

Reachability Bounds for Global and Robust Optimization of Parametric ODEs Via Implicit Methods

Matthew Wilhelm, PhD Candidate Matthew Stuber, Assistant Professor





Dynamic Optimization

One common form of the dynamic optimization problem

$$\phi^* = \min_{\mathbf{p} \in P \subset \mathbb{R}^{n_p}} \phi(\mathbf{x}(\mathbf{p}, t_f), \mathbf{p})$$

s.t. $\dot{\mathbf{x}} = \mathbf{f}(\mathbf{x}(\mathbf{p}, t), \mathbf{p}, t), \quad \forall t \in I = [t_0, t_f]$
 $\mathbf{x}(\mathbf{p}, t) = \mathbf{x}_0(\mathbf{p})$
 $\mathbf{g}(\mathbf{x}(\mathbf{p}, t), \mathbf{p}) \leq \mathbf{0}, \quad \forall t \in I = [t_0, t_f]$

- That is we seek an optimal satisfying the ODE-IVP ۲ problem along with any constraints.
 - \circ **x** state variables
 - **p** decision variables Ο

Global optimization of stiff dynamical systems

Matthew E. Wilhelm¹ Anne V. Le² Matthew D. Stuber¹

Abstract

Department of Chemical Engineering, Texas A&M University, College Station, Texas Correspondence Matthew D. Stuber, PhD, Process Systems and Operations Research Laboratory, University of cticut, Storrs, Mansfield, CT 0626

Process Systems and Operations Research

Connecticut, Storrs, Mansfield, Connecticut

Laboratory, Department of Chemical and Biomolecular Engineering, University of

We present a deterministic global optimization method for nonlinear programming formulations constrained by stiff systems of ordinary differential equation (ODE) ini tial value problems (IVPs). The examples arise from dynamic optimization problems exhibiting both fast and slow transient phenomena commonly encountered in modelbased systems engineering applications. The proposed approach utilizes unconditionally stable implicit integration methods to reformulate the ODE-constrained problem

Wilhelm, ME; Le, AV; and Stuber. MD. "Global Optimization of Stiff Dynamical Systems." In Press for AIChE Journal: Futures Issue (2019)

1 | INTRODUCTION

Dynamic optimization problems of the form $\phi^* = \min_{\mathbf{p} \in \mathbf{P} \cap \mathbf{P}^*} \phi(\mathbf{x}(\mathbf{p}, t_f), \mathbf{p})$ s.t. $\dot{\mathbf{x}}(\mathbf{p},t) = \mathbf{f}(\mathbf{x}(\mathbf{p},t),\mathbf{p},t), \forall t \in I = [t_0,t_f]$

 $x(p,t_0) = x_0(p)$

 $g(x(p,t_f),p) \leq 0$

are of extreme importance to process systems engineers and the broader model-based systems engineering community as they can be However, calculating rigorous lower bounds poses significant chalformulated for a variety of systems whose transient behavior is of lenges as this step requires that rigorous and accurate global bounds particular interest, from optimal control to mechanistic model valida- are known or are readily calculable for all variables and functions of tion. The first major complicating detail of the optimization formula(1). For standard nonconvex nonlinear programs (NLPs) (i.e., without tion (1) is that it is constrained by a system of ordinary differential dynamical systems constraints), rigorous lower bounds on the optimal equation-initial value problems (ODE-IVPs). Therefore, simply verifying a feasible point requires the solution of a system of ODE-IVPs. The second major complicating detail is that (1) is a nonconvex

program, in general, and therefore verifying optimality requires deter ministic global optimization. The focus of this paper is on solving (1) to guaranteed global optimality (or declaration of infeasibility). The methods developed in this work are of specific importance when the ODF-IVP system is stiff

Methods for solving (1) rigorously to global optimality rely on the spatial branch-and-bound (B&B) framework^{1,2} or some variant. The B&B algorithm requires the ability to calculate rigorous upper and lower bounds on the global optimal solution value. An upper bound can be calculated by simply evaluating $\phi(\mathbf{x}(\cdot, t_t), \cdot)$ at any feasible point solution value are obtained by calculating convex and concave relaxa tions of the functions and solving a corresponding convex lowerbounding problem. Applying this approach to a dynamic optimization

AlChE Journal. 2019;e16836. wileyonlinelibrary.com/journal/aic https://doi.org/10.1002/aic.16836

© 2019 American Institute of Chemical Engineers 1 of 20



Parametric ODE-IVP

Parametric ODE-IVP Formulation

$$\dot{\mathbf{x}} = \frac{d\mathbf{x}}{dt}(\mathbf{p}, t) = \mathbf{f}(\mathbf{x}(\mathbf{p}, t), \mathbf{p}, t), \ t \in I = [t_0, t_f], \ \mathbf{p} \in P$$

 $\mathbf{x}(\mathbf{p}, t_0) = \mathbf{x}_0(\mathbf{p}), \ \mathbf{p} \in P$

where $\mathbf{f}: D \times \Pi \times T \to \mathbb{R}^{n_x}$ and $x_0: P \to D$ with $D \subset \mathbb{R}^{n_x}$, $\Pi \subset \mathbb{R}^{n_p}$, and $T \subset \mathbb{R}$ open with $P \in I\Pi$ and $I \in IT$

Assumptions (Well-posed)

1. $\mathbf{x}_0: P \to D$ is locally Lipschitz continuous on P

2. **f** is continuously differential on $D \times \Pi \times T$

Parametric ODE-IVP Solution



A solution is any continuous $\mathbf{x} : P \times I \to D$ such that, for every $\mathbf{p} \in P$, $\mathbf{x}(\mathbf{p}, \cdot) : T \to D$ is continuous differentiable and satisfies the parametric ODE-IVP on I.



Review (Dynamic Optimization)

Discretization¹:

- \circ The domain is separated into *K* finite elements.
- The derivative terms are approximated by difference forms wherever they appear within the formulation.

$$\frac{dx}{dt}\Big|_{t_{k+1}} = \frac{x_{k+1} - x_k}{h}$$
$$g\left(\frac{dx}{dt}\Big|_{t_{k+1}}, f(x_{k+1}, p_{k+1})\right) = 0$$

k = 0, ..., K - 1

Collocation¹:

- The domain is separated into *K* finite elements.
- The solution in each finite element is then approximated by a polynomial of order N + 1

$$\frac{dx}{dt}\Big|_{t_{kj}} = \frac{1}{h_i} \sum_{j=0}^{K} \frac{dl_j(\tau_i)}{d\tau}, \qquad x_{i+1,0} = \sum_{j=0}^{N} l_j(1) x_{kj}$$
$$g\left(\frac{dx}{dt}\Big|_{t_{kj}}, f(x_{ki}, p_{ki})\right) = 0, \qquad t_{kj} = t_{k-1} + \tau_j h$$

 $k = 0, \dots, K - 1, \quad i = 0, \dots, N - 1,$

1. Biegler, L.T.: Nonlinear Programming: Concepts, Algorithms, and Applications to Chemical Processes. SIAM, Philadelphia (2010)

Review (Dynamic Optimization)

Discretization¹:

difference form

 $k = 0, \dots, K - 1$

- \circ The domain is separated into *K* finite elements.
- The derivative terms are approximated by

Collocation¹:

- The domain is separated into K finite elements.
 The colution in each finite element is then
 - al of order N + 1
- the formulatic Either method introduces a significant number dx| of variables and potentially nonlinear equations
 - Results in complicated high dimension problems that may be challenging for modern global optimizers to solve.

