6. Numerical methods for mathematical programs with complementarity constraints

Armin Nurkanović

Winter School on Numerical Methods for Optimal Control of Nonsmooth Systems
École des Mines de Paris
February 3-5, 2025, Paris, France

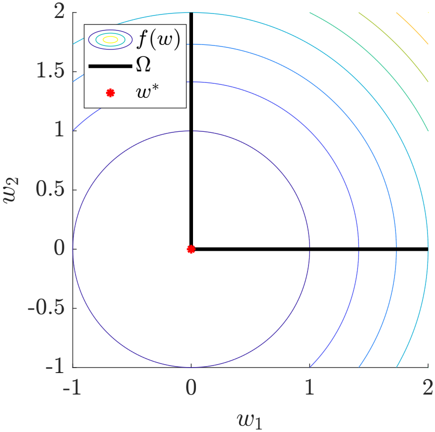
universität freiburg

Outline of the lecture

Numerical methods for MPCCs

MPEC

$$\min_{w\in\mathbb{R}^n} \quad f(w) \tag{1a}$$
$$\text{s.t.} \quad g(w) = 0, \tag{1b}$$
$$h(w) \geq 0, \tag{1c}$$
$$0 \leq w_1 \perp w_2 \geq 0, \tag{1d}$$



$$w = (w_0, w_1, w_2) \in \mathbb{R}^n, \; w_0 \in \mathbb{R}^p, \; w_1, w_2 \in \mathbb{R}^m,$$

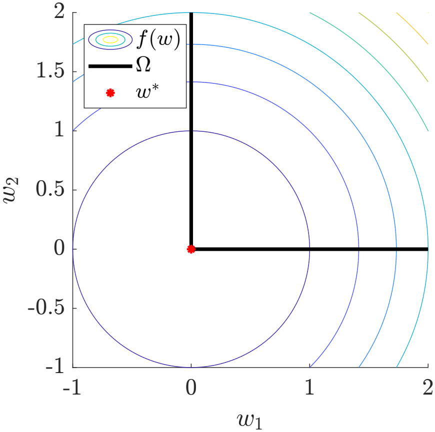
$$\Omega = \{x \in \mathbb{R}^n \mid g(w) = 0, h(w) \geq 0, \; 0 \leq w_1 \perp w_2 \geq 0\},$$

- Standard NLP methods solve the KKT conditions.
- MPECs violate constraint qualifications, and the KKT conditions may not be necessary.
- There are many stationary concepts for MPECs, and not all are useful.

Numerical methods for MPCCs

MPEC

$$\min_{w\in\mathbb{R}^n} \quad f(w) \tag{1a}$$
$$\text{s.t.} \quad g(w) = 0, \tag{1b}$$
$$h(w) \geq 0, \tag{1c}$$
$$0 \leq w_1 \perp w_2 \geq 0, \tag{1d}$$



$$w = (w_0, w_1, w_2) \in \mathbb{R}^n, \; w_0 \in \mathbb{R}^p, \; w_1, w_2 \in \mathbb{R}^m,$$

$$\Omega = \{x \in \mathbb{R}^n \mid g(w) = 0, h(w) \geq 0, \; 0 \leq w_1 \perp w_2 \geq 0\},$$

- Standard NLP methods solve the KKT conditions.
- MPECs violate constraint qualifications, and the KKT conditions may not be necessary.
- There are many stationary concepts for MPECs, and not all are useful.
- **Workaround/main idea**: solve a (finite) sequence of more regular problems.

Two families of MPCC methods

1. **Regularization, smoothing, penalty methods**: solve a sequence of regular nonlinear programs parametrized by a parameter $\sigma^k > 0$.
   - Pro: easy to implement.
   - Con: ill-conditioning, weaker theoretical properties.

Two families of MPCC methods

1. **Regularization, smoothing, penalty methods**: solve a sequence of regular nonlinear programs parametrized by a parameter $\sigma^k > 0$.
   - Pro: easy to implement.
   - Con: ill-conditioning, weaker theoretical properties.

2. **Combinatorial, active-set, pivoting methods**: solve a sequence of piece MPCCs (TNLP, BNLP), until a piece problem is found that solves also the MPCC.
   - Pro: strong theoretical properties, can be combined with 1.
   - Con: more difficult to implement, worst-case combinatorial complexity.

Two families of MPCC methods

1. **Regularization, smoothing, penalty methods**: solve a sequence of regular nonlinear programs parametrized by a parameter $\sigma^k > 0$.
   - Pro: easy to implement.
   - Con: ill-conditioning, weaker theoretical properties.
2. **Combinatorial, active-set, pivoting methods**: solve a sequence of piece MPCCs (TNLP, BNLP), until a piece problem is found that solves also the MPCC.
   - Pro: strong theoretical properties, can be combined with 1.
   - Con: more difficult to implement, worst-case combinatorial complexity.

Both classes of methods are available in `nosnoc`'s `mpccsol()` function:

`github.com/nosnoc/nosnoc`

To get started see: `github.com/nosnoc/nosnoc/tree/main/examples/generic_mpcc`

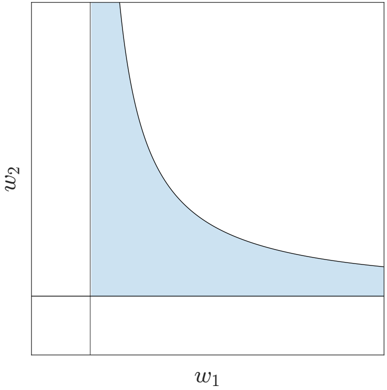
Outline of the lecture

Scholtes' global relaxation method

The easiest to implement and the most efficient regularization method [Scholtes, 2001].

Reg(σ^k)

$$\begin{aligned} \min_{w \in \mathbb{R}^n} \quad & f(w) \\ \text{s.t.} \quad & g(w) = 0, \\ & h(w) \geq 0, \\ & w_1, w_2 \geq 0, \\ & w_{1,i} w_{2,i} \leq \sigma^k, \ i = 1, \dots, m. \end{aligned}$$

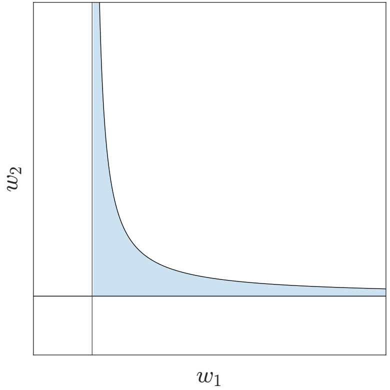


Scholtes' global relaxation method

The easiest to implement and the most efficient regularization method [Scholtes, 2001].

Reg(σ^k)

$$
\begin{aligned}
\min_{w \in \mathbb{R}^n} \quad & f(w) \\
\text{s.t.} \quad & g(w) = 0, \\
& h(w) \geq 0, \\
& w_1, w_2 \geq 0, \\
& w_{1,i} w_{2,i} \leq \sigma^k, \ i = 1, \ldots, m.
\end{aligned}
$$

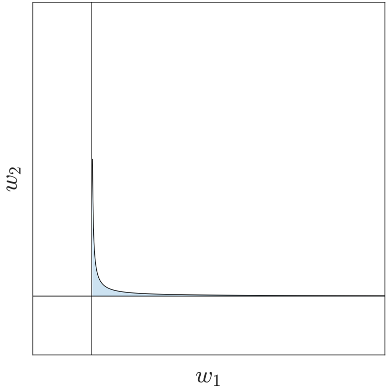


Scholtes' global relaxation method

The easiest to implement and the most efficient regularization method [Scholtes, 2001].

Reg(σ^k)

$$
\begin{aligned}
\min_{w \in \mathbb{R}^n} \quad & f(w) \\
\text{s.t.} \quad & g(w) = 0, \\
& h(w) \geq 0, \\
& w_1, w_2 \geq 0, \\
& w_{1,i} w_{2,i} \leq \sigma^k, \ i = 1, \dots, m.
\end{aligned}
$$

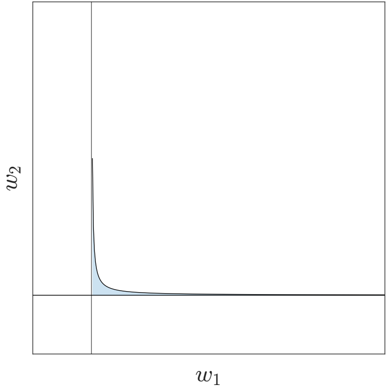


Scholtes' global relaxation method

The easiest to implement and the most efficient regularization method [Scholtes, 2001].

Reg(σ^k)

$$
\begin{aligned}
\min_{w \in \mathbb{R}^n} \quad & f(w) \\
\text{s.t.} \quad & g(w) = 0, \\
& h(w) \geq 0, \\
& w_1, w_2 \geq 0, \\
& w_{1,i} w_{2,i} \leq \sigma^k, \ i = 1, \dots, m.
\end{aligned}
$$



Theorem ([Scholtes, 2001, Hoheisel et al., 2013])

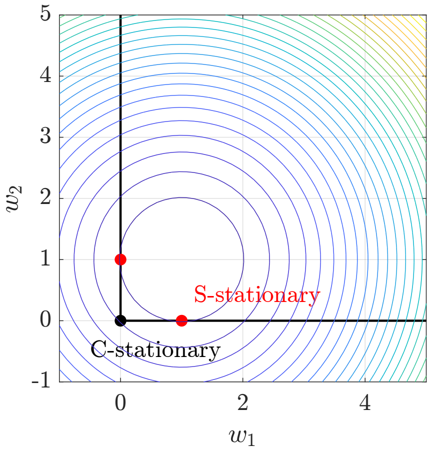
Let $\{\sigma^k\} \downarrow 0$ and let w^k be a stationary point of Reg(σ^k) with $w^k \to w^*$ such that MPEC-MFCQ holds at w^*. Then w^* is a C-stationary point of the the MPEC (1).

C-stationarity allows first-order descent directions

Consider the two-dimensional MPEC:

$$\begin{aligned} \min_{w \in \mathbb{R}^2} \quad & (w_1 - 1)^2 + (w_2 - 1)^2 \\ \text{s.t.} \quad & 0 \leq w_1 \perp w_2 \geq 0. \end{aligned}$$

- The origin $w^* = 0$ is an C-stationary point with the optimal multipliers $\nu = -2$, $\xi = 2$.

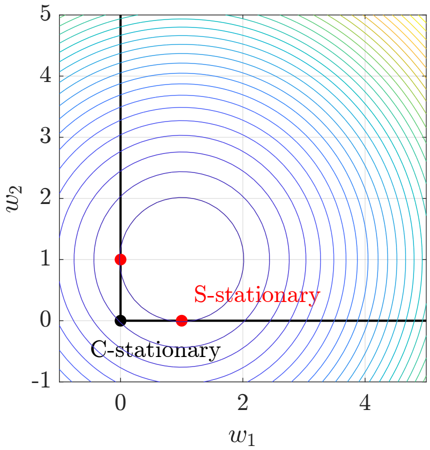


C-stationarity allows first-order descent directions

Consider the two-dimensional MPEC:

$$\begin{aligned} \min_{w \in \mathbb{R}^2} \quad & (w_1 - 1)^2 + (w_2 - 1)^2 \\ \text{s.t.} \quad & 0 \leq w_1 \perp w_2 \geq 0. \end{aligned}$$

- The origin $w^* = 0$ is an C-stationary point with the optimal multipliers $\nu = -2$, $\xi = 2$.
- There exists two descent direction $d = (1, 0)$ and $d = (0, 1)$
- The origin is not B-stationary, $w = (1, 0)$ and $w = (0, 1)$ are S-stationary.



Smoothing method have drawbacks

No benefit in using them. Have elaborate convergence theory [Ralph and Wright, 2004].

Instead of relaxing $w_1, w_2 \geq 0 \; w_1 w_2 \leq \sigma$, use smoothing $w_1, w_2 \geq 0$, $w_1 w_2 = \sigma$.

