}{y_\alpha^{n+1}} = 0, \quad \alpha = 1, \ldots, N+1,$$

$$\sum_{\alpha=1}^{N+1} y_\alpha^{n+1}\mathbf{x}_\alpha^{n+1} - \mathbf{b}^{n+1} = 0, \quad \mathbf{e}^T\mathbf{x}_\alpha^{n+1} - 1 = 0.$$

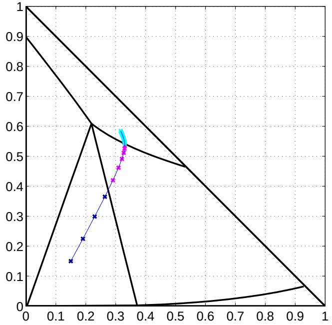
Newton System

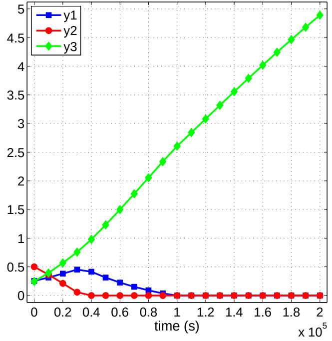
$$\begin{bmatrix} \mathbf{H_b} & 0 & 0 & \mathbf{B} & 0 \\ 0 & y_\alpha \nabla^2 g(\mathbf{x}_\alpha) & \nabla g(\mathbf{x}_\alpha) + \boldsymbol{\lambda} & y_\alpha & \mathbf{e} \\ 0 & (\nabla g(\mathbf{x}_\alpha) + \boldsymbol{\lambda})^T & \dfrac{\nu}{(y_\alpha)^2} & (\mathbf{x}_\alpha)^T & 0 \\ -\mathbf{I} & (y_\alpha)^T & \mathbf{x}_\alpha & 0 & 0 \\ 0 & \mathbf{e}^T & 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} \mathbf{p_b} \\ \mathbf{p_x} \\ \mathbf{p}_y \\ \mathbf{p}_{\boldsymbol{\lambda}} \\ \mathbf{p}_{\zeta_\alpha} \end{bmatrix} = \begin{bmatrix} \mathbf{r_b} \\ \mathbf{r_x} \\ \mathbf{r}_y \\ \mathbf{r}_{\boldsymbol{\lambda}} \\ \mathbf{r}_{\zeta_\alpha} \end{bmatrix}$$

- Karush-Kuhn-Tucker Newton system is block-structured and ill-conditioned.
- $\mathbf{H_b}$ is positive definite.
- Design of numerical algebra techniques, via sequential quadratic programming or Schur complement techniques.
- System can be symmetrized with a change of variables.

Numerical Results

- Trajectory of the (normalized) feed vector $\mathbf{b}(t)$. Convergence to a stationary solution.

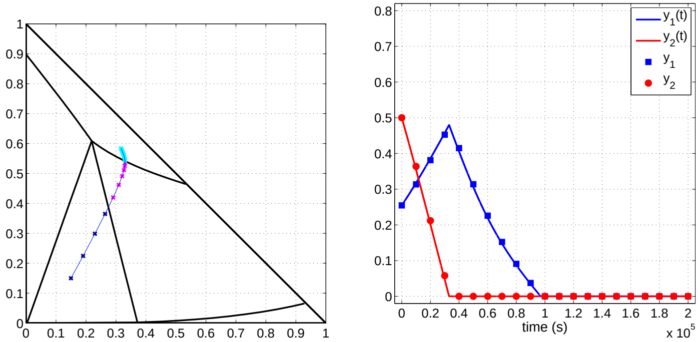




- Detection of the phase separations with cold starts.

Numerical Results

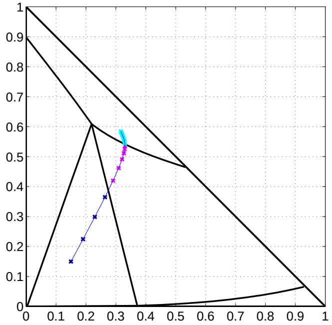
- Trajectory of the (normalized) feed vector $\mathbf{b}(t)$. Convergence to a stationary solution.

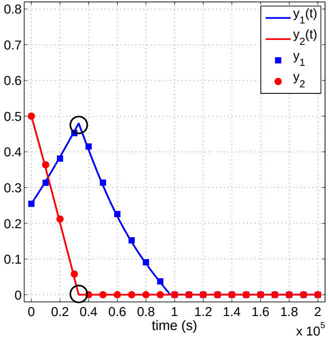


- Detection of the phase separations with cold starts.