 $r_{i+1,0} = \sum_{j=0}^{N} l_j(1) x_{kj}$

 $t_{kj} = t_{k-1} + \tau_j h$



Review (Relaxing ODEs)

Taylor-Series Integrators³

$$\mathbf{x}(\tau_{q+1},\mathbf{p}) \in \underbrace{\mathbf{x}(\tau_{q},\mathbf{p}) + \sum_{j=1}^{p} \frac{h^{j}}{j!} \mathbf{f}^{(j)}(\mathbf{x}(\tau_{q},\mathbf{p}),\mathbf{p})}_{\text{Taylor Series}} + \underbrace{\frac{h^{p+1}}{(p+1)!} \mathbf{f}^{(p+1)}(\mathbf{X}(\tau_{q}),\mathbf{P})}_{\text{Remainder Bound}}$$

Differential Inequalities⁴

$$\dot{\mathbf{x}}^{cv}(t,\mathbf{p}) = \mathbf{f}^{cv}(t,\mathbf{p},\mathbf{x}^{cv}(t,\mathbf{p}),\mathbf{x}^{cc}(t,\mathbf{p})), \quad \mathbf{x}^{cv}(t_0,\mathbf{p}) = \mathbf{x}_0^{cv}(\mathbf{p})$$
$$\dot{\mathbf{x}}^{cc}(t,\mathbf{p}) = \mathbf{f}^{cc}(t,\mathbf{p},\mathbf{x}^{cv}(t,\mathbf{p}),\mathbf{x}^{cc}(t,\mathbf{p})), \quad \mathbf{x}^{cc}(t_0,\mathbf{p}) = \mathbf{x}_0^{cc}(\mathbf{p})$$

Discretize-and-Relax⁵

- A two-stage method.
- Calculates valid relaxations across a time step and then refines relaxation at each point.

- Each of these approaches allows relaxation of ODE to be calculated with respect to p.
- □ Results in a lower dimensional problem.

3. Rihm, Robert. Interval methods for initial value problems in ODEs. Topics in Validated Computations (1994): 173-207.

Joseph K Scott, Paul I Barton. Improved relaxations for the parametric solutions of ODEs using differential inequalities. *Journal of Global Optimization*. 2013 (57): 143–176.
 A.M. Sahlodin, Benoît Chachaut. Discretize-then-relax approach for convex/concave relaxations of the solutions of parametric ODEs. *Applied Numerical Mathematics*, 61 (179): 803 – 820, 2011

Implicit Integration of Problem



7

Implicit Linear Multistep Method

• An s-step PILMS method can be defined by the equation^{6,7}:

$$\widehat{\mathbf{z}}_{k+s} + \sum_{i=0}^{s-1} a_i \widehat{\mathbf{z}}_{k+i} - \Delta t \sum_{j=0}^s b_j \mathbf{f}(\widehat{\mathbf{z}}_{k+j}, \mathbf{p}, t_{k+j}) = \mathbf{0}$$

$$\mathbf{h}_{1} = \boldsymbol{\xi}_{0}^{1} = (h_{1}, h_{2}, h_{3}, h_{4}, h_{5})$$

$$\mathbf{h}_{2} = \boldsymbol{\theta}_{0}^{2} = (h_{6}, h_{7}, h_{8}, h_{9}, h_{10})$$

$$\mathbf{h}_{3} = \boldsymbol{\theta}_{1}^{2} = (h_{11}, h_{12}, h_{13}, h_{14}, h_{15})$$

$$\vdots \qquad \vdots \qquad \vdots \qquad \vdots \qquad \\ \mathbf{h}_{K} = \boldsymbol{\theta}_{K-2}^{2} = (h_{5K-4}, h_{5K-3}, h_{5K-2}, h_{5K-2}, h_{5K})$$

$$\mathbf{z}_{0} \mathbf{p} \quad \overbrace{\boldsymbol{\xi}_{0}^{1}}^{\mathbf{z}_{1}} \qquad \overbrace{\boldsymbol{\xi}_{0}^{2}}^{\mathbf{z}_{2}} \qquad \overbrace{\boldsymbol{\theta}_{0}^{2}}^{\mathbf{z}_{2}} \qquad \overbrace{\boldsymbol{\theta}_{1}^{2}}^{\mathbf{z}_{3}} \qquad \overbrace{\boldsymbol{\theta}_{2}^{2}}^{\mathbf{z}_{4}} \qquad \overbrace{\boldsymbol{\theta}_{3}^{2}}^{\mathbf{z}_{4}} \cdots$$

- Specific methods for each choice of $\{a_i\}_{i=0}^{s-1}$ and $\{b_j\}_{j=0}^s$.
- Adams Moulton arises from $a_{s-1} = -1$, $\{a_i\}_{i=0}^{s-2} = 0$.

$$\overline{\boldsymbol{\xi}_{k}^{s}(\hat{\mathbf{z}}_{k+s},\ldots,\hat{\mathbf{z}}_{k},\mathbf{p})} = \hat{\mathbf{z}}_{k+s} + \sum_{i=0}^{s-1} a_{i}\hat{\mathbf{z}}_{k+i} - \Delta t b_{j}\mathbf{f}(\hat{\mathbf{z}}_{k+s},\mathbf{p},t_{k+s}) = \mathbf{0}$$

• Backwards Difference Formula comes from

$$\boldsymbol{\zeta}_{k}^{s}(\hat{\mathbf{z}}_{k+s},\ldots,\hat{\mathbf{z}}_{k},\mathbf{p}) = \hat{\mathbf{z}}_{k+s} - \hat{\mathbf{z}}_{k+s-1} - \Delta t \sum_{i=0}^{s} b_{j}\mathbf{f}(\hat{\mathbf{z}}_{k+j},\mathbf{p},t_{k+j}) = \mathbf{0}$$

6. Gautschi W. Numerical Analysis. Springer Science & Business Media, New York; 2012.

7. Hairer E, Wanner G. Solving Ordinary Differential Equations II: Stiff and Differential-Algebraic Problems. Springer, Heidelberg; 1991.