Consider the two-dimensional MPEC:

$$\begin{aligned} \min_{w \in \mathbb{R}^2} \quad & (w_1 + 1)^2 + (w_2 + 1)^2 \\ \text{s.t.} \quad & 0 \leq w_1 \perp w_2 \geq 0. \end{aligned}$$

Smoothing method have drawbacks

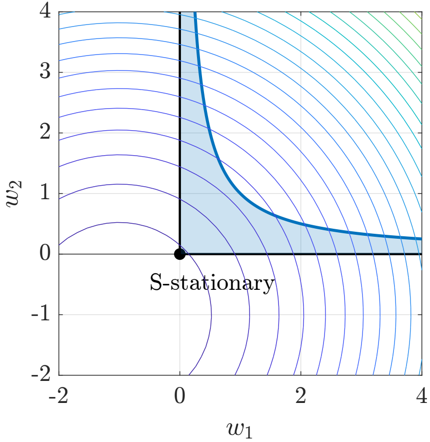
No benefit in using them. Have elaborate convergence theory [Ralph and Wright, 2004].

Instead of relaxing $w_1, w_2 \geq 0$ $w_1 w_2 \leq \sigma$, use smoothing $w_1, w_2 \geq 0$, $w_1 w_2 = \sigma$.

Consider the two-dimensional MPEC:

$$\begin{aligned} \min_{w \in \mathbb{R}^2} \quad & (w_1 + 1)^2 + (w_2 + 1)^2 \\ \text{s.t.} \quad & 0 \leq w_1 \perp w_2 \geq 0. \end{aligned}$$

- Relaxation method converges in one iteration to $w^* = (0, 0)$.



Smoothing method have drawbacks

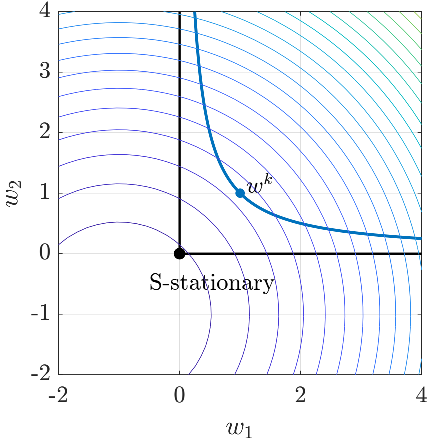
No benefit in using them. Have elaborate convergence theory [Ralph and Wright, 2004].

Instead of relaxing $w_1, w_2 \geq 0 \; w_1 w_2 \leq \sigma$, use smoothing $w_1, w_2 \geq 0$, $w_1 w_2 = \sigma$.

Consider the two-dimensional MPEC:

$$\min_{w \in \mathbb{R}^2} \quad (w_1 + 1)^2 + (w_2 + 1)^2$$
$$\text{s.t.} \quad 0 \leq w_1 \perp w_2 \geq 0.$$

- Relaxation method converges in one iteration to $w^* = (0, 0)$.
- Smoothing needs an infinite sequence for convergence.



Smoothing method have drawbacks

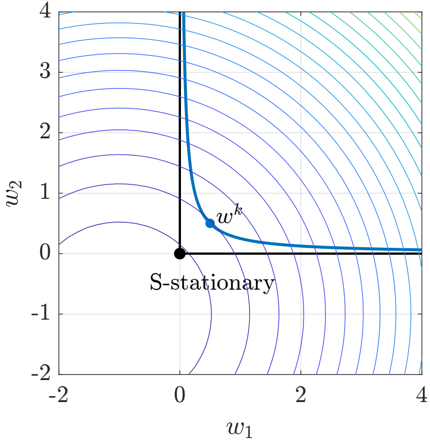
No benefit in using them. Have elaborate convergence theory [Ralph and Wright, 2004].

Instead of relaxing $w_1, w_2 \geq 0 \; w_1 w_2 \leq \sigma$, use smoothing $w_1, w_2 \geq 0$, $w_1 w_2 = \sigma$.

Consider the two-dimensional MPEC:

$$\min_{w \in \mathbb{R}^2} \quad (w_1 + 1)^2 + (w_2 + 1)^2$$
$$\text{s.t.} \quad 0 \leq w_1 \perp w_2 \geq 0.$$

- Relaxation method converges in one iteration to $w^* = (0, 0)$.
- Smoothing needs an infinite sequence for convergence.



Smoothing method have drawbacks

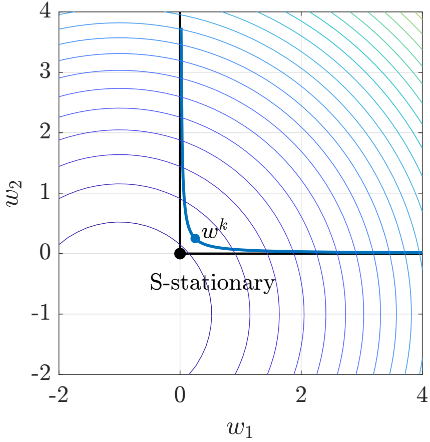
No benefit in using them. Have elaborate convergence theory [Ralph and Wright, 2004].

Instead of relaxing $w_1, w_2 \geq 0\ w_1 w_2 \leq \sigma$, use smoothing $w_1, w_2 \geq 0$, $w_1 w_2 = \sigma$.

Consider the two-dimensional MPEC:

$$\begin{aligned} \min_{w \in \mathbb{R}^2} \quad & (w_1+1)^2 + (w_2+1)^2 \\ \text{s.t.} \quad & 0 \leq w_1 \perp w_2 \geq 0. \end{aligned}$$

- Relaxation method converges in one iteration to $w^* = (0, 0)$.
- Smoothing needs an infinite sequence for convergence.



Smoothing method have drawbacks

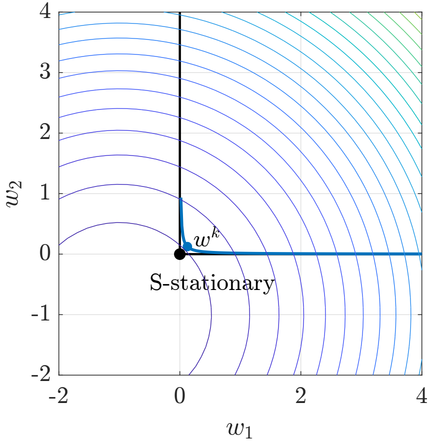
No benefit in using them. Have elaborate convergence theory [Ralph and Wright, 2004].

Instead of relaxing $w_1, w_2 \geq 0\ w_1 w_2 \leq \sigma$, use smoothing $w_1, w_2 \geq 0$, $w_1 w_2 = \sigma$.

Consider the two-dimensional MPEC:

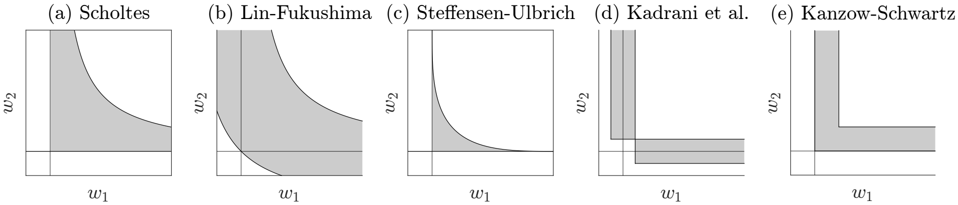
$$\begin{aligned} \min_{w \in \mathbb{R}^2} \quad & (w_1+1)^2 + (w_2+1)^2 \\ \text{s.t.} \quad & 0 \leq w_1 \perp w_2 \geq 0. \end{aligned}$$

- Relaxation method converges in one iteration to $w^* = (0, 0)$.
- Smoothing needs an infinite sequence for convergence.



Other regularization methods

There exist many elaborate ways to relax the L-shaped set. Convergence theory in [Hoheisel et al., 2013]

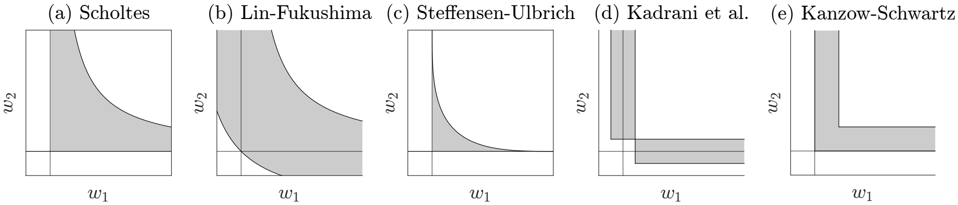


- They have better convergence properties than Scholtes' method if the NLP's are solved exactly.
- In practice, they perform better only on easier problems [Nurkanović et al., 2024].

:

Other regularization methods

There exist many elaborate ways to relax the L-shaped set. Convergence theory in [Hoheisel et al., 2013]



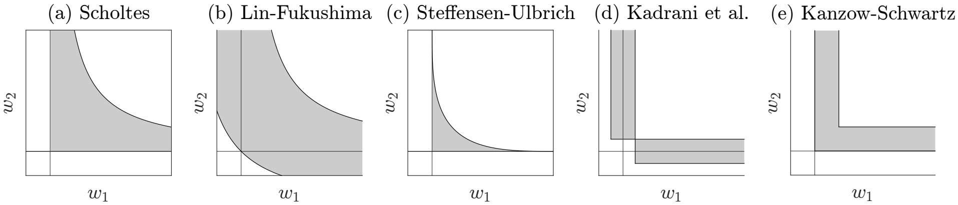
- They have better convergence properties than Scholtes' method if the NLP's are solved exactly.
- In practice, they perform better only on easier problems [Nurkanović et al., 2024].

Convergence results [Hoheisel et al., 2013]:

a) [Scholtes, 2001]: C-stationarity under MPEC-MFCQ.

Other regularization methods

There exist many elaborate ways to relax the L-shaped set. Convergence theory in [Hoheisel et al., 2013]



- They have better convergence properties than Scholtes' method if the NLP's are solved exactly.
- In practice, they perform better only on easier problems [Nurkanović et al., 2024].

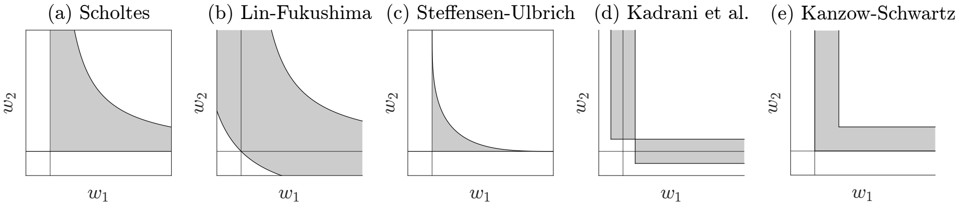
Convergence results [Hoheisel et al., 2013]:

a) [Scholtes, 2001]: C-stationarity under MPEC-MFCQ.

b) [Lin and Fukushima, 2005]: C-stationarity under MPEC-MFCQ..

Other regularization methods

There exist many elaborate ways to relax the L-shaped set. Convergence theory in [Hoheisel et al., 2013]



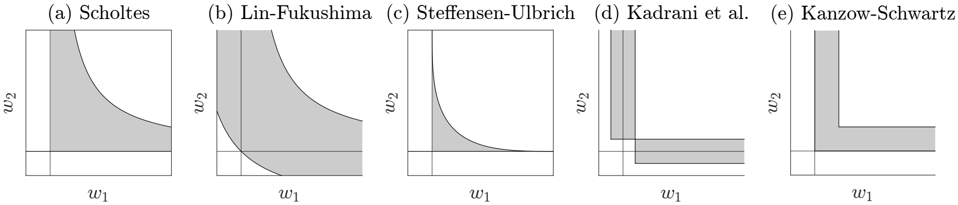
- They have better convergence properties than Scholtes' method if the NLP's are solved exactly.
- In practice, they perform better only on easier problems [Nurkanović et al., 2024].

Convergence results [Hoheisel et al., 2013]:

a) [Scholtes, 2001]: C-stationarity under MPEC-MFCQ.
b) [Lin and Fukushima, 2005]: C-stationarity under MPEC-MFCQ..
c) [Steffensen and Ulbrich, 2010]: C-stationarity under MPEC-CPLD.

Other regularization methods

There exist many elaborate ways to relax the L-shaped set. Convergence theory in [Hoheisel et al., 2013]



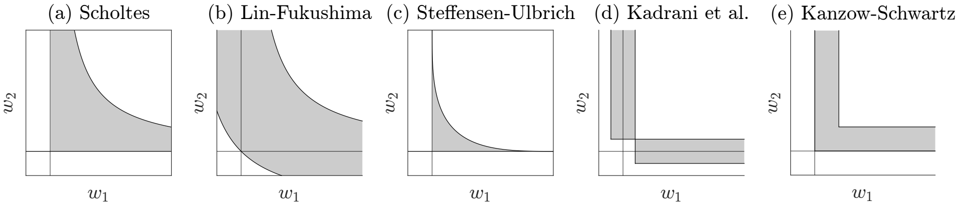
- They have better convergence properties than Scholtes' method if the NLP's are solved exactly.
- In practice, they perform better only on easier problems [Nurkanović et al., 2024].