Numerical Results

- Trajectory of the (normalized) feed vector $\mathbf{b}(t)$. Convergence to a stationary solution.

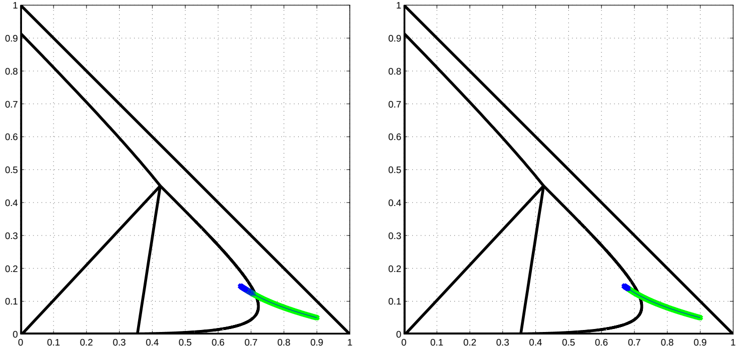




- Detection of the phase separations with cold starts.
- What happens with warm starts ??

Warm-Start Effect

- Warm-start approach or faster convergence:



Cold-starts Warm-starts

- Various warm-starts techniques miss the time of activation/deactivation!

Tracking Algorithm

- Warm-start techniques with backtracking and detection of the bifurcation points.
  - Perform the algorithm with warm-start techniques.
  - Detect the time interval $[t^n, t^{n+1}]$ where the activation/deactivation happens.
  - Use tracking techniques to determine $h^\star$ and the point of impact.
  - Restart with warm-start techniques from the discontinuity point.

Tracking Algorithm

- Warm-start techniques with backtracking and detection of the bifurcation points.
  - Perform the algorithm with warm-start techniques.
  - Detect the time interval $[t^n, t^{n+1}]$ where the activation/deactivation happens.
  - Use tracking techniques to determine $h^\star$ and the point of impact.
  - Restart with warm-start techniques from the discontinuity point.

Tracking Algorithm

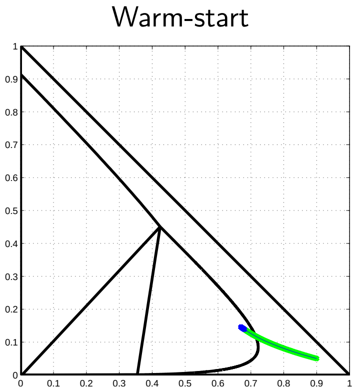
- Warm-start techniques with backtracking and detection of the bifurcation points.
  - Perform the algorithm with warm-start techniques.
  - Detect the time interval $[t^n, t^{n+1}]$ where the activation/deactivation happens.
  - Use tracking techniques to determine $h^\star$ and the point of impact.
  - Restart with warm-start techniques from the discontinuity point.

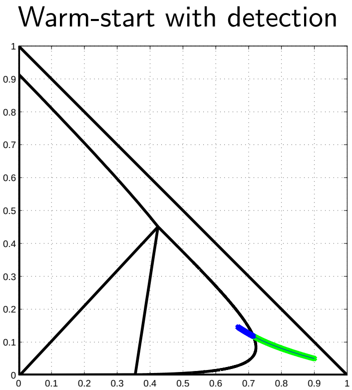
Tracking Algorithm

- Warm-start techniques with backtracking and detection of the bifurcation points.
  - Perform the algorithm with warm-start techniques.
  - Detect the time interval $[t^n, t^{n+1}]$ where the activation/deactivation happens.
  - Use tracking techniques to determine $h^\star$ and the point of impact.
  - Restart with warm-start techniques from the discontinuity point.

Comparison of Tracking Results

- Evolution of $\mathbf{b}(t)$:

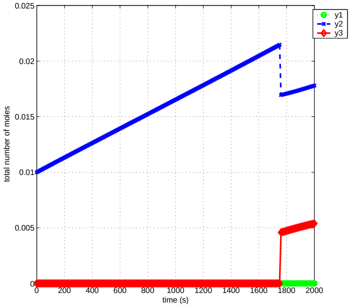




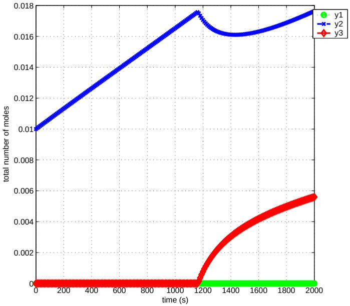
Comparison of Tracking Results

- Evolution of $y_\alpha(t)$:

Warm-start



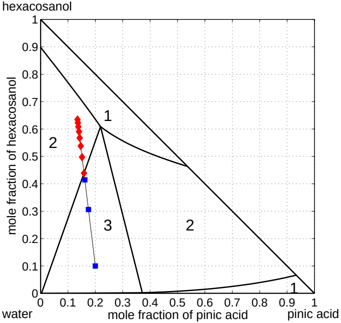
Warm-start with detection



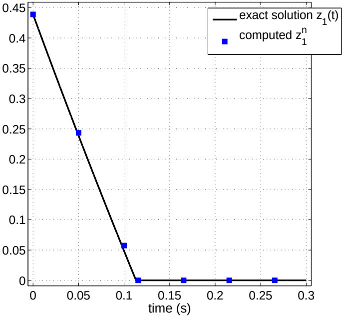
Convergence Results - Euler implicit

- When all constraints are active $y_\alpha > 0$, $\dfrac{d\mathbf{b}(t)}{dt} = f(\mathbf{b}(t))$ and the exact solution is known.

$\mathbf{b}^n$



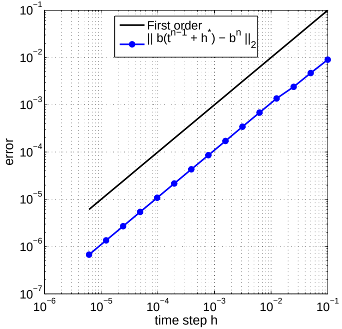
y_1^n

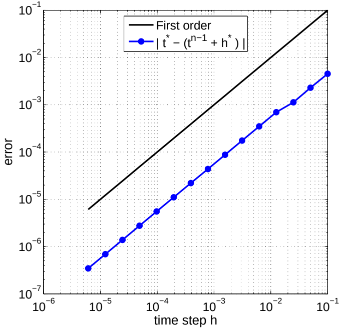


Convergence Results - Euler implicit

- When all constraints are active $y_\alpha > 0$, $\dfrac{d\mathbf{b}(t)}{dt} = f(\mathbf{b}(t))$ and the exact solution is known.

$$\left|\left|\mathbf{b}(t^n + h^\star) - \mathbf{b}^{n+1}\right|\right|_\infty \qquad\qquad \left|t^\star - (t^{n+1} + h^\star)\right|$$

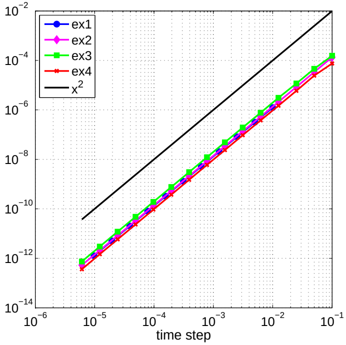


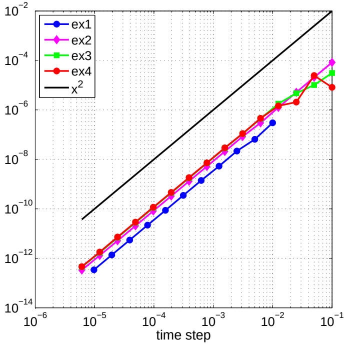


Convergence Results - Crank-Nicolson

- When all constraints are active $y_\alpha > 0$, $\dfrac{d\mathbf{b}(t)}{dt} = f(\mathbf{b}(t))$ and the exact solution is known.

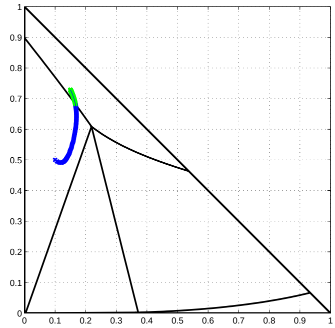
$$\left|\left|\mathbf{b}(t^n + h^\star) - \mathbf{b}^{n+1}\right|\right|_\infty \qquad \left|t^\star - (t^{n+1} + h^\star)\right|$$



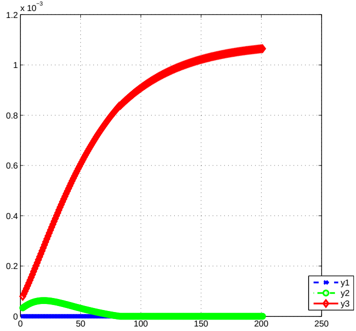


Constraint Activation - General Case

- Time of impact $t^\star$ when $y_\alpha = 0$. Transition from 2 phases to 1 phase.