Two Step PILMS Methods

- Two-step PILMS methods exhibit unconditional A-stability^{6,7}
- The three implicit approaches considered can be expressed in terms of a series of block solves given by:

Implicit Euler^{6,7}*:*

$$\boldsymbol{\xi}_{k}^{1}(\hat{\mathbf{z}}_{k+1}, \hat{\mathbf{z}}_{k}, \mathbf{p}) = \hat{\mathbf{z}}_{k+1} - \hat{\mathbf{z}}_{k} - \Delta t \mathbf{f}(\hat{\mathbf{z}}_{k+1}, \mathbf{p}, t_{k+1})$$

Two-step Adam's Moulton method^{6,7}:

$$\boldsymbol{\xi}_{k}^{2}(\hat{\mathbf{z}}_{k+2}, \hat{\mathbf{z}}_{k+1}, \hat{\mathbf{z}}_{k}, \mathbf{p}) = \hat{\mathbf{z}}_{k+2} - \frac{4}{3}\hat{\mathbf{z}}_{k+1} + \frac{1}{3}\hat{\mathbf{z}}_{k} - \frac{2}{3}\Delta t \mathbf{f}(\hat{\mathbf{z}}_{k+2}, \mathbf{p}, t_{k+2}, \mathbf{p})$$

Two-step BDF method^{6,7}*:*

$$\boldsymbol{\zeta}_{k}^{2}(\hat{\boldsymbol{z}}_{k+2}, \hat{\boldsymbol{z}}_{k+1}, \hat{\boldsymbol{z}}_{k}, \boldsymbol{p}) = \hat{\boldsymbol{z}}_{k+2} - \hat{\boldsymbol{z}}_{k+1} - \frac{1}{2}\Delta t \big(\mathbf{f}(\hat{\boldsymbol{z}}_{k+2}, \boldsymbol{p}, t_{k+2}) + \mathbf{f}(\hat{\boldsymbol{z}}_{k+1}, \boldsymbol{p}, t_{k+1}) \big)$$

7. Hairer E, Wanner G. Solving Ordinary Differential Equations II: Stiff and Differential-Algebraic Problems. Springer, Heidelberg; 1991.



^{6.} Gautschi W. Numerical Analysis. Springer Science & Business Media, New York; 2012.

Assumptions

- 1. There exists a unique function $\mathbf{z}: P \to D^{K+1}$ with $\mathbf{h}(\mathbf{z}(\mathbf{p}), \mathbf{p}) = \mathbf{0}$, and an interval $X \in ID$ such that $\mathbf{z}(\mathbf{p})$ is unique in X^{K+1} for all $\mathbf{p} \in P$.
- 2. Derivative information ∇h_i , $i = 1, ..., n_x K$ is available, say by automatic differentiation, and is factorable.
- 3. A matrix $\mathbf{Y}_k \in \mathbb{R}^{n_x \times n_x}$ is known such that $M_k = \mathbf{Y}_k J_k^s (X, P)$ satisfies $0 \notin M_{k,ii}$ for all $i \in \{1, ..., n_x\}$ and for all k, where $J_k^s (X, P)$ is an inclusion monotonic interval extension of \mathbf{J}_k^s on $X \times P$.

Hardest assumption to verify:

- We've developed block sequential analogs to the sufficient conditions presented in Neumaier.⁸
- Parametric interval methods can be used to provide sharper tests.^{8,9}
- These can further be combined with bisection methods.^{10,11}

Can the interval extension of J_k^s be preconditioned such that it contains no singular matrices on its domain?





Development of Relaxations

• The parametric mean-valued form¹² of a block is given by:

 $\left|\nabla_{\mathbf{x}}h_{j}(\mathbf{y}^{j}(\mathbf{p}),\mathbf{p})^{T}(\mathbf{z}_{k}(\mathbf{p})-\boldsymbol{\gamma}(\mathbf{p}))=-h_{j}(\boldsymbol{\gamma}(\mathbf{p}),\mathbf{p}), \quad j=(k-1)n_{x}+1,...,kn_{x}\right|$

where $\gamma: P \to D$ and $\mathbf{y}^j: P \to D$ such that for some $\lambda: P \to (0,1)$,

 $\mathbf{y}^{j}(\mathbf{p}) = \lambda(\mathbf{p})\mathbf{z}_{k}(\mathbf{p}) + (1 - \lambda(\mathbf{p}))\gamma(\mathbf{p}), \quad \forall \mathbf{p} \in P$

• This suggests that a parametric analog of the Newton-Raphson fixed point method may be used to compute relaxation.