Convergence results [Hoheisel et al., 2013]:

a) [Scholtes, 2001]: C-stationarity under MPEC-MFCQ.
b) [Lin and Fukushima, 2005]: C-stationarity under MPEC-MFCQ..
c) [Steffensen and Ulbrich, 2010]: C-stationarity under MPEC-CPLD.
d) [Kadrani et al., 2009]: M-stationarity under MPEC-CPLD.

Other regularization methods

There exist many elaborate ways to relax the L-shaped set. Convergence theory in [Hoheisel et al., 2013]



- They have better convergence properties than Scholtes' method if the NLP's are solved exactly.
- In practice, they perform better only on easier problems [Nurkanović et al., 2024].

Convergence results [Hoheisel et al., 2013]:

a) [Scholtes, 2001]: C-stationarity under MPEC-MFCQ.
b) [Lin and Fukushima, 2005]: C-stationarity under MPEC-MFCQ..
c) [Steffensen and Ulbrich, 2010]: C-stationarity under MPEC-CPLD.
d) [Kadrani et al., 2009]: M-stationarity under MPEC-CPLD.
e) [Kanzow and Schwartz, 2013]: M-stationarity under MPEC-CPLD.

“The price of inexactness” [Kanzow and Schwartz, 2015]

- Most convergence results derived under the assumption that every subproblem for a fixed σ^k is solved exactly.
- If NLPs are solved to some $\epsilon > 0$ threshold, than most methods have weaker convergence properties.

“The price of inexactness” [Kanzow and Schwartz, 2015]

- Most convergence results derived under the assumption that every subproblem for a fixed σ^k is solved exactly.
- If NLPs are solved to some $\epsilon > 0$ threshold, than most methods have weaker convergence properties.
- Steffensen-Ulbrich, Kadrani, Kanzow-Schwartz converge only to **weakly** stationary points!

"The price of inexactness" [Kanzow and Schwartz, 2015]

- Most convergence results derived under the assumption that every subproblem for a fixed σ^k is solved exactly.
- If NLPs are solved to some $\epsilon > 0$ threshold, than most methods have weaker convergence properties.
- Steffensen-Ulbrich, Kadrani, Kanzow-Schwartz converge only to **weakly** stationary points!
- Scholtes' method still converges to C-stationary points.
- A nice implementation for Scholtes' method is in pair it with interior-point methods, where the barrier τ and homotopy σ parameters are jointly driven to zero [Raghunathan and Biegler, 2005] (`IPOPT-C`).

Penalty methods

Closely related to regularization methods, sometimes with a one-to-one correspondence [Leyffer et al., 2006]

Main idea: put difficult part into objective.

The ℓ_1 reformulation

$$
\begin{aligned}
\min_{w \in \mathbb{R}^n} \quad & f(w) + \rho w_1^\top w_2 \\
\text{s.t.} \quad & g(w) = 0, \\
& h(w) \geq 0, \\
& w_1, w_2 \geq 0.
\end{aligned}
$$

The ℓ_∞ reformulation

$$
\begin{aligned}
\min_{w \in \mathbb{R}^n, s \in \mathbb{R}} \quad & f(w) + \rho s \\
\text{s.t.} \quad & g(w) = 0, \\
& h(w) \geq 0, \\
& w_1, w_2 \geq 0, \\
& w_{1,i} w_{2,i} \leq s, \; i = 1 \dots m, \\
& 0 \leq s \leq \bar{s}.
\end{aligned}
$$

Penalty methods

Closely related to regularization methods, sometimes with a one-to-one correspondence [Leyffer et al., 2006]

Main idea: put difficult part into objective.

The ℓ_1 reformulation

$$\begin{aligned} \min_{w \in \mathbb{R}^n} \quad & f(w) + \rho w_1^\top w_2 \\ \text{s.t.} \quad & g(w) = 0, \\ & h(w) \geq 0, \\ & w_1, w_2 \geq 0. \end{aligned}$$

The ℓ_∞ reformulation

$$\begin{aligned} \min_{w \in \mathbb{R}^n, s \in \mathbb{R}} \quad & f(w) + \rho s \\ \text{s.t.} \quad & g(w) = 0, \\ & h(w) \geq 0, \\ & w_1, w_2 \geq 0, \\ & w_{1,i} w_{2,i} \leq s, \; i = 1 \dots m, \\ & 0 \leq s \leq \bar{s}. \end{aligned}$$

- If the penalty parameter ρ large enough, solve single NLP to solve MPEC [Anitescu, 2005a, Anitescu, 2005b, Ralph and Wright, 2004].
- Good implementations update the penalty parameter in a homotopy $\rho^k = \frac{1}{\sigma^k}$
- In practice, usually converges faster than Scholtes' method, but has a lower success rate.

Steering the homotopy parameter to zero

The practical performance depends on the update rate.

Approach: Solve a sequence of regularized NLPs (σ^k), warm start the next iteration with $w^*(\sigma^{k-1})$. Update the homotopy parameter via:

- Linear update rule:

$$\sigma^{k+1} = \kappa\sigma^k, \ \kappa \in (0,1)$$

- For faster convergence, superlinear rules may also be helpful:

$$\sigma^{k+1} = \min(\kappa\sigma^k, (\sigma^k)^\eta), \ \kappa \in (0,1), \ \eta > 1$$

Steering the homotopy parameter to zero

The practical performance depends on the update rate.

Approach: Solve a sequence of regularized NLPs (σ^k), warm start the next iteration with $w^*(\sigma^{k-1})$. Update the homotopy parameter via:

- Linear update rule:

$$\sigma^{k+1} = \kappa\sigma^k, \;\; \kappa \in (0,1)$$

- For faster convergence, superlinear rules may also be helpful:

$$\sigma^{k+1} = \min(\kappa\sigma^k, (\sigma^k)^\eta), \;\; \kappa \in (0,1), \;\; \eta > 1$$

Some best practices:

- Start with not too large, and not too small σ^0. For well-scaled problems $\sigma^0 = 1$ usually works well.
- Do not update too aggressively nor top conservatively. Good choices are $\kappa = 0.1$, $\kappa = 0.2$ and $\eta = 1.5$.
- Update rules are usually monotone. To increase robustness, if an iteration fails, go back and update with $\kappa^+ = \gamma\kappa$, e.g., $\gamma = 2$.

Outline of the lecture

Difficulties:

1. Regularization methods, also under very strong assumptions, converge to points that are weaker than S-stationary, which are possibly not B-stationary.
2. Regularized NLPs, small or large σ, may be extremely difficult to solve (much more than TNLPs/BNLPs).

Motivation for active set methods for MPECs

Difficulties:

1. Regularization methods, also under very strong assumptions, converge to points that are weaker than S-stationary, which are possibly not B-stationary.
2. Regularized NLPs, small or large σ, may be extremely difficult to solve (much more than TNLPs/BNLPs).

Main idea of active-set methods:

1. Find a feasible point $w \in \Omega$ of the MPEC (1).
2. Solved the branch NLP with some partition of the index set, e.g. $\mathcal{I}_1(w)$, $\mathcal{I}_2(w)$
3. Call an oracle to verify B-stationary or to give a better $\mathcal{I}_1(w^+)$ and $\mathcal{I}_2(w^+)$ (e.g. an LPEC)
4. If not B-stationary, go to 2.

A brief history of active-set methods

- **Late 1990s, early 2000s**: selecting the next TNLP/BNLP based on signs of multipliers [Fukushima and Tseng, 2002, Giallombardo and Ralph, 2008, Izmailov and Solodov, 2008, Jiang and Ralph, 1999, Lin and Fukushima, 2006, Liu et al., 2006, Luo et al., 1996, Scholtes and Stöhr, 1999]. Convergence to B-stationarity can only be guaranteed under MPEC-LICQ.
- **2007**: In [Leyffer and Munson, 2007] for the first time suggested to use LPECs as a stopping criteria and for step computation.
- **2022**: SQP-type methods with LPECs, developed for MPECs with bound constraints [Kirches et al., 2022] (B-stationary), extension for general constraints via augmented Lagrangian [Guo and Deng, 2022] (M-stationary).
- **2024-2025**: for general MPECs (B-stationary) in [Kazi et al., 2024] and [N. and Leyffer, 2025].

Refresher on the the Branch Nonlinear Program (BNLP)

MPEC active sets

$$\mathcal{I}_{+0}(w) = \{i \in \{1,\dots,m\} \mid w_{1,i} > 0, w_{2,i} = 0\},$$
$$\mathcal{I}_{0+}(w) = \{i \in \{1,\dots,m\} \mid w_{1,i} = 0, w_{2,i} > 0\},$$
$$\mathcal{I}_{00}(w) = \{i \in \{1,\dots,m\} \mid w_{1,i} = 0, w_{2,i} = 0\}.$$

Partitioning of the degenerate set $\mathcal{I}_{00}(w$

$\mathcal{D}_1(w) \cup \mathcal{D}_2(w) = \mathcal{I}_{00}(w)$,
$\mathcal{D}_1(w) \cap \mathcal{D}_2(w) = \emptyset$,
$\mathcal{I}_1(w) := \mathcal{I}_{0+}(w) \cup \mathcal{D}_1(w)$,
$\mathcal{I}_2(w) := \mathcal{I}_{+0}(w) \cup \mathcal{D}_2(w)$,

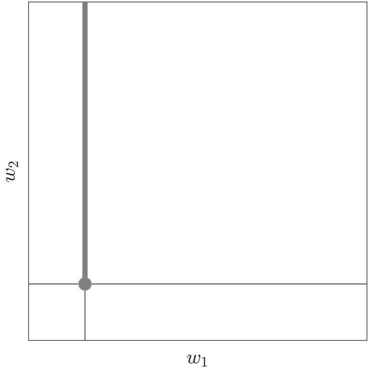
Refresher on the the Branch Nonlinear Program (BNLP)

MPEC active sets

$$\mathcal{I}_{+0}(w) = \{i \in \{1,\dots,m\} \mid w_{1,i} > 0, w_{2,i} = 0\},$$
$$\mathcal{I}_{0+}(w) = \{i \in \{1,\dots,m\} \mid w_{1,i} = 0, w_{2,i} > 0\},$$
$$\mathcal{I}_{00}(w) = \{i \in \{1,\dots,m\} \mid w_{1,i} = 0, w_{2,i} = 0\}.$$

Partitioning of the degenerate set $\mathcal{I}_{00}(w$

$$\mathcal{D}_1(w) \cup \mathcal{D}_2(w) = \mathcal{I}_{00}(w),$$
$$\mathcal{D}_1(w) \cap \mathcal{D}_2(w) = \emptyset,$$
$$\mathcal{I}_1(w) := \mathcal{I}_{0+}(w) \cup \mathcal{D}_1(w),$$
$$\mathcal{I}_2(w) := \mathcal{I}_{+0}(w) \cup \mathcal{D}_2(w),$$



Branch NLP (BNLP) at w^*

$$\begin{aligned} \min_{w \in \mathbb{R}^n} \quad & f(w) \\ \text{s.t.} \quad & g(w) = 0, \\ & h(w) \geq 0, \\ & w_{1,i} = 0,\ w_{2,i} \geq 0,\ i \in \mathcal{I}_1(w^*) \\ & w_{1,i} \geq 0,\ w_{2,i} = 0,\ i \in \mathcal{I}_2(w^*). \end{aligned}$$

Refresher on the the Branch Nonlinear Program (BNLP)

MPEC active sets

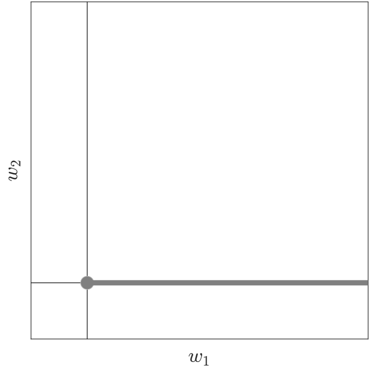
$$\mathcal{I}_{+0}(w) = \{i \in \{1,\dots,m\} \mid w_{1,i} > 0, w_{2,i} = 0\},$$
$$\mathcal{I}_{0+}(w) = \{i \in \{1,\dots,m\} \mid w_{1,i} = 0, w_{2,i} > 0\},$$
$$\mathcal{I}_{00}(w) = \{i \in \{1,\dots,m\} \mid w_{1,i} = 0, w_{2,i} = 0\}.$$

Partitioning of the degenerate set $\mathcal{I}_{00}(w$

$$\mathcal{D}_1(w) \cup \mathcal{D}_2(w) = \mathcal{I}_{00}(w),$$
$$\mathcal{D}_1(w) \cap \mathcal{D}_2(w) = \emptyset,$$
$$\mathcal{I}_1(w) := \mathcal{I}_{0+}(w) \cup \mathcal{D}_1(w),$$
$$\mathcal{I}_2(w) := \mathcal{I}_{+0}(w) \cup \mathcal{D}_2(w),$$

Branch NLP (BNLP) at w^*

$$\begin{aligned}
\min_{w \in \mathbb{R}^n} \quad & f(w) \\
\text{s.t.} \quad & g(w) = 0, \\
& h(w) \geq 0, \\
& w_{1,i} = 0,\ w_{2,i} \geq 0,\ i \in \mathcal{I}_1(w^*) \\
& w_{1,i} \geq 0,\ w_{2,i} = 0,\ i \in \mathcal{I}_2(w^*).
\end{aligned}$$



Refresher on LPECs

If there are no complementarity constraints, the LPEC reduces to an LP.