Trajectory



Values of y_α

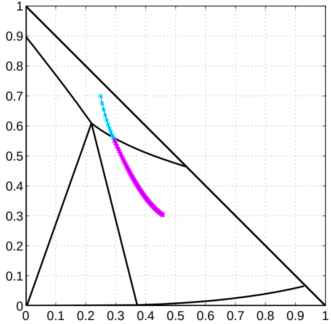
Current Work

Constraint Deactivation - Algorithm

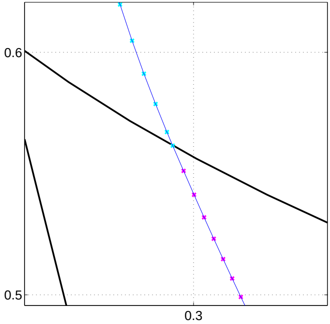
- The transition $y_\alpha = 0 \rightarrow y_\alpha > 0$ cannot be treated with a Taylor expansion of the function $y_\alpha(t)$ before the discontinuity.
- No dual variables available to replace y_α (numerical instability of dual variables).
- Geometric reconstruction of the interface with least-squares, interpolation, etc.
- Computation of intersection point of the trajectory with given reconstructed interface.

Constraint Deactivation - Preliminary Results

- Transition from 1 to 2 phases.



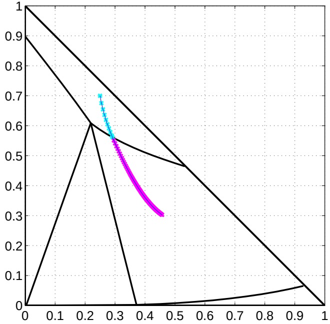
Trajectory



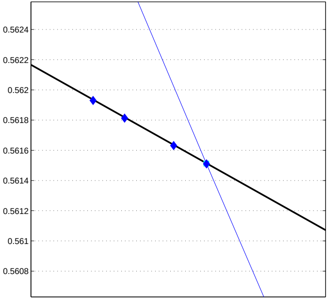
Zoom

Constraint Deactivation - Preliminary Results

- Transition from 1 to 2 phases.



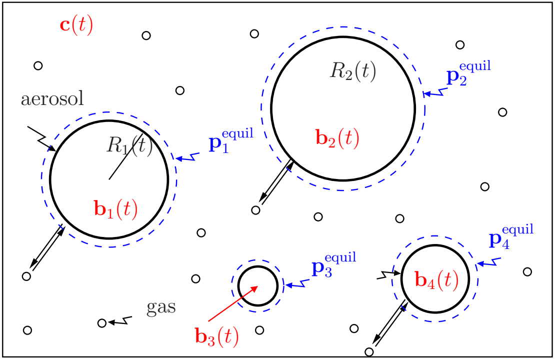
Trajectory



4 points to build the interface

Current Work - Sequence of Optimization Problems

- Population of aerosol particles.



- Differences of sizes/reaction speeds/modeling of internal energy increase the stiffness of the problem.

Modeling with Sequences of Optimization Problems

- Population of M aerosol particles ($i = 1, \ldots, M$).

$$\frac{d}{dt}\mathbf{c}(t) = -\sum_{i=1}^{M} h(r_i)\left(\mathbf{c}(t) - \eta(r_i)\exp\left(\nabla g_i(\mathbf{x}_\alpha^i)\right)\right), \quad \mathbf{c}(0) = \mathbf{c}_0$$

$$\frac{d}{dt}\mathbf{b}_i(t) = h(r_i)\left(\mathbf{c}(t) - \eta(r_i)\exp\left(\nabla g_i(\mathbf{x}_\alpha^i)\right)\right), \quad \mathbf{b}_i(0) = \mathbf{b}_{0,i}$$

$$\min_{y_\alpha^i, \mathbf{x}_\alpha^i} \quad \sum_{\alpha=1}^{P_i} y_\alpha^i g_i(\mathbf{x}_\alpha^i)$$

$$\text{s. t.} \quad \sum_{\alpha=1}^{P} y_\alpha^i \mathbf{x}_\alpha^i = \mathbf{b}_i(t),$$

$$y_\alpha^i \geq 0, \quad \mathbf{e}^T \mathbf{x}_\alpha^i = 1, \quad \mathbf{x}_\alpha^i > \mathbf{0}, \quad \alpha = 1, \ldots, P_i.$$

Time discretization and KKT system

- Time discretization, first order optimality conditions and Newton method lead to large block-structured linear systems.