1. Initialize the relaxations, \mathbf{z}_k^0 , and subgradients, $\mathbf{s}_{\mathbf{z}_k}^0$, from interval bounds.



1: procedure $BLOCKRELAX(\mathbf{h}_k, X, P, \mathbf{z}_{k-1}^{cv}, \mathbf{z}_{k-1}^{cc}, \mathbf{s}_{\mathbf{z}_{k-1}}^{cv}, \mathbf{z}_{k-2}^{cc}, \mathbf{z}_{k-2}^{cv}, \mathbf{z}_{k-2}^{cc}, \mathbf{s}_{\mathbf{z}_{k-2}}^{cv}, \mathbf{s}_{\mathbf{z}_{k-2}}^{cc}, \mathbf{s}_{\mathbf{z}_{k-2}}^{cv}, \mathbf{s}_{\mathbf{z}_{k-2}}^{cc}, \mathbf{s}_{\mathbf{z}_{k-2}}^{cv}, \mathbf$ $\mathbf{z}_{k}^{0,cv}, \mathbf{z}_{k}^{0,cc}, \mathbf{s}_{\mathbf{z}_{k}}^{0,cv}, \mathbf{s}_{\mathbf{z}_{k}}^{0,cc} \leftarrow \mathbf{x}^{L}, \mathbf{x}^{U}, \mathbf{0}, \mathbf{0}$ for $j \leftarrow 0$ to r - 1 do $\lambda \leftarrow \lambda \in [0, 1], \bar{\mathbf{p}} \leftarrow \mathbf{p} \in P$ $c, C, s_c, s_c \leftarrow Aff(\ldots)$ Subroutine defined in [23] $\mathbf{z}_{L}^{i,a}(\mathbf{p}) \leftarrow \mathbf{c} + \mathbf{s}_{\mathbf{c}}^{\mathsf{T}}(\mathbf{p} - \bar{\mathbf{p}}), \forall \mathbf{p} \in P$ Affine relaxation lower bound $\mathbf{z}_{L}^{j,A}(\mathbf{p}) \leftarrow \mathbf{C} + \mathbf{s}_{\mathbf{C}}^{\mathsf{T}}(\mathbf{p} - \bar{\mathbf{p}}), \forall \mathbf{p} \in P$ Affine relaxation upper bound $\gamma^{j}(\cdot) \leftarrow \lambda \mathbf{z}_{k}^{j,a}(\cdot) + (1-\lambda)\mathbf{z}_{k}^{j,A}(\cdot)$ $\mathbf{s}_{\gamma}^{j} \leftarrow \lambda \mathbf{s}_{c} + (1 - \lambda) \mathbf{s}_{C}$ $\mathsf{M}^{j,cv}(\cdot) \leftarrow \mathsf{u}_{\mathsf{B}}(\mathsf{z}^{j,a}_{\iota}(\cdot),\mathsf{z}^{j,A}_{\iota}(\cdot),\ldots,\mathsf{z}^{j,a}_{\iota}(\cdot),\mathsf{z}^{j,A}_{\iota}(\cdot),\cdot)$ B matrix defined in [23] w.r.t. h_k and J_k $\mathbf{M}^{j,cc}(\cdot) \leftarrow \mathbf{o}_{\mathbf{B}}(\mathbf{z}_{k}^{j,\vartheta}(\cdot), \mathbf{z}_{k}^{j,A}(\cdot), \dots, \dots, \mathbf{z}_{k}^{j,\vartheta}(\cdot), \mathbf{z}_{k}^{j,A}(\cdot), \cdot)$ 11: $\mathbf{s}_{\mathbf{M}}^{j,cv}(\cdot) \leftarrow \mathcal{S}_{\mathbf{U}_{\mathcal{B}}}(\mathbf{z}_{\mathbf{k}}^{j,a}(\cdot),\mathbf{z}_{\mathbf{k}}^{j,A}(\cdot),\mathbf{s}_{\mathbf{c}},\mathbf{s}_{\mathbf{C}},\ldots,\mathbf{z}_{\mathbf{k}}^{j,a}(\cdot),\mathbf{z}_{\mathbf{k}}^{j,A}(\cdot),\mathbf{s}_{\mathbf{c}},\mathbf{s}_{\mathbf{C}},\cdot)$ 12: $\mathbf{s}_{\mathbf{M}}^{j,cc}(\cdot) \leftarrow \mathcal{S}_{\mathbf{0}_{R}}(\mathbf{z}_{\mathbf{k}}^{j,a}(\cdot), \mathbf{z}_{\mathbf{k}}^{j,A}(\cdot), \mathbf{s}_{\mathbf{c}}, \mathbf{s}_{\mathbf{C}}, \dots, \mathbf{z}_{\mathbf{k}}^{j,a}(\cdot), \mathbf{z}_{\mathbf{k}}^{j,A}(\cdot), \mathbf{s}_{\mathbf{c}}, \mathbf{s}_{\mathbf{C}}, \cdot)$ 13: $\mathbf{z}_{k}^{j+1,cv}(\cdot) \leftarrow \bar{\mathbf{u}}_{\mathbf{w}}(\boldsymbol{\gamma}^{j}(\cdot), \mathbf{\gamma}^{j}(\cdot), \mathbf{M}^{j,cv}(\cdot), \mathbf{M}^{j,cc}(\cdot), \mathbf{z}_{k}^{j,cv}(\cdot), \mathbf{z}_{k}^{j,cc}(\cdot), \mathbf{z}_{k-1}^{cv}(\cdot) \mathbf{z}_{k-2}^{cc}(\cdot), \mathbf{z}_{k-2}^{cv}(\cdot) \mathbf{z}_{k-2}^{cc}(\cdot), \mathbf{z}_{k-2}^{cv}(\cdot), \mathbf{z}_{k-2}$ 15: $\mathbf{z}_{k}^{j+1,cc}(\cdot) \leftarrow \bar{\mathbf{o}}_{\mathbf{w}}(\gamma^{j}(\cdot),\gamma^{j}(\cdot),\mathbf{M}^{j,cv}(\cdot),\mathbf{M}^{j,cc}(\cdot),\mathbf{z}_{k}^{j,cv}(\cdot),\mathbf{z}_{k}^{j,cc}(\cdot),\mathbf{z}_{k-1}^{cv}(\cdot)\mathbf{z}_{k-1}^{cc}(\cdot),\mathbf{z}_{k-2}^{cv}(\cdot)\mathbf{z}_{k-2}^{cc}(\cdot),\cdot)$ $\mathbf{s}_{\mathbf{z}}^{j+1,cv}(\cdot) \leftarrow S_{\hat{\mathbf{u}}_{\mathbf{w}}}(\gamma^{j}(\cdot),\gamma^{j}(\cdot),\mathbf{s}_{\mathbf{y}}^{j},\mathbf{s}_{\mathbf{y}}^{j},\mathbf{M}^{j,cv}(\cdot),\mathbf{M}^{j,cc}(\cdot),\mathbf{z}_{\mathbf{k}}^{j,cv}(\cdot),\mathbf{z}_{\mathbf{k}}^{j,cc}(\cdot),\mathbf{s}_{\mathbf{z}}^{j,cv}(\cdot),\mathbf{s}_$ 16: $\mathbf{s}_{\mathbf{z}}^{j+1,cc}(\cdot) \leftarrow S_{\mathbf{\bar{0}}_{\mathbf{w}}}(\gamma^{j}(\cdot),\gamma^{j}(\cdot),\mathbf{s}_{\mathbf{y}}^{j},\mathbf{s}_{\mathbf{y}}^{j},\mathbf{M}^{j,cv}(\cdot),\mathbf{M}^{j,cc}(\cdot),\mathbf{z}_{\mathbf{z}}^{j,cv}(\cdot),\mathbf{z}_{\mathbf{z}}^{j,cv}(\cdot),\mathbf{s}_$ end for 18:

19: return $\mathbf{z}_{k}^{r,cc}(\cdot), \mathbf{z}_{k}^{r,cc}(\cdot), \mathbf{s}_{\mathbf{z}}^{r,cv}(\cdot), \mathbf{s}_{\mathbf{z}}^{r,cc}(\cdot)$