Theorem ([Luo et al., 1996])

Let $\mathcal{T}_\Omega(w) = \mathcal{T}_\Omega^{\mathrm{MPEC}}(w)$, *and* $w^* \in \Omega$ *be a local minimizer of the MPEC* (1), *then it holds that*

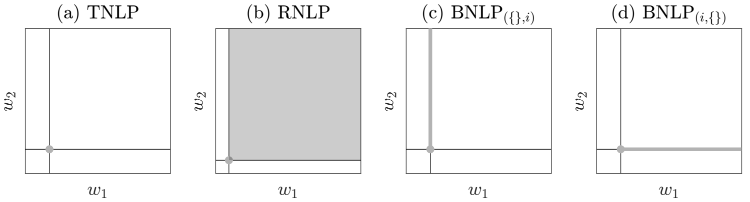
$$\nabla f(w^*)^\top d \geq 0 \;\; \textit{for all } d \in \mathcal{T}_\Omega^{\mathrm{MPEC}}(w^*), \tag{2}$$

or equivalently, $d = 0$ *is a local minimizer of the following optimization problem:*

$$\begin{aligned}
\min_{d \in \mathbb{R}^n} \quad & \nabla f(w^k)^\top d \\
s.t. \quad & g(w^k) + \nabla g(w^k)^\top d = 0, \\
& h(w^k) + \nabla h(w^k)^\top d \geq 0, \\
& w_{1,i}^k + d_{1,i} = 0, \forall i \in \mathcal{I}_{0+}^k, \\
& w_{2,i}^k + d_{2,i} = 0, \forall i \in \mathcal{I}_{+0}^k, \\
& 0 \leq w_{1,i}^k + d_{1,i} \perp w_{2,i}^k + d_{2,i} \geq 0, \forall i \in \mathcal{I}_{00}^k
\end{aligned}$$

MPEC feasible set relations

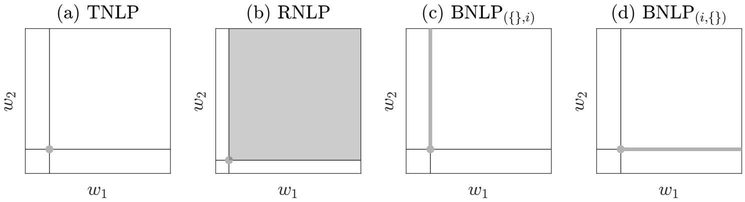
$$\Omega_{\mathrm{TNLP}} = \bigcap_{(\mathcal{I}_1,\mathcal{I}_2)} \Omega_{\mathrm{BNLP}(\mathcal{I}_1,\mathcal{I}_2)} \subset \Omega = \bigcup_{(\mathcal{I}_1,\mathcal{I}_2)} \Omega_{\mathrm{BNLP}(\mathcal{I}_1,\mathcal{I}_2)}.$$



Refresher on the relation between MPEC pieces

MPEC feasible set relations

$$\Omega_{\mathrm{TNLP}} = \bigcap_{(\mathcal{I}_1,\mathcal{I}_2)} \Omega_{\mathrm{BNLP}(\mathcal{I}_1,\mathcal{I}_2)} \subset \Omega = \bigcup_{(\mathcal{I}_1,\mathcal{I}_2)} \Omega_{\mathrm{BNLP}(\mathcal{I}_1,\mathcal{I}_2)}.$$



Summary:

- The TNLP, RNLP, and BNLPs are regular nonlinear optimization problems.
- If we know the right TNLP/BNLP, we can just solve a regular NLP to solve the MPEC.
- There are $2^{|\mathcal{I}_{00}|}$ BNLPs - highlighting the combinatorial nature.

MPECopt: An algorithm for computing a B-stationary points

Algorithm `MPECopt` (simplified)

Input: w^0,$\rho^0 \in [\rho^{\mathrm{lb}}, \rho^{\mathrm{ub}}]$, $\gamma^{\mathrm{L}} \in (0,1)$, $\gamma^{\mathrm{U}} > 1$

Call Phase I Algorithm to find a feasible point $w^0 \in \Omega$
for $k = 0, \ldots$ **do** // Major/outer loop
  Possibly reset trust region radius $\rho^{k,0} \in [\rho^{\mathrm{lb}}, \rho^{\mathrm{ub}}]$ **for** $l = 0, \ldots$ **do** // Minor/inner loop
    Solve $\mathrm{LPEC}(w^k, \rho^{k,l})$ for $d^{k,l}$ **if** $d^{k,l} = 0$ **then**
      terminate // (B-stationary point found)
    else if *Active set changed* **then**
      Solve $\mathrm{BNLP}(w^k + d^{k,l})$ for $w^{k,l}$
      if $f(w^{k,l}) < f(w^k)$ **then**
        $w^{k+1} = w^{k,l}$; $\rho^{k,l+1} = \gamma^{\mathrm{U}} \rho^{k,l}$; **break** // step accepted
      else $\rho^{k,l+1} = \gamma^{\mathrm{L}} \rho^{k,l}$ // reduce TR-radius
    else $\rho^{k,l+1} = \gamma^{\mathrm{L}} \rho^{k,l}$ // reduce TR-radius

Algorithm for computing a B-stationary point of MPECs

$\mathrm{LPEC}(w^k, \rho)$ - reduced

$$
\begin{aligned}
\min_{d \in \mathbb{R}^n} \quad & \nabla f(w^k)^\top d \\
\text{s.t.} \quad & g(w^k) + \nabla g(w^k)^\top d = 0, \\
& h(w^k) + \nabla h(w^k)^\top d \geq 0, \\
& w_{1,i}^k + d_{1,i} = 0, \ w_{2,i}^k + d_{2,i} \geq 0, \forall i \in \mathcal{I}_{0+}, \\
& w_{1,i}^k + d_{1,i} \geq 0, \ w_{2,i}^k + d_{2,i} = 0, \forall i \in \mathcal{I}_{+0} \\
& 0 \leq w_{1,i}^k + d_{1,i} \perp w_{2,i}^k + d_{2,i} \geq 0, \forall i \in \mathcal{I}_{00}^k, \\
& \|d\|_\infty \leq \rho,
\end{aligned}
$$

Algorithm for computing a B-stationary point of MPECs

LPEC(w^k, ρ) - full

$$
\begin{aligned}
\min_{d \in \mathbb{R}^n} \quad & \nabla f(w^k)^\top d \\
\text{s.t.} \quad & g(w^k) + \nabla g(w^k)^\top d = 0, \\
& h(w^k) + \nabla h(w^k)^\top d \geq 0, \\
& 0 \leq w_{1,i}^k + d_{1,i} \perp w_{2,i}^k + d_{2,i} \geq 0, \forall i, \\
& \|d\|_\infty \leq \rho,
\end{aligned}
$$

Algorithm for computing a B-stationary point of MPECs

$\mathrm{LPEC}(w^k, \rho)$ - full

$$\begin{aligned}
\min_{d \in \mathbb{R}^n} \quad & \nabla f(w^k)^\top d \\
\text{s.t.} \quad & g(w^k) + \nabla g(w^k)^\top d = 0, \\
& h(w^k) + \nabla h(w^k)^\top d \geq 0, \\
& 0 \leq w_{1,i}^k + d_{1,i} \perp w_{2,i}^k + d_{2,i} \geq 0, \forall i, \\
& \|d\|_\infty \leq \rho,
\end{aligned}$$

$\mathrm{BNLP}(w^k)$

$$\begin{aligned}
\min_{x \in \mathbb{R}^n} \quad & f(x) \\
\text{s.t.} \quad & g(w) = 0, \\
& h(w) \geq 0, \\
& w_{1,i} = 0, \ w_{2,i} \geq 0, \ i \in \mathcal{I}_1^k, \\
& w_{1,i} \geq 0, \ w_{2,i} = 0, \ i \in \mathcal{I}_2^k.
\end{aligned}$$

Algorithm for computing a B-stationary point of MPECs

$\mathrm{LPEC}(w^k, \rho)$ - full

$$\begin{aligned}
\min_{d \in \mathbb{R}^n} \quad & \nabla f(w^k)^\top d \\
\text{s.t.} \quad & g(w^k) + \nabla g(w^k)^\top d = 0, \\
& h(w^k) + \nabla h(w^k)^\top d \geq 0, \\
& 0 \leq w^k_{1,i} + d_{1,i} \perp w^k_{2,i} + d_{2,i} \geq 0, \forall i, \\
& \|d\|_\infty \leq \rho,
\end{aligned}$$

$\mathrm{BNLP}(w^k)$

$$\begin{aligned}
\min_{x \in \mathbb{R}^n} \quad & f(x) \\
\text{s.t.} \quad & g(w) = 0, \\
& h(w) \geq 0, \\
& w_{1,i} = 0, \; w_{2,i} \geq 0, \; i \in \mathcal{I}_1^k, \\
& w_{1,i} \geq 0, \; w_{2,i} = 0, \; i \in \mathcal{I}_2^k.
\end{aligned}$$

Main steps:

1. Solve $\mathrm{LPEC}(w^k, \rho)$ to determine the active set for $\mathrm{BNLP}(w^k + d)$ - or verify B-stationarity. Don't use d for iterate update.

Algorithm for computing a B-stationary point of MPECs

$\mathrm{LPEC}(w^k, \rho)$ - full

$$
\begin{aligned}
\min_{d \in \mathbb{R}^n} \quad & \nabla f(w^k)^\top d \\
\text{s.t.} \quad & g(w^k) + \nabla g(w^k)^\top d = 0, \\
& h(w^k) + \nabla h(w^k)^\top d \geq 0, \\
& 0 \leq w_{1,i}^k + d_{1,i} \perp w_{2,i}^k + d_{2,i} \geq 0, \forall i, \\
& \|d\|_\infty \leq \rho,
\end{aligned}
$$

$\mathrm{BNLP}(w^k)$

$$
\begin{aligned}
\min_{x \in \mathbb{R}^n} \quad & f(x) \\
\text{s.t.} \quad & g(w) = 0, \\
& h(w) \geq 0, \\
& w_{1,i} = 0,\ w_{2,i} \geq 0,\ i \in \mathcal{I}_1^k, \\
& w_{1,i} \geq 0,\ w_{2,i} = 0,\ i \in \mathcal{I}_2^k.
\end{aligned}
$$

Main steps:

1. Solve $\mathrm{LPEC}(w^k, \rho)$ to determine the active set for $\mathrm{BNLP}(w^k + d)$ - or verify B-stationarity. Don't use d for iterate update.
2. Solve $\mathrm{BNLP}(w^k + d)$ for accuracy: if decrease in objective accept step, else: resolve LPEC with smaller ρ.