$$
\left[\begin{array}{ccc|ccc|cc|c}
H_1 & 0 & B_1 & 0 & 0 & 0 & \dots & \dots & C_1 \\
0 & O_1 & A_1 & 0 & 0 & 0 & & & 0 \\
E_1 & A_1^T & 0 & 0 & 0 & 0 & & & 0 \\
\hline
0 & 0 & 0 & H_2 & 0 & B_2 & & & C_2 \\
0 & 0 & 0 & 0 & O_2 & A_2 & & & 0 \\
0 & 0 & 0 & E_2 & A_2^T & 0 & & & 0 \\
\hline
\vdots & & & & & & \ddots & & \vdots \\
\vdots & & & & & & & \ddots & \vdots \\
\hline
0 & 0 & D_1 & 0 & 0 & D_2 & \dots & \dots & H_0
\end{array}\right]
\left[\begin{array}{c}
p_{c_1} \\ p_{x_1} \\ p_{\lambda_1} \\ \hline p_{c_2} \\ p_{x_2} \\ p_{\lambda_2} \\ \hline \vdots \\ \vdots \\ \hline p_{c_0}
\end{array}\right]
=
\left[\begin{array}{c}
r_{c_1} \\ r_{x_1} \\ r_{\lambda_1} \\ \hline r_{c_2} \\ r_{x_2} \\ r_{\lambda_2} \\ \hline \vdots \\ \vdots \\ \hline r_{c_0}
\end{array}\right]
$$

- Schur complement techniques.

Schur Complement Techniques

- Construct the Schur complement system

$$\mathcal{S} p_{c_0} = \mathcal{R}$$

$$\mathcal{S} = H_0 - \sum_{i=1}^{M} \begin{pmatrix} 0 & 0 & D_i \end{pmatrix} \begin{pmatrix} H_i & 0 & B_i \\ 0 & O_i & A_i \\ E_i & A_i^T & 0 \end{pmatrix}^{-1} \begin{pmatrix} C_i \\ 0 \\ 0 \end{pmatrix}$$

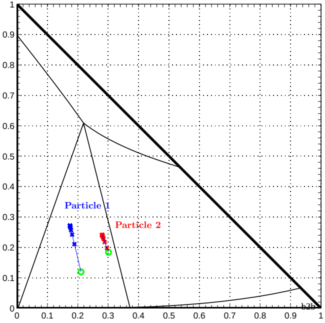
$$\mathcal{R} = r_{c_0} - \sum_{i=1}^{M} \begin{pmatrix} 0 & 0 & D_i \end{pmatrix} \begin{pmatrix} H_i & 0 & B_i \\ 0 & O_i & A_i \\ E_i & A_i^T & 0 \end{pmatrix}^{-1} \begin{pmatrix} r_{c_i} \\ r_{x_i} \\ r_{\lambda_i} \end{pmatrix}$$

- Each system corresponds to one particle.
- Size of the Schur complement $\mathcal{S}$ is reasonable.
- Followed a by sequence of M individual blocks.

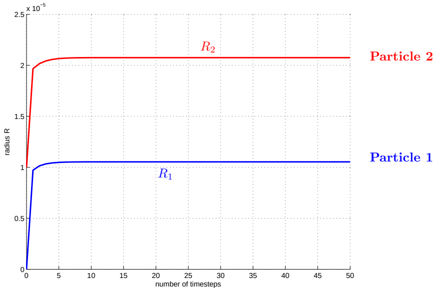
Preliminary Results

- Two particles without tracking of discontinuities.

Trajectories $\mathbf{b}_i$



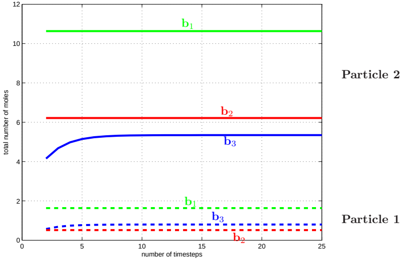
Radii R_i



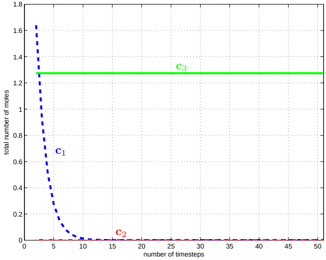
Preliminary Results

- Two particles without tracking of discontinuities.

Concentrations **b** in the particles



Concentrations **c** in the gas phase



Conclusions

- Coupling optimization problem with differential equations.
- Warm-starting heuristics.
- Numerical techniques for the tracking of activation/deactivation in the trajectories.

- Improvement of warm-starting heuristics.
- Improvement of interior-point parameter ν.
- Extension to sequences of optimization problems.

Acknowledgements:

C. Landry (EPFL, Switzerland)

A. Leonard (University of Houston)