2. Compute affine function between current relaxations and respective subgradients



1: procedure BLOCKRELAX($\mathbf{h}_k, X, P, \mathbf{z}_{k-1}^{cv}, \mathbf{z}_{k-1}^{cc}, \mathbf{s}_{\mathbf{z}_{k-1}}^{cv}, \mathbf{s}_{\mathbf{z}_{k-1}}^{cc}, \mathbf{z}_{k-2}^{cv}, \mathbf{z}_{k-2}^{cc}, \mathbf{s}_{\mathbf{z}_{k-2}}^{cv}, \mathbf{s}_{\mathbf{z}_{k-2}}^{cc}, \mathbf{s}_{\mathbf{z}_{k-2}}^{cv}, \mathbf{s}_{\mathbf{z}_{k-2}}^{cc}, \mathbf{s}_{\mathbf{z}_{k-2}}^{cv}, \mathbf{s}_{\mathbf{z}_{k-2}}^{cc}, \mathbf{s}_{\mathbf{z}_{k-2}}^{cv}, \mathbf{s}_{\mathbf{z}_{k-2}}^{cc}, \mathbf{s}_{\mathbf{z}_{k-2}}^{cv}, \mathbf{s}_{\mathbf{z}_{k-2}}^$ $\mathbf{z}_k^{0,cv}, \mathbf{z}_k^{0,cc}, \mathbf{s}_{\mathbf{z}_k}^{0,cv}, \mathbf{s}_{\mathbf{z}_k}^{0,cc} \leftarrow \mathbf{x}^L, \mathbf{x}^U, \mathbf{0}, \mathbf{0}$ for $j \leftarrow 0$ to r - 1 do $\lambda \leftarrow \lambda \in [0, 1], \bar{\mathbf{p}} \leftarrow \mathbf{p} \in P$ $c, C, s_c, s_c \leftarrow Aff(\ldots)$ Subroutine defined in [23] 2. $\mathbf{z}_{L}^{i,a}(\mathbf{p}) \leftarrow \mathbf{c} + \mathbf{s}_{\mathbf{c}}^{\mathsf{T}}(\mathbf{p} - \bar{\mathbf{p}}), \forall \mathbf{p} \in P$ Affine relaxation lower bound $\mathbf{z}_{L}^{j,A}(\mathbf{p}) \leftarrow \mathbf{C} + \mathbf{s}_{\mathbf{C}}^{\mathsf{T}}(\mathbf{p} - \bar{\mathbf{p}}), \forall \mathbf{p} \in P$ Affine relaxation upper bound $\gamma^{j}(\cdot) \leftarrow \lambda \mathbf{z}_{k}^{j,a}(\cdot) + (1-\lambda)\mathbf{z}_{k}^{j,A}(\cdot)$ $s_{\gamma}^{j} \leftarrow \lambda s_{c} + (1 - \lambda)s_{C}$ $\mathsf{M}^{j,cv}(\cdot) \leftarrow \mathsf{u}_{\mathsf{B}}(\mathsf{z}^{j,a}_{\scriptscriptstyle L}(\cdot),\mathsf{z}^{j,A}_{\scriptscriptstyle L}(\cdot),\ldots,\ldots,\mathsf{z}^{j,a}_{\scriptscriptstyle L}(\cdot),\mathsf{z}^{j,A}_{\scriptscriptstyle L}(\cdot),\cdot)$ B matrix defined in [23] w.r.t. h_k and J_k $\mathbf{M}^{j,cc}(\cdot) \leftarrow \mathbf{o}_{\mathbf{B}}(\mathbf{z}_{k}^{j,\vartheta}(\cdot), \mathbf{z}_{k}^{j,A}(\cdot), \dots, \dots, \mathbf{z}_{k}^{j,\vartheta}(\cdot), \mathbf{z}_{k}^{j,A}(\cdot), \cdot)$ 11: $\mathbf{s}_{\mathbf{M}}^{j,cv}(\cdot) \leftarrow \mathcal{S}_{\mathbf{u}_{\mathcal{B}}}(\mathbf{z}_{\mathbf{k}}^{j,\theta}(\cdot), \mathbf{z}_{\mathbf{k}}^{j,A}(\cdot), \mathbf{s_{c}}, \mathbf{s_{C}}, \dots, \mathbf{z}_{\mathbf{k}}^{j,\theta}(\cdot), \mathbf{z}_{\mathbf{k}}^{j,A}(\cdot), \mathbf{s_{c}}, \mathbf{s_{C}}, \cdot)$ 12: $\mathbf{s}_{\mathbf{M}}^{j,cc}(\cdot) \leftarrow \mathcal{S}_{\mathbf{0}_{R}}(\mathbf{z}_{\mathbf{k}}^{j,a}(\cdot), \mathbf{z}_{\mathbf{k}}^{j,A}(\cdot), \mathbf{s}_{\mathbf{c}}, \mathbf{s}_{\mathbf{C}}, \dots, \mathbf{z}_{\mathbf{k}}^{j,a}(\cdot), \mathbf{z}_{\mathbf{k}}^{j,A}(\cdot), \mathbf{s}_{\mathbf{c}}, \mathbf{s}_{\mathbf{C}}, \cdot)$ 13: $\mathbf{z}_{k}^{j+1,cv}(\cdot) \leftarrow \bar{\mathbf{u}}_{\mathbf{w}}(\boldsymbol{\gamma}^{j}(\cdot), \mathbf{\gamma}^{j}(\cdot), \mathbf{M}^{j,cv}(\cdot), \mathbf{M}^{j,cc}(\cdot), \mathbf{z}_{k}^{j,cv}(\cdot), \mathbf{z}_{k}^{j,cc}(\cdot), \mathbf{z}_{k-1}^{cv}(\cdot) \mathbf{z}_{k-2}^{cc}(\cdot), \mathbf{z}_{k-2}^{cv}(\cdot) \mathbf{z}_{k-2}^{cc}(\cdot), \mathbf{z}_{k-2}^{cv}(\cdot), \mathbf{z}_{k-2}$ 15: $\mathbf{z}_{k}^{j+1,cc}(\cdot) \leftarrow \bar{\mathbf{o}}_{\mathbf{w}}(\gamma^{j}(\cdot),\gamma^{j}(\cdot),\mathbf{M}^{j,cv}(\cdot),\mathbf{M}^{j,cc}(\cdot),\mathbf{z}_{k}^{j,cv}(\cdot),\mathbf{z}_{k}^{j,cc}(\cdot),\mathbf{z}_{k-1}^{cv}(\cdot)\mathbf{z}_{k-1}^{cc}(\cdot),\mathbf{z}_{k-2}^{cv}(\cdot)\mathbf{z}_{k-2}^{cc}(\cdot),\cdot)$ $\mathbf{s}_{\mathbf{z}}^{j+1,cv}(\cdot) \leftarrow S_{\hat{\mathbf{u}}_{\mathbf{w}}}(\gamma^{j}(\cdot),\gamma^{j}(\cdot),\mathbf{s}_{\mathbf{y}}^{j},\mathbf{s}_{\mathbf{y}}^{j},\mathbf{M}^{j,cv}(\cdot),\mathbf{M}^{j,cc}(\cdot),\mathbf{z}_{\mathbf{k}}^{j,cv}(\cdot),\mathbf{z}_{\mathbf{k}}^{j,cc}(\cdot),\mathbf{s}_{\mathbf{z}}^{j,cv}(\cdot),\mathbf{s}_$ 16: $\mathbf{s}_{\mathbf{z}}^{j+1,cc}(\cdot) \leftarrow S_{\mathbf{\tilde{o}}_{\mathbf{w}}}(\gamma^{j}(\cdot),\gamma^{j}(\cdot),\mathbf{s}_{\mathbf{y}}^{j},\mathbf{s}_{\mathbf{y}}^{j},\mathbf{M}^{j,cv}(\cdot),\mathbf{M}^{j,cc}(\cdot),\mathbf{z}_{\mathbf{z}}^{j,cv}(\cdot),\mathbf{z}_{\mathbf{z}}^{j,cv}(\cdot),\mathbf{s}_$ end for 18:

19: return $\mathbf{z}_{k}^{r,cc}(\cdot), \mathbf{z}_{k}^{r,cc}(\cdot), \mathbf{s}_{\mathbf{z}}^{r,cv}(\cdot), \mathbf{s}_{\mathbf{z}}^{r,cv}(\cdot)$



3. Compute **M**, the relaxation of a preconditioned Jacobian and respective subgradients

The functions $\mathbf{M}_k: P \to M_k$, and $\mathbf{B}_k: X \times \cdots \times X \times P \to M_k$ be defined for $k \in \{1, \dots, K\}$ corresponding to each timestep are defined by:

$$\mathbf{M}_{k}(\cdot) = \mathbf{B}_{k}\left(\mathbf{y}^{(k-1)n_{x}+1}(\cdot), \dots, \mathbf{y}^{kn_{x}}(\cdot), \cdot\right) \equiv \mathbf{Y}_{k}\begin{pmatrix} \nabla_{\mathbf{x}}h_{(k-1)n_{x}+1}(\mathbf{y}^{(k-1)n_{x}+1}(\cdot), \cdot)^{\mathrm{T}} \\ \nabla_{\mathbf{x}}h_{(k-1)n_{x}+2}(\mathbf{y}^{(k-1)n_{x}+2}(\cdot), \cdot)^{\mathrm{T}} \\ \vdots \\ \nabla_{\mathbf{x}}h_{kn_{x}}(\mathbf{y}^{kn_{x}}(\cdot), \cdot)^{\mathrm{T}} \end{pmatrix}$$

- The points $(\mathbf{y}^{j}(\mathbf{p}), \mathbf{p}) \in X \times P$ are the points at which the gradients of $\nabla_{x} h_{j}(\cdot, \cdot)$ are evaluated.
- \mathbf{Y}_k is a matrix which preconditions the system satisfying assumption 3 (\mathbf{M}_k enclosures no singular matrix)

	1: pr	1: procedure $BLOCKRELAX(\mathbf{h}_{k}, X, P, \mathbf{z}_{k-1}^{cv}, \mathbf{z}_{k-1}^{cc}, \mathbf{s}_{\mathbf{z}_{k-1}}^{cv}, \mathbf{z}_{k-1}^{cc}, \mathbf{z}_{k-2}^{cv}, \mathbf{z}_{k-2}^{cc}, \mathbf{s}_{\mathbf{z}_{k-2}}^{cv}, \mathbf{s}_{\mathbf{z}_{k-2}}^{cc}, \mathbf{s}_{\mathbf{z}_{k-2}}^{cv}, \mathbf{s}_{\mathbf{z}_{k-2}}^{cc}, \mathbf{r})$								
1.	2:	2: $z_k^{0,cv}, z_k^{0,cv}, \mathbf{s}_{z_k}^{0,cv}, \mathbf{s}_{z_k}^{0,cv} \leftarrow \mathbf{x}^L, \mathbf{x}^U, 0, 0$								
	3:	3: for $j \leftarrow 0$ to $r - 1$ do								
	4:	$\lambda \leftarrow \lambda \in [0, 1], \tilde{\mathbf{p}} \leftarrow \mathbf{p} \in P$								
2.	5:	$c,C,s_c,s_C \leftarrow \texttt{Aff}(\ldots) \qquad \qquad \texttt{bSubroutine defined in [23]}$								
	6:	$\mathbf{Z}_{k}^{j,p}(\mathbf{p}) \leftarrow \mathbf{c} + \mathbf{s}_{\mathbf{c}}^{T}(\mathbf{p} - \mathbf{\tilde{p}}), \ \forall \mathbf{p} \in \mathcal{P} \qquad \qquad \triangleright \text{ Affine relaxation lower bound}$								
	7:	$\boldsymbol{z}_{k}^{j,A}(\mathbf{p}) \leftarrow \mathbf{C} + \mathbf{s}_{\mathbf{C}}^{T}(\mathbf{p} - \bar{\mathbf{p}}), \ \forall \mathbf{p} \in \boldsymbol{P} \qquad \qquad \triangleright \text{Affine relaxation upper bound}$								
	8:	$\gamma^{i}(\cdot) \leftarrow \lambda \mathbf{z}_{k}^{i,a}(\cdot) + (1-\lambda)\mathbf{z}_{k}^{i,A}(\cdot)$								
	9:	$\mathbf{s}_{\gamma}^{l} \leftarrow \lambda \mathbf{s}_{\mathbf{c}} + (1 - \lambda) \mathbf{s}_{\mathbf{c}}$								
3.	10:	$M^{j,\mathrm{cv}}(\cdot) \leftarrow u_B(z^{j,a}_k(\cdot),z^{j,A}_k(\cdot),\ldots,z^{j,a}_k(\cdot),z^{j,A}_k(\cdot),\cdot) \qquad \qquad b \text{ B matrix defined in [23] w.r.t. } \mathbf{h}_k \text{ and } \mathbf{J}_k$								
	11:	$M^{j,cc}(\cdot) \leftarrow o_{B}(z_k^{j,s}(\cdot), z_k^{j,A}(\cdot), \dots, z_k^{j,s}(\cdot), z_k^{j,A}(\cdot), \cdot)$								
	12:	$\mathbf{s}_{\mathbf{M}}^{j,\varepsilon \nu}(\cdot) \leftarrow \mathcal{S}_{\mathbf{u}_{\mathcal{B}}}(\mathbf{z}_{k}^{j,g}(\cdot),\mathbf{z}_{k}^{j,A}(\cdot),\mathbf{s}_{c},\mathbf{s}_{C},\ldots,\mathbf{z}_{k}^{j,g}(\cdot),\mathbf{z}_{k}^{j,A}(\cdot),\mathbf{s}_{c},\mathbf{s}_{C},\cdot)$								
	13:	$\mathbf{s}_{\mathbf{M}}^{j,cc}(\cdot) \leftarrow S_{0_{\mathcal{B}}}(\mathbf{z}_{k}^{j,g}(\cdot),\mathbf{z}_{k}^{j,A}(\cdot),\mathbf{s}_{c},\mathbf{s}_{C},\ldots,\mathbf{z}_{k}^{j,a}(\cdot),\mathbf{z}_{k}^{j,A}(\cdot),\mathbf{s}_{c},\mathbf{s}_{C},\cdot)$								
	14:	$\mathbf{z}_{k}^{i_{1}i_{c}v}(\cdot) \leftarrow \bar{\mathbf{u}}_{\mathbf{y}}(\gamma^{j}(\cdot), \gamma^{j}(\cdot), \mathbf{M}^{j,cv}(\cdot), \mathbf{M}^{j,cc}(\cdot), \mathbf{z}_{k}^{j,cv}(\cdot), \mathbf{z}_{k}^{j,c}(\cdot), \mathbf{z}_{k-1}^{cv}(\cdot)\mathbf{z}_{k-1}^{cc}(\cdot), \mathbf{z}_{k-2}^{cv}(\cdot)\mathbf{z}_{k-2}^{cc}(\cdot), \mathbf{z}_{k-1}^{cv}(\cdot)\mathbf{z}_{k-2}^{cv}(\cdot)\mathbf{z}_{k-2}^{cv}(\cdot), \mathbf{z}_{k-1}^{cv}(\cdot)\mathbf{z}_{k-2}^{cv}(\cdot)\mathbf{z}_{k-2}^{cv}(\cdot)\mathbf{z}_{k-2}^{cv}(\cdot), \mathbf{z}_{k-2}^{cv}(\cdot)\mathbf{z}_{k-2}^{cv}(\cdot), \mathbf{z}_{k-2}^{cv}(\cdot)\mathbf{z}_{k-2}^{cv}(\cdot), \mathbf{z}_{k-2}^{cv}(\cdot)\mathbf{z}_{k-2}^{cv}(\cdot), \mathbf{z}_{k-2}^{cv}(\cdot), \mathbf{z}_{k-2}$								
	15:	$\mathbf{z}_{k}^{i+1,cc}(\cdot) \leftarrow \bar{\mathbf{o}}_{\mathbf{V}}(\gamma^{j}(\cdot),\gamma^{j}(\cdot),\mathbf{M}^{j,c\mathbf{v}}(\cdot),\mathbf{M}^{j,cc}(\cdot),\mathbf{z}_{k}^{i,c\mathbf{v}}(\cdot),\mathbf{z}_{k}^{i,c\mathbf{v}}(\cdot),\mathbf{z}_{k-1}^{c\mathbf{v}}(\cdot)\mathbf{z}_{k-2}^{cc}(\cdot)$								
	16:	$\mathbf{s}_{\mathbf{z}}^{j+1,c\nu}(\cdot) \leftarrow \mathcal{S}_{0_{\mathbf{y}}}(\gamma^{j}(\cdot),\gamma^{j}(\cdot),\mathbf{s}_{\mathbf{y}}^{j},\mathbf{s}_{\mathbf{y}}^{j},\mathbf{M}^{j,c\nu}(\cdot),\mathbf{M}^{j,cc}(\cdot),\mathbf{z}_{k}^{j,c\nu}(\cdot),\mathbf{z}_{k}^{j,cc}(\cdot),\mathbf{s}_{\mathbf{z}}^{j,c\nu}(\cdot),\mathbf{s}_{\mathbf{z}}$								
	17:	$\mathbf{s}_{\mathbf{z}}^{j+1,cc}(\cdot) \leftarrow \mathcal{S}_{0_{\mathbf{y}}}(\gamma^{j}(\cdot),\gamma^{j}(\cdot),\mathbf{s}_{\mathbf{y}}^{j},\mathbf{s}_{\mathbf{y}}^{j},\mathbf{M}^{j,cv}(\cdot),\mathbf{M}^{j,cc}(\cdot),\mathbf{z}_{k}^{j,cv}(\cdot),\mathbf{z}_{k}^{j,cc}(\cdot),\mathbf{s}_{\mathbf{z}}^{j,cv}(\cdot),\mathbf{s}_{\mathbf{z}}$								
	18:	end for								