Full vs reduced LPEC

$\mathrm{LPEC}(w^k, \rho)$ - full

$$
\begin{aligned}
\min_{d \in \mathbb{R}^n} \quad & \nabla f(w^k)^\top d \\
\text{s.t.} \quad & g(w^k) + \nabla g(w^k)^\top d = 0, \\
& h(w^k) + \nabla h(w^k)^\top d \geq 0, \\
& 0 \leq w_{1,i}^k + d_{1,i} \perp w_{2,i}^k + d_{2,i} \geq 0, \forall i, \\
& \|d\|_\infty \leq \rho,
\end{aligned}
$$

- If w^k infeasible, then not clear how to define index sets in reduced LPEC.
- If w^k feasible, then $d = 0$ feasible.
- If w^k not B-stationary, then $\nabla f(w^k)^\top d < 0$.
- Use $w^k + d$ for next active set guess.
- More complementarity constraints, more computationally expensive? Depends on ρ.

Full vs reduced LPEC

$\mathrm{LPEC}(w^k, \rho)$ - full

$$
\begin{aligned}
\min_{d \in \mathbb{R}^n} \quad & \nabla f(w^k)^\top d \\
\text{s.t.} \quad & g(w^k) + \nabla g(w^k)^\top d = 0, \\
& h(w^k) + \nabla h(w^k)^\top d \geq 0, \\
& 0 \leq w^k_{1,i} + d_{1,i} \perp w^k_{2,i} + d_{2,i} \geq 0, \forall i, \\
& \|d\|_\infty \leq \rho,
\end{aligned}
$$

- If w^k infeasible, then not clear how to define index sets in reduced LPEC.
- If w^k feasible, then $d = 0$ feasible.
- If w^k not B-stationary, then $\nabla f(w^k)^\top d < 0$.
- Use $w^k + d$ for next active set guess.
- More complementarity constraints, more computationally expensive? Depends on ρ.

Lemma

Let $w \in \Omega$ be a feasible point of the MPEC (1). For all trust region radii that satisfy

$$
0 < \rho < \bar{\rho} = \min\{\{w_{1,i} \mid i \in \mathcal{I}^k_{+0}\} \cup \{w_{2,i} \mid i \in \mathcal{I}^k_{0+}\}\}, \tag{3}
$$

the sets of the local minimizers of the reduced and full LPECs are identical. In the special case of $\mathcal{I}_{00} = \{1, \ldots, m\}$ the reduced and full LPEC coincide and in that case, $\bar{\rho} = \infty$.

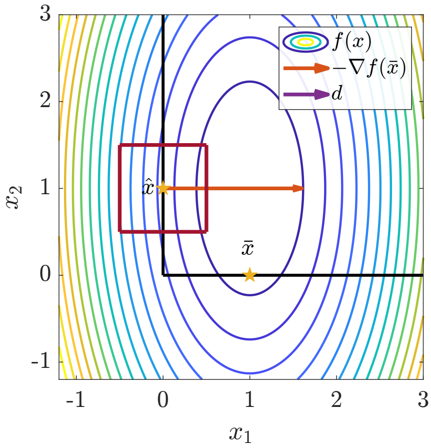
Example: nonconvexity of LPECs and better/worse minima

Assumption in example: LPEC solved to global optimality (turns out: not so restrictive in practice).

Consider the two-dimensional MPEC:

$$\begin{aligned} \min_{w \in \mathbb{R}^2} \quad & 4(w_1 - 1)^2 + (w_2 - 1)^2 \\ \text{s.t.} \quad & 0 \leq w_1 \perp w_2 \geq 0. \end{aligned}$$

- Two B-stationary points $\bar{w} = (1, 0)$ and $\hat{w} = (0, 1)$, with $f(\bar{w}) = 1$ and $f(\hat{w}) = 4$.
- If ρ is sufficiently small, $\hat{w}$ with $f(\hat{w}) = 4$ is verified.



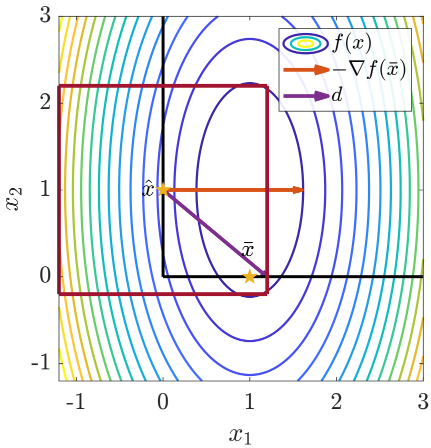
Example: nonconvexity of LPECs and better/worse minima

Assumption in example: LPEC solved to global optimality (turns out: not so restrictive in practice).

Consider the two-dimensional MPEC:

$$\min_{w \in \mathbb{R}^2} \quad 4(w_1 - 1)^2 + (w_2 - 1)^2$$
$$\text{s.t.} \quad 0 \leq w_1 \perp w_2 \geq 0.$$

- Two B-stationary points $\bar{w} = (1, 0)$ and $\hat{w} = (0, 1)$, with $f(\bar{w}) = 1$ and $f(\hat{w}) = 4$.
- If ρ is sufficiently small, $\hat{w}$ with $f(\hat{w}) = 4$ is verified.
- If ρ large, the globally optimal LPEC solution finds a BNLP with $f(\bar{w}) = 1$.

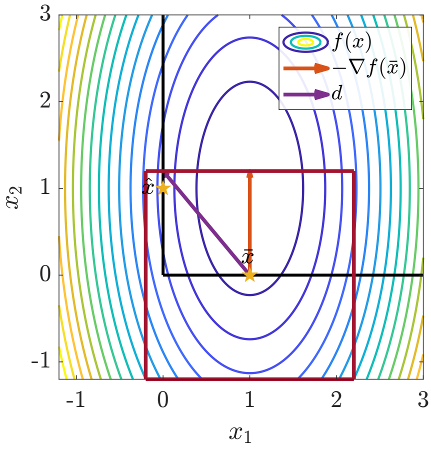


Example: nonconvexity of LPECs and better/worse minima

Assumption in example: LPEC solved to global optimality (turns out: not so restrictive in practice).

Consider the two-dimensional MPEC:

$$\begin{aligned} \min_{w \in \mathbb{R}^2} \quad & 4(w_1 - 1)^2 + (w_2 - 1)^2 \\ \text{s.t.} \quad & 0 \leq w_1 \perp w_2 \geq 0. \end{aligned}$$



- Two B-stationary points $\bar{w} = (1, 0)$ and $\hat{w} = (0, 1)$, with $f(\bar{w}) = 1$ and $f(\hat{w}) = 4$.
- If ρ is sufficiently small, $\hat{w}$ with $f(\hat{w}) = 4$ is verified.
- If ρ large, the globally optimal LPEC solution finds a BNLP with $f(\bar{w}) = 1$.
- Conversely, start $\bar{w}$, the full LPEC finds a BNLP with a larger objective, step rejected, ρ reduced.

On solving LPECs

Solution methods for LPECs:

1. **Regularization and penalty methods**.
   - Pro: easy to implement.
   - Con: not guaranteed to verify $d = 0$ as B-stationary point of LPEC.

On solving LPECs

Solution methods for LPECs:

1. **Regularization and penalty methods**.
   - Pro: easy to implement.
   - Con: not guaranteed to verify $d = 0$ as B-stationary point of LPEC.
2. **Tailored LPEC methods**, extensions of the simplex method, branch and cut.
   - Pro: guaranteed to verify $d = 0$ as B-stationary point of LPEC.
   - Con: difficult to implement.

On solving LPECs

Solution methods for LPECs:

1. **Regularization and penalty methods**.
   - Pro: easy to implement.
   - Con: not guaranteed to verify $d = 0$ as B-stationary point of LPEC.
2. **Tailored LPEC methods**, extensions of the simplex method, branch and cut.
   - Pro: guaranteed to verify $d = 0$ as B-stationary point of LPEC.
   - Con: difficult to implement.
3. **Mixed-integer reformulation:** reformulate into equivalent mixed-integer linear program (MILP).
   - Pro: trivial to implement, find global minimum.
   - Con: worst case exponential complexity.

LPEC as MILP

$\mathrm{LPEC}(w^k, \rho)$ as MILP

$$\min_{\substack{d\in\mathbb{R}^n,\\ y\in\{0,1\}^m}} \quad \nabla f(w^k)^\top d \tag{4a}$$
$$\text{s.t.} \quad g(w^k) + \nabla g(w^k)^\top d = 0, \tag{4b}$$
$$h(w^k) + \nabla h(w^k)^\top d \geq 0, \tag{4c}$$
$$0 \leq w_1^k + d_1 \leq yM, \tag{4d}$$
$$0 \leq w_2^k + d_2 \leq (e-y)M, \tag{4e}$$
$$\|d\|_\infty \leq \rho. \tag{4f}$$

- Trust region $\|d\| \leq \rho$ makes feasible set compact.
- bigM dominated by ρ, just needs to be large enough for feasiblity $M = \max((w_1^k, w_2^k)) + \rho$.
- At feasible points, ρ can be arbitaly small - very tight relaxations.
- BNLP index sets:

$$\mathcal{I}_1(w^k + d) = \{i \mid y_i = 0\},$$
$$\mathcal{I}_2(w^k + d) = \{i \mid y_i = 1\}.$$

LPEC as MILP

LPEC(w^k, ρ) as MILP

$$\min_{\substack{d \in \mathbb{R}^n,\\ y \in \{0,1\}^m}} \quad \nabla f(w^k)^\top d \tag{4a}$$

$$\text{s.t.} \quad g(w^k) + \nabla g(w^k)^\top d = 0, \tag{4b}$$

$$h(w^k) + \nabla h(w^k)^\top d \geq 0, \tag{4c}$$

$$0 \leq w_1^k + d_1 \leq yM, \tag{4d}$$

$$0 \leq w_2^k + d_2 \leq (e - y)M, \tag{4e}$$

$$\|d\|_\infty \leq \rho. \tag{4f}$$

- Trust region $\|d\| \leq \rho$ makes feasible set compact.
- bigM dominated by ρ, just needs to be large enough for feasiblity $M = \max((w_1^k, w_2^k)) + \rho$.
- At feasible points, ρ can be arbitarly small - very tight relaxations.
- BNLP index sets:

$$\mathcal{I}_1(w^k + d) = \{i \mid y_i = 0\},$$
$$\mathcal{I}_2(w^k + d) = \{i \mid y_i = 1\}.$$

Lemma

Let $w^* \in \Omega$ be a S-stationary point of the MPEC (1). For a sufficiently small $\rho > 0$, a global minimizer of the relaxed MILP (4) is $d = 0$, and any $y \in \{0,1\}^m$ such that (4d) and (4e) hold.

Phase II - finding feasible points

Crossover strategy

1. Use regularization or penalty based method with σ^k.
2. If $\|\mathrm{diag}(w_1)w_2\|_\infty < \rho^0$, solve LPEC($w^*(\sigma^k), \rho^0$).
3. Solve BNLP($w^*(\sigma^k) + d$), if successful, feasible point found.
4. If not, reduce σ^k and go to 1.

Phase II - finding feasible points

Crossover strategy

1. Use regularization or penalty based method with σ^k.
2. If $\|\mathrm{diag}(w_1)w_2\|_\infty < \rho^0$, solve LPEC$(w^*(\sigma^k), \rho^0)$.
3. Solve BNLP$(w^*(\sigma^k) + d)$, if successful, feasible point found.
4. If not, reduce σ^k and go to 1.

Theorem (Feasibility (informal), N. & Leyffer, 2025)

Under suitable technical assumptions, if $w(\sigma^k)$ is close enough to a feasible point of the MPEC, then for a sufficently large ρ, every feasible point d of LPEC$(w^*(\sigma^k), \rho^0)$ predicts a feasible BNLP.

Phase II - finding feasible points

Crossover strategy

1. Use regularization or penalty based method with σ^k.
2. If $\|\mathrm{diag}(w_1)w_2\|_\infty < \rho^0$, solve LPEC($w^*(\sigma^k), \rho^0$).
3. Solve BNLP($w^*(\sigma^k) + d$), if successful, feasible point found.
4. If not, reduce σ^k and go to 1.

Theorem (Feasibility (informal), N. & Leyffer, 2025)

Under suitable technical assumptions, if $w(\sigma^k)$ is close enough to a feasible point of the MPEC, then for a suffcently large ρ, every feasible point d of LPEC($w^*(\sigma^k), \rho^0$) predicts a feasible BNLP.

Theorem (Convergence (informal), N. & Leyffer, 2025)

Under suitable technical assumptions, given a feasible point $w^0 \in \Omega$, the MPECopt algorithm finds a B-stationary point in a finite number of iterations.

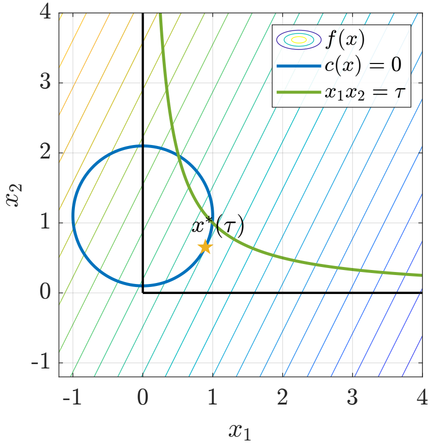
Example: identifying a feasible BNLP

Combination of LPECs with regularization methods.

Consider the two-dimensional MPEC:

$$\begin{aligned} \min_{w \in \mathbb{R}^2} \quad & -2w_1 + w_2 \\ \text{s.t.} \quad & -w_1 - (w_2 - a)^2 + 1 \geq 0 \\ & 0 \leq w_1 \perp w_2 \geq 0 \end{aligned}$$

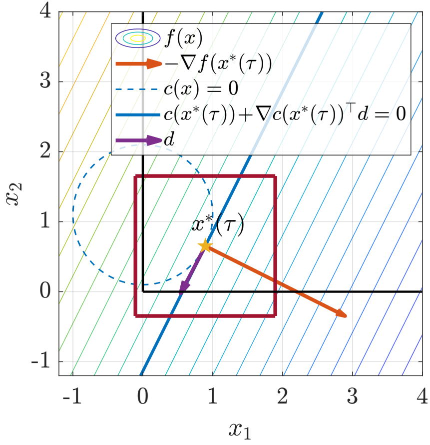
- In the example, we set $a = 1.1$.



Consider the two-dimensional MPEC:

$$\begin{aligned} \min_{w \in \mathbb{R}^2} \quad & -2w_1 + w_2 \\ \text{s.t.} \quad & -w_1 - (w_2 - a)^2 + 1 \geq 0 \\ & 0 \leq w_1 \perp w_2 \geq 0 \end{aligned}$$

- In the example, we set $a = 1.1$.
- If σ is not small enough, LPEC selects an infeasible BNLP.



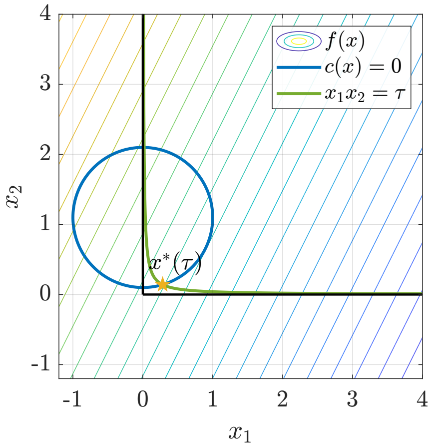
Example: identifying a feasible BNLP

Combination of LPECs with regularization methods.

Consider the two-dimensional MPEC:

$$\begin{aligned}\min_{w\in\mathbb{R}^2} \quad & -2w_1 + w_2 \\ \text{s.t.} \quad & -w_1 - (w_2 - a)^2 + 1 \geq 0 \\ & 0 \leq w_1 \perp w_2 \geq 0\end{aligned}$$

- In the example, we set $a = 1.1$.
- If σ is not small enough, LPEC selects an infeasible BNLP.
- For smaller τ LPEC predicts correct BNLP.



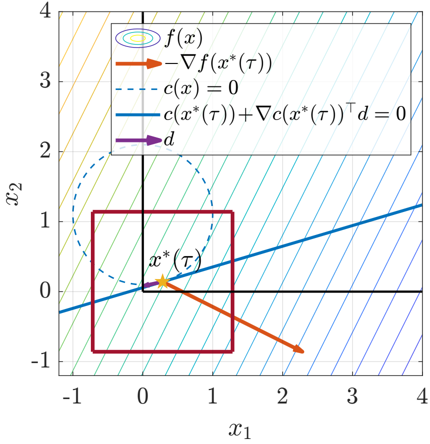
Example: identifying a feasible BNLP

Combination of LPECs with regularization methods.

Consider the two-dimensional MPEC:

$$\begin{aligned} \min_{w \in \mathbb{R}^2} \quad & -2w_1 + w_2 \\ \text{s.t.} \quad & -w_1 - (w_2 - a)^2 + 1 \geq 0 \\ & 0 \leq w_1 \perp w_2 \geq 0 \end{aligned}$$

- In the example, we set $a = 1.1$.
- If σ is not small enough, LPEC selects an infeasible BNLP.
- For smaller τ LPEC predicts correct BNLP.
- In practice, often for large σ the LPEC finds a feasible BNLP.



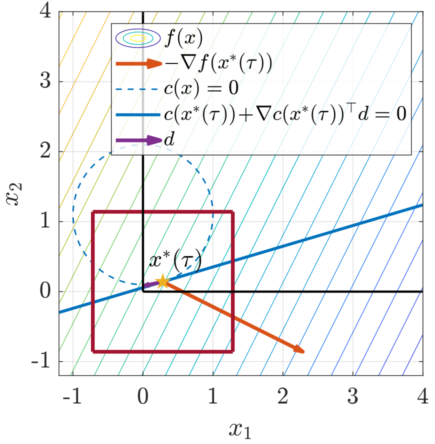
Example: identifying a feasible BNLP

Combination of LPECs with regularization methods.

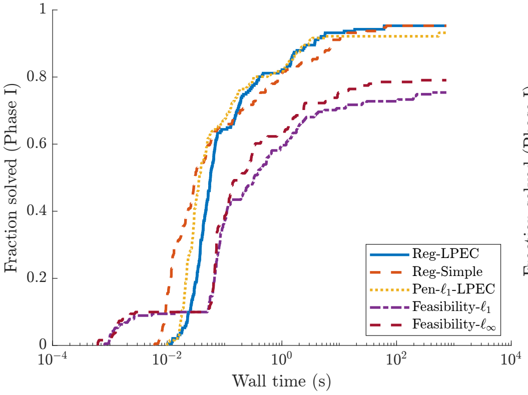
Consider the two-dimensional MPEC:

$$\begin{aligned} \min_{w \in \mathbb{R}^2} \quad & -2w_1 + w_2 \\ \text{s.t.} \quad & -w_1 - (w_2 - a)^2 + 1 \geq 0 \\ & 0 \leq w_1 \perp w_2 \geq 0 \end{aligned}$$

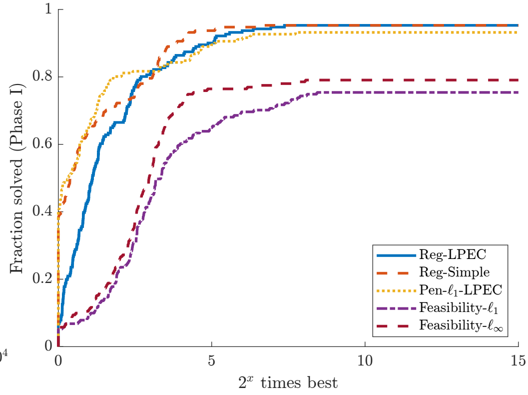
- In the example, we set $a = 1.1$.
- If σ is not small enough, LPEC selects an infeasible BNLP.
- For smaller τ LPEC predicts correct BNLP.
- In practice, often for large σ the LPEC finds a feasible BNLP.
- Moreover, often the solution of this BNLP coincides with the solution of the MPEC.



Outline of the lecture



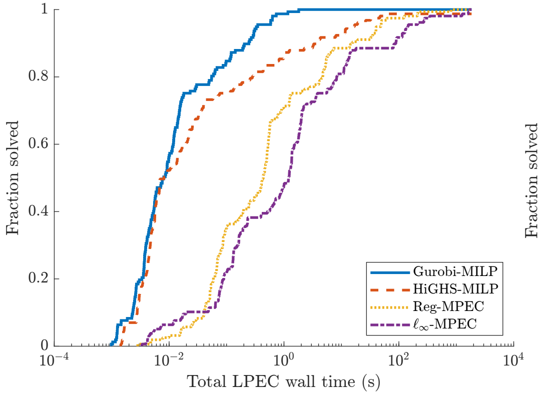
(a) Absolute timings.



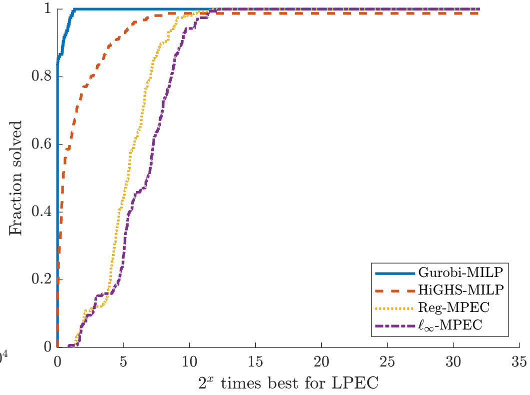
(b) Relative timings.

Figure: Evaluating different Phase I algorithms in MPECopt on the `MacMPEC` test set.

Comparison of LPEC methods on macMPEC (1/3)



(a) Absolute timings.



(b) Relative timings.

Figure: Evaluating different LPEC algorithms in MPECopt on the `MacMPEC` test set.

Comparison of LPEC methods on macMPEC (2/3)

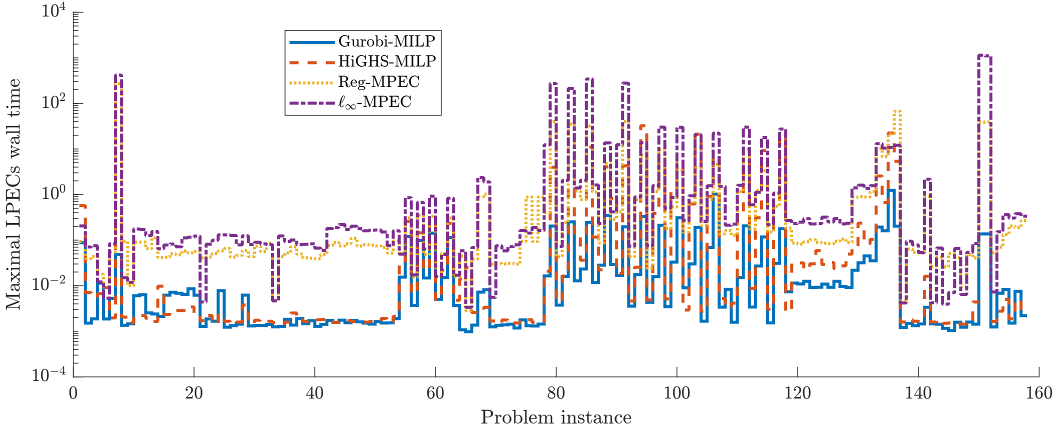


Figure: Maximal solution times for different LPEC algorithms in MPECopt on the `MacMPEC` test set.