19: return $\mathbf{z}_{k}^{r,cc}(\cdot), \mathbf{z}_{k}^{r,cc}(\cdot), \mathbf{s}_{\mathbf{z}}^{r,cv}(\cdot), \mathbf{s}_{\mathbf{z}}^{r,cc}(\cdot)$



4(a). Compute relaxation of $\boldsymbol{\psi}_k$ (a valid relaxation of \mathbf{z}_k^{J+1}) and the respective subgradients thereof.

Let $\mathbf{b}_k: X \times X \times X \times P \to \mathbb{R}^{n_x}$ such that $\mathbf{b}_k = \mathbf{Y}_k \theta_{k-2}^s$ with $\theta \in \{\xi, \zeta\}, s \in \{1,2\}$, and $2 \le k \le K$. Define the function $\psi_k: X \times M \times X \times X \times P \to \mathbb{R}^{n_x}$ such that $\forall (\tilde{\gamma}, \tilde{\mathbf{M}}, \tilde{\mathbf{z}}_k, \tilde{\mathbf{z}}_{k-1}, \tilde{\mathbf{z}}_{k-2}, \mathbf{p}) \in X \times M \times X \times X \times P$, we have $\psi_k(\tilde{\gamma}, \tilde{\mathbf{M}}, \tilde{\mathbf{z}}_k, \tilde{\mathbf{z}}_{k-1}, \tilde{\mathbf{z}}_k) = \tilde{\mathbf{z}}_k^*$, where the ith component of $\tilde{\mathbf{z}}_k^*$ is given by the loop:

for
$$i = 1, ..., n_x$$
 do
 $\tilde{z}_{k,i}^* \coloneqq \tilde{\gamma}_i - \frac{b_{k,i}(\tilde{\gamma}, \mathbf{z}_{k-1}(\mathbf{p}), \mathbf{z}_{k-2}(\mathbf{p}), \mathbf{p}) + \sum_{j < i} \tilde{m}_{ij} \left(\overline{z}_{k,j}^* - \tilde{\gamma}_j\right) + \sum_{j > i} \tilde{m}_{ij} \left(\tilde{z}_{k,j} - \tilde{\gamma}_j\right)}{\tilde{m}_{ii}}$