Comparison of LPEC methods on macMPEC (3/3)

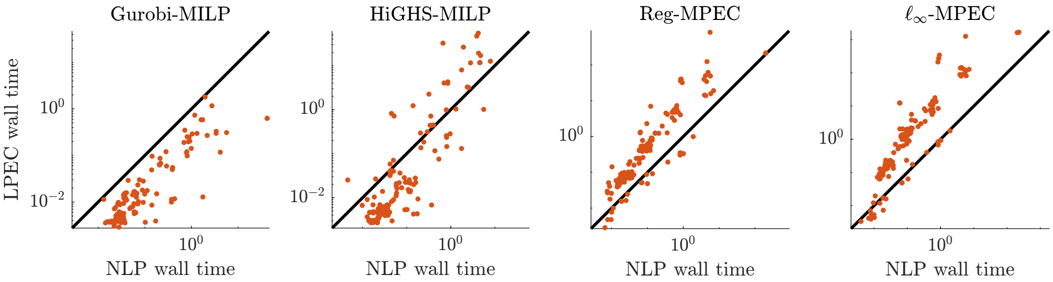
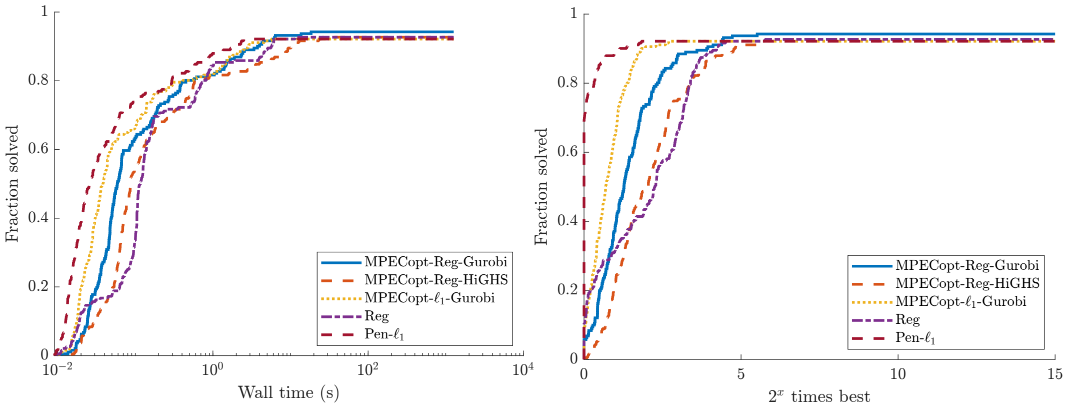


Figure: Comparison of total NLP and LPEC computation times on the MacMPEC.

Active-set vs regularization methods on macMPEC



(a) Absolute timings.

(b) Relative timings.

Figure: Evaluating different MPEC solution methods on the MacMPEC test set in terms of finding a stationary point.

Number of NLP and LPEC solves

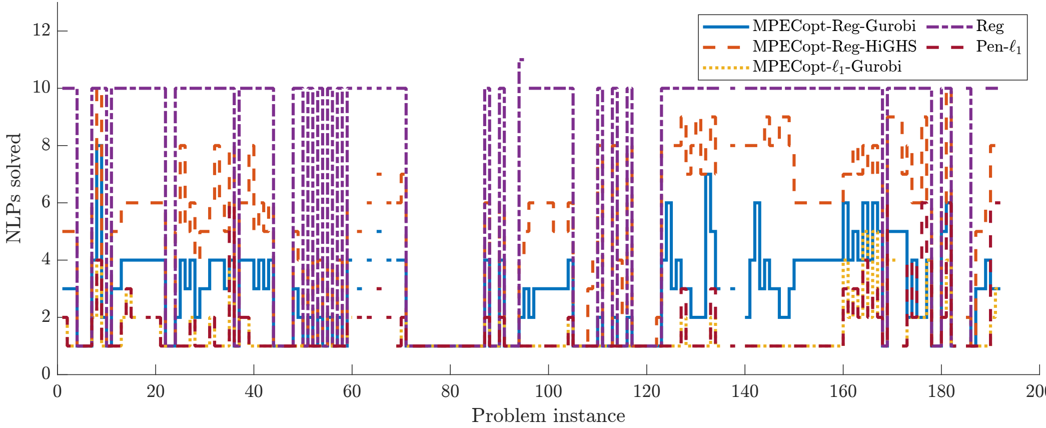


Figure: Number of NLPs (top plot) and NLPs (bottom plots) solved in MacMPEC on all problem instances.

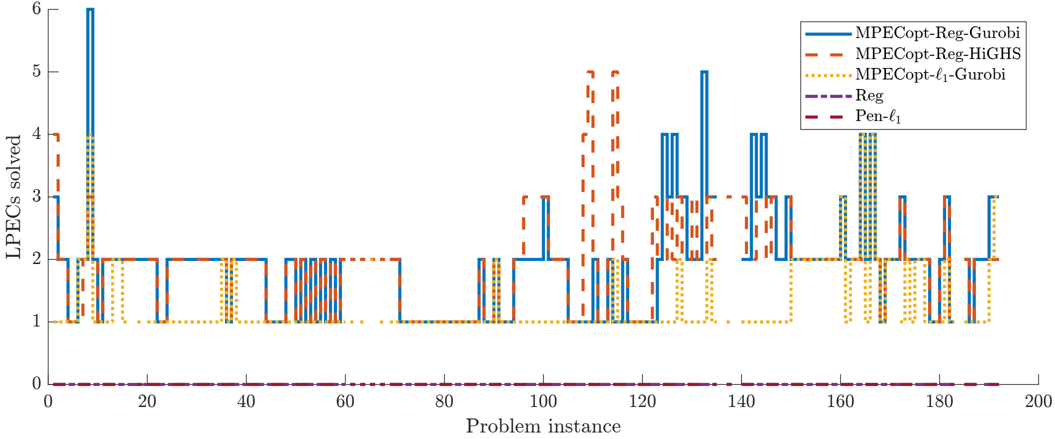


Figure: Number of NLPs (top plot) and NLPs (bottom plots) solved in MacMPEC on all problem instances.

Distribution of stationary points

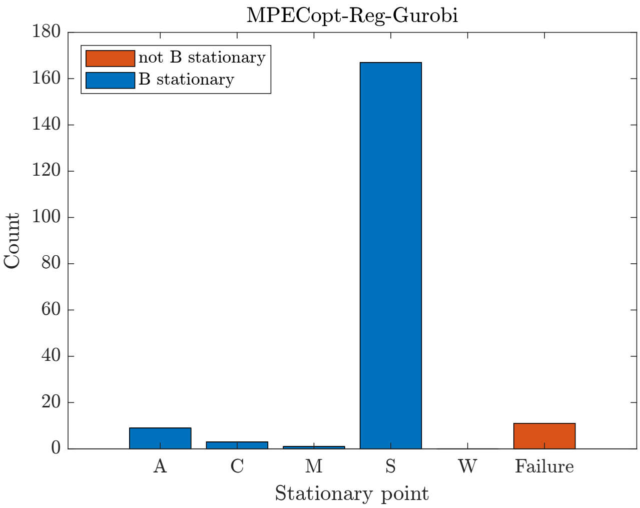


Figure: Distribution of stationary points on `MacMPEC` for different solution methods. Failure counts the number of infeasible problems.

Distribution of stationary points

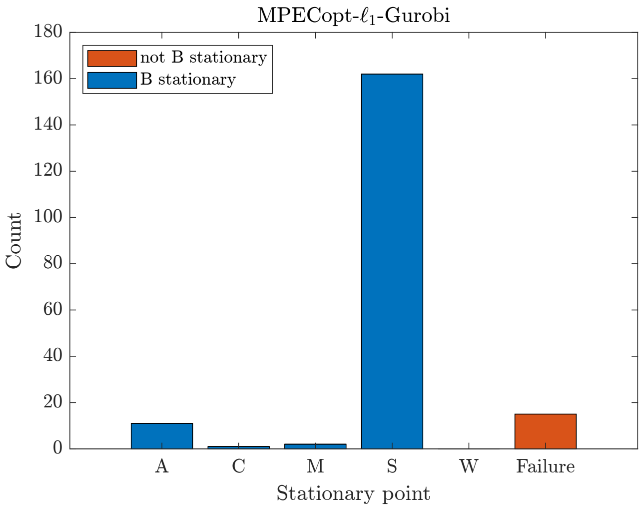


Figure: Distribution of stationary points on MacMPEC for different solution methods. Failure counts the number of infeasible problems.

Distribution of stationary points

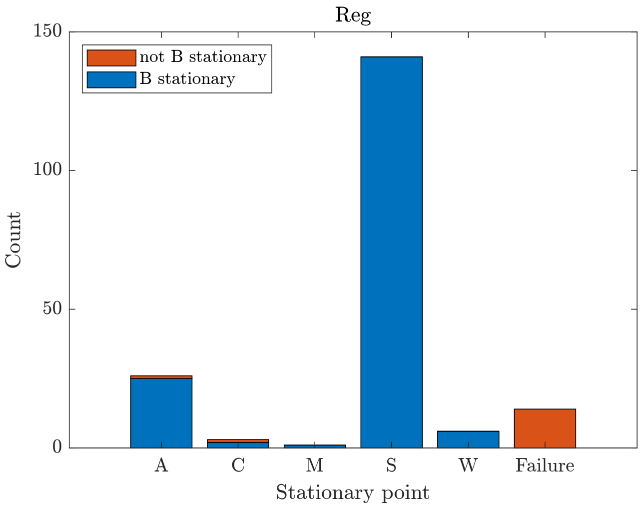


Figure: Distribution of stationary points on MacMPEC for different solution methods. Failure counts the number of infeasible problems.

Distribution of stationary points

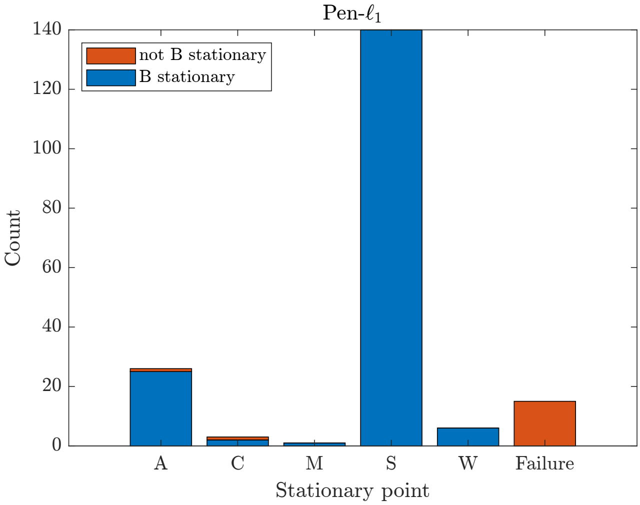
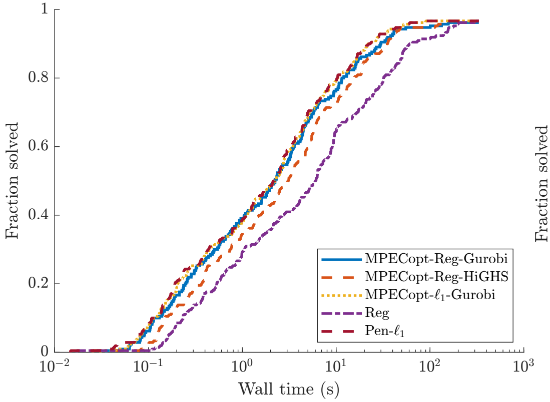
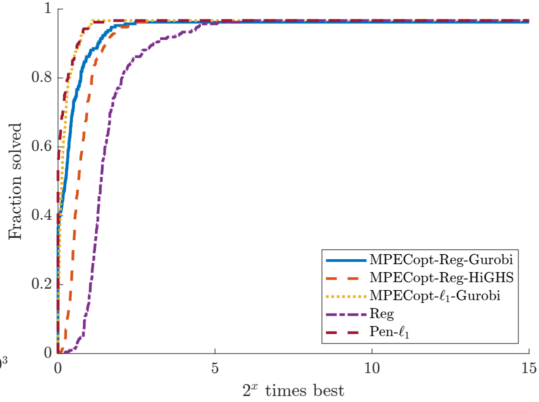


Figure: Distribution of stationary points on `MacMPEC` for different solution methods. Failure counts the number of infeasible problems.

Active set vs regularization methods on random MPECs



(a) Absolute timings.



(b) Relative timings.

Figure: Evaluating different MPEC solution methods on the synthetic test set in terms of finding a stationary point.

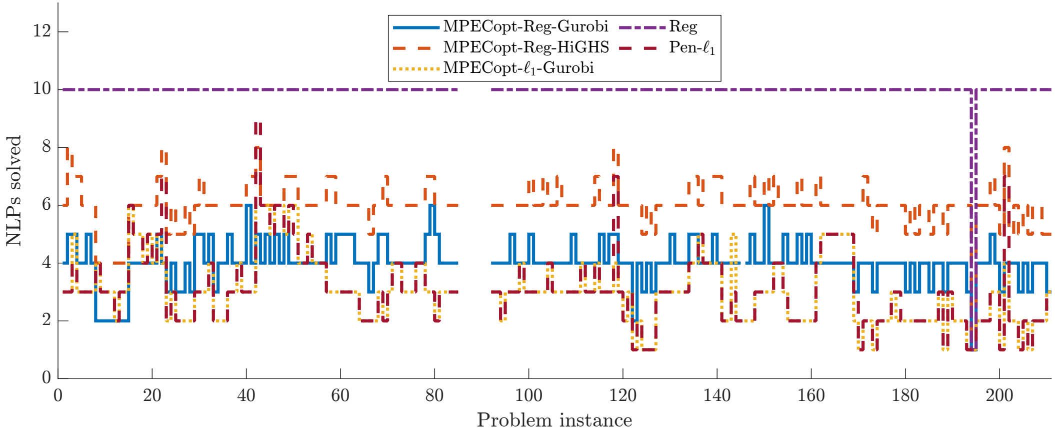


Figure: Number of NLPs (top plot) and NLPs (bottom plots) solved in the synthetic test set on all problem instances.

Number of NLP and LPEC solves

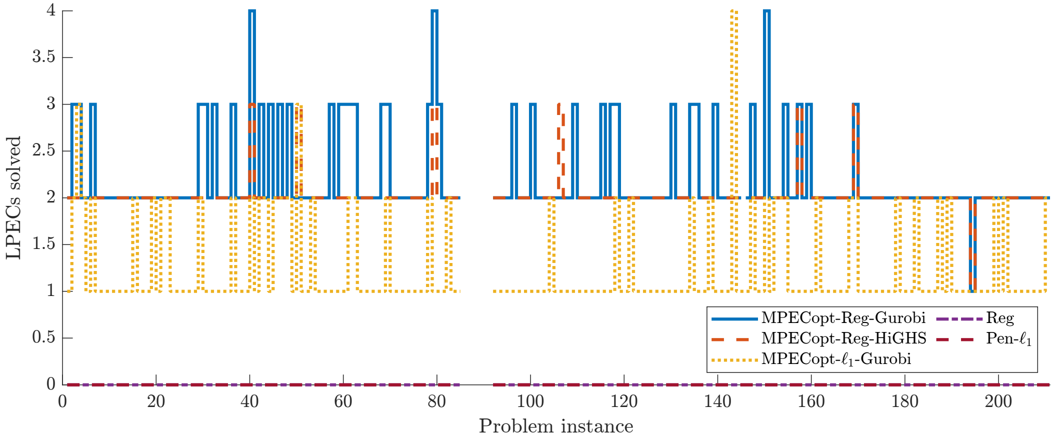


Figure: Number of NLPs (top plot) and NLPs (bottom plots) solved in the synthetic test set on all problem instances.

Distribution of stationary points

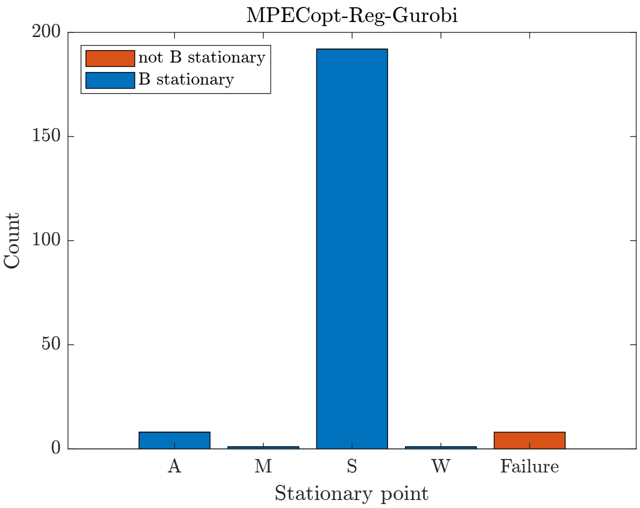


Figure: Distribution of stationary points on the synthetic test set for different solution methods. Failure counts the number of infeasible problems.

Distribution of stationary points

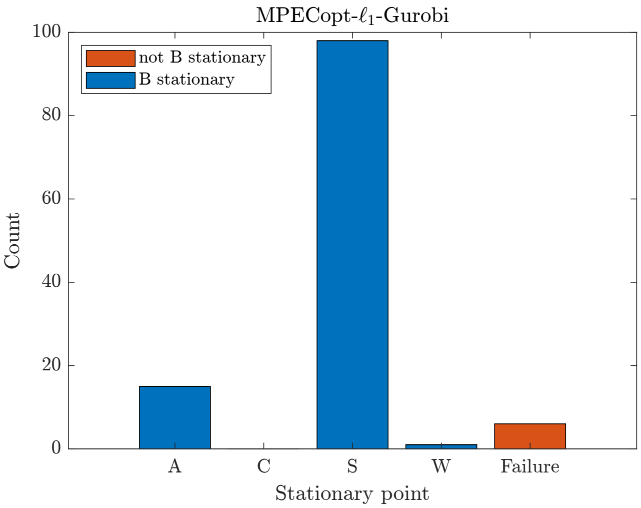


Figure: Distribution of stationary points on the synthetic test set for different solution methods. Failure counts the number of infeasible problems.

Distribution of stationary points

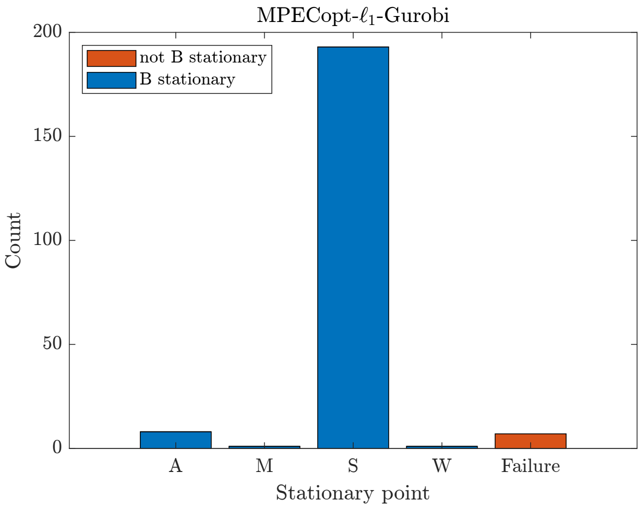


Figure: Distribution of stationary points on the synthetic test set for different solution methods. Failure counts the number of infeasible problems.

Distribution of stationary points

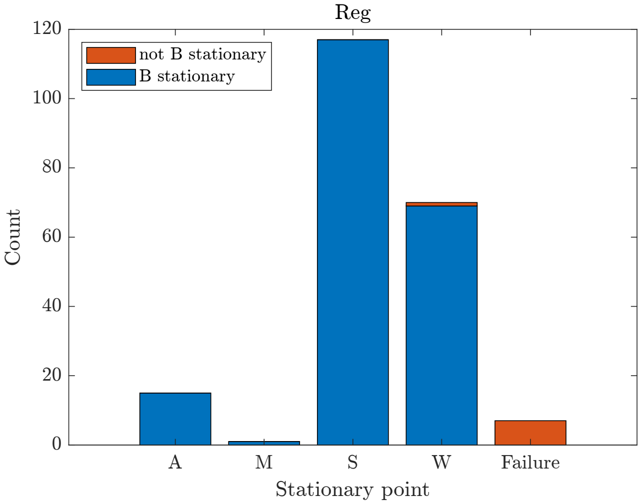


Figure: Distribution of stationary points on the synthetic test set for different solution methods. Failure counts the number of infeasible problems.

Summary

- MPEC methods solve (approximately) a sequence of more regular NLP.
- Regularization/penalty methods: easy to implement, but may converge to spurious stationary points.
- In practice, they do not converge often to spurious stationary points.
- However, if MPEC-LICQ does not hold we can only know if we solve an LPEC.
- Active-set methods solve a sequence of branch NLPs. Converge always to B-stationary points.
- LPECs of reasonable size can be efficiently solved as MILPs, because small ρ means not so much branching.

Thank you for your attention!

For more info on our work see summer school course material.

Cited references I

Anitescu, M. (2005a).
Global convergence of an elastic mode approach for a class of mathematical programs with complementarity constraints.
SIAM Journal on Optimization, 16:120–145.

Anitescu, M. (2005b).
On using the elastic mode in nonlinear programming approaches to mathematical programs with complementarity constraints.
SIAM Journal on Optimization, 15(4):1203–1236.

Fukushima, M. and Tseng, P. (2002).
An implementable active-set algorithm for computing a b-stationary point of a mathematical program with linear complementarity constraints.
SIAM Journal on Optimization, 12(3):724–739.

Cited references II

Giallombardo, G. and Ralph, D. (2008).
Multiplier convergence in trust-region methods with application to convergence of decomposition methods for mpecs.
Mathematical Programming, 112(2):335–369.

Guo, L. and Deng, Z. (2022).
A new augmented lagrangian method for mpccs—theoretical and numerical comparison with existing augmented lagrangian methods.
Mathematics of Operations Research, 47(2):1229–1246.

Hoheisel, T., Kanzow, C., and Schwartz, A. (2013).
Theoretical and numerical comparison of relaxation methods for mathematical programs with complementarity constraints.
Mathematical Programming, 137(1):257–288.

Cited references III

Izmailov, A. F. and Solodov, M. V. (2008).
An active-set newton method for mathematical programs with complementarity constraints.
SIAM Journal on Optimization, 19(3):1003–1027.

Jiang, H. and Ralph, D. (1999).
Qpecgen, a matlab generator for mathematical programs with quadratic objectives and affine variational inequality constraints.
Computational Optimization and Applications, 13:25–59.

Kadrani, A., Dussault, J.-P., and Benchakroun, A. (2009).
A new regularization scheme for mathematical programs with complementarity constraints.
SIAM Journal on Optimization, 20(1):78–103.

Cited references IV

Kanzow, C. and Schwartz, A. (2013).
A new regularization method for mathematical programs with complementarity constraints with strong convergence properties.
SIAM Journal on Optimization, 23(2):770–798.

Kanzow, C. and Schwartz, A. (2015).
The price of inexactness: convergence properties of relaxation methods for mathematical programs with complementarity constraints revisited.
Mathematics of Operations Research, 40(2):253–275.

Kazi, S. R., Thombre, M., and Biegler, L. (2024).
Globally convergent method for optimal control of hybrid dynamical systems.
IFAC-PapersOnLine, 58(14):868–873.

Kirches, C., Larson, J., Leyffer, S., and Manns, P. (2022).
Sequential linearization method for bound-constrained mathematical programs with complementarity constraints.
SIAM Journal on Optimization, 32(1):75–99.

Cited references V

Leyffer, S., López-Calva, G., and Nocedal, J. (2006).
Interior methods for mathematical programs with complementarity constraints.
SIAM Journal on Optimization, 17(1):52–77.

Leyffer, S. and Munson, T. S. (2007).
A globally convergent filter method for mpecs.
Preprint ANL/MCS-P1457-0907, Argonne National Laboratory, Mathematics and Computer Science Division.

Lin, G.-H. and Fukushima, M. (2005).
A modified relaxation scheme for mathematical programs with complementarity constraints.
Annals of Operations Research, 133(1):63–84.

Lin, G.-H. and Fukushima, M. (2006).
Hybrid approach with active set identification for mathematical programs with complementarity constraints.
Journal of optimization theory and applications, 128(1):1–28.

Cited references VI

Liu, X., Perakis, G., and Sun, J. (2006).
A robust sqp method for mathematical programs with linear complementarity constraints.
Computational Optimization and Applications, 34(1):5–33.

Luo, Z., Pang, J., and Ralph, D. (1996).
Mathematical Programs with Equilibrium Constraints.
Cambridge University Press, Cambridge.

Nurkanović, A., Pozharskiy, A., and Diehl, M. (2024).
Solving mathematical programs with complementarity constraints arising in nonsmooth optimal control.
Vietnam Journal of Mathematics, pages 1–39.

Raghunathan, A. U. and Biegler, L. T. (2005).
An interior point method for mathematical programs with complementarity constraints (mpccs).
SIAM Journal on Optimization, 15(3):720–750.

Cited references VII

Ralph, D. and Wright, S. J. (2004).
Some properties of regularization and penalization schemes for mpecs.
Optimization Methods and Software, 19(5):527–556.

Scholtes, S. (2001).
Convergence properties of a regularization scheme for mathematical programs with complementarity constraints.
SIAM Journal on Optimization, 11(4):918–936.

Scholtes, S. and Stöhr, M. (1999).
Exact penalization of mathematical programs with equilibrium constraints.
SIAM Journal on Control and Optimization, 37(2):617–652.

Steffensen, S. and Ulbrich, M. (2010).
A new relaxation scheme for mathematical programs with equilibrium constraints.
SIAM Journal on Optimization, 20(5):2504–2539.