	1: p	1: procedure $BLOCKRELAX(\mathbf{h}_{k}, X, P, \mathbf{z}_{k-1}^{cv}, \mathbf{z}_{k-1}^{cc}, \mathbf{s}_{k-1}^{cv}, \mathbf{s}_{k-1}^{cc}, \mathbf{z}_{k-2}^{cv}, \mathbf{z}_{k-2}^{cc}, \mathbf{s}_{k-2}^{cv}, \mathbf{s}_{k-2}^{cc}, \mathbf{r})$						
1.	2:	2: $\mathbf{z}_{k}^{0,cv}, \mathbf{z}_{k}^{0,cc}, \mathbf{s}_{\mathbf{z}_{k}}^{0,cv}, \mathbf{s}_{\mathbf{z}_{k}}^{0,cc} \leftarrow \mathbf{x}^{t}, \mathbf{x}^{U}, 0, 0$						
	3:	for $j \leftarrow 0$ to $r - 1$ do						
	4:	$\lambda \leftarrow \lambda \in [0,1], \tilde{\mathbf{p}} \leftarrow \mathbf{p} \in P$						
2.	5:	$c,C,s_c,s_C \leftarrow \texttt{Aff}(\ldots) \qquad \qquad \texttt{> Subroutine defined in [23]}$						
	6:	$\mathbf{z}_{k}^{j,s}(\mathbf{p}) \leftarrow \mathbf{c} + \mathbf{s}_{\mathbf{c}}^{T}(\mathbf{p} - \hat{\mathbf{p}}), \ \forall \mathbf{p} \in P$ \triangleright Affine relaxation lower bound						
	7:	$z_k^{j,A}(\mathbf{p}) \gets \mathbf{C} + \mathbf{s}_{\mathbf{C}}^{T}(\mathbf{p} - \mathbf{\tilde{p}}), \ \forall \mathbf{p} \in \mathcal{P} \qquad \qquad \triangleright \text{ Affine relaxation upper bound}$						
	8:	$\gamma^{j}(\cdot) \leftarrow \lambda \overline{x}_{k}^{j,g}(\cdot) + (1-\lambda)\overline{x}_{k}^{j,A}(\cdot)$						
	9:	$\mathbf{s}_{\mathbf{\gamma}}^{j} \leftarrow \lambda \mathbf{s}_{\mathbf{c}} + (1 - \lambda) \mathbf{s}_{\mathbf{C}}$						
3.	10:	$M^{j, c_{Y}}(\cdot) \leftarrow u_{B}(\mathbf{z}_{k}^{j, a}(\cdot), \mathbf{z}_{k}^{j, a}(\cdot), \dots, \dots, \mathbf{z}_{k}^{j, a}(\cdot), \mathbf{z}_{k}^{j, A}(\cdot), \cdot) \qquad \qquad b \text{ B matrix defined in [23] w.r.t. } \mathbf{h}_{k} \text{ and } \mathbf{J}_{k}$						
	11:	$M^{j,\varepsilon\varepsilon}(\cdot) \leftarrow og(x_k^{j,\vartheta}(\cdot),x_k^{j,\lambda}(\cdot),\ldots,\ldots,x_k^{j,\vartheta}(\cdot),x_k^{j,\lambda}(\cdot),\cdot)$						
	12:	$\mathbf{s}_{\mathbf{M}}^{j,\varepsilon \gamma}(\cdot) \leftarrow \mathcal{S}_{\mathbf{u}g}(\mathbf{z}_{k}^{j,s}(\cdot),\mathbf{z}_{k}^{j,A}(\cdot),\mathbf{s}_{\mathbf{C}},\mathbf{s}_{\mathbf{C}},\ldots,\mathbf{z}_{k}^{j,s}(\cdot),\mathbf{z}_{k}^{j,A}(\cdot),\mathbf{s}_{\mathbf{c}},\mathbf{s}_{\mathbf{C}},\cdot)$						
	13:	$\mathbf{s}_{\mathbf{M}}^{j,cc}(\cdot) \leftarrow \mathcal{S}_{0_{\mathcal{B}}}(\mathbf{z}_{k}^{j,\sigma}(\cdot), \mathbf{z}_{k}^{j,A}(\cdot), \mathbf{s}_{\mathbf{c}}, \mathbf{s}_{\mathbf{C}}, \dots, \mathbf{z}_{k}^{j,a}(\cdot), \mathbf{z}_{k}^{j,A}(\cdot), \mathbf{s}_{\mathbf{c}}, \mathbf{s}_{\mathbf{C}}, \cdot)$						
	14:	$\mathbf{z}_{k}^{j+1,c\mathbf{v}}(\cdot) \leftarrow \hat{\mathbf{u}}_{\mathbf{v}}(\boldsymbol{\gamma}^{j}(\cdot), \boldsymbol{\gamma}^{j}(\cdot), \mathbf{M}^{j,c\mathbf{v}}(\cdot), \mathbf{M}^{j,c\mathbf{c}}(\cdot), \mathbf{z}_{k}^{j,c\mathbf{v}}(\cdot), \mathbf{z}_{k-1}^{j,c\mathbf{v}}(\cdot) \mathbf{z}_{k-1}^{c\mathbf{v}}(\cdot) \mathbf{z}_{k-2}^{c\mathbf{v}}(\cdot) $						
4.	15:	$\mathbf{z}_{k}^{j+1,cc}(\cdot) \leftarrow \bar{\mathbf{o}}_{\psi}(\gamma^{j}(\cdot),\gamma^{j}(\cdot),\mathbf{M}^{j,cv}(\cdot),\mathbf{M}^{j,cc}(\cdot),\mathbf{z}_{k}^{j,cv}(\cdot),\mathbf{z}_{k}^{j,cv}(\cdot),\mathbf{z}_{k-1}^{j,c}(\cdot)\mathbf{z}_{k-1}^{cc}(\cdot),\mathbf{z}_{k-2}^{cc}(\cdot)\mathbf{z}_{k-2}^{cc}(\cdot),\mathbf{z}_{k-2}^{cc}(\cdot)\mathbf{z}_{k-2}^{cc}(\cdot),\mathbf{z}_{k-2}^{cc}(\cdot)\mathbf{z}_{k-2}^{cc}(\cdot),\mathbf{z}_{k-2}^{cc}(\cdot)\mathbf{z}_{k-2}^{cc}(\cdot),\mathbf{z}_{k-2}^{cc}(\cdot)\mathbf{z}_{k-2}^{cc}(\cdot),\mathbf{z}_{k-2}^{cc}(\cdot)\mathbf{z}_{k-2}^{cc}(\cdot),\mathbf{z}_{k-2}^{cc}(\cdot)\mathbf{z}_{k-2}^{cc}(\cdot),\mathbf{z}_{k-2}^{cc}(\cdot)\mathbf{z}_{k-2}^{cc}(\cdot),\mathbf{z}_{k-2}^{cc}(\cdot)\mathbf{z}_{k-2}^{cc}(\cdot),\mathbf{z}_{k-2}^{cc}(\cdot)\mathbf{z}_{k-2}^{cc}(\cdot),\mathbf{z}_{k-2}^{cc$						
	16:	$\mathbf{s}_{\mathbf{z}}^{j+1,cv}(\cdot) \leftarrow \mathcal{S}_{0_{\mathbf{y}}}(\gamma^{j}(\cdot),\gamma^{j}(\cdot),\mathbf{s}_{\mathbf{y}}^{j},\mathbf{s}_{\mathbf{y}}^{j},\mathbf{M}^{j,cv}(\cdot),\mathbf{M}^{j,cc}(\cdot),\mathbf{z}_{k}^{j,cv}(\cdot),\mathbf{z}_{k}^{j,cv}(\cdot),\mathbf{s}_{\mathbf{z}}$						
	17:	$s_z^{j+1,cc}(\cdot) \leftarrow \mathcal{S}_{0_{\psi}}(\gamma^j(\cdot),\overline{\gamma^j}(\cdot),\overline{s_y'},\overline{s_y'},M^{j,c\psi}(\cdot),M^{j,cc}(\cdot),z_k^{j,c\psi}(\cdot),\overline{z}_k^{j,c\psi}(\cdot),\overline{s}_z^{j,c\psi}(\cdot),\overline{s}_z^{j,c\psi}(\cdot),\cdot)$						
	18:	end for						

 $19: \quad \operatorname{return} \mathbf{z}_{k}^{r,cc}(\cdot), \mathbf{z}_{k}^{r,cc}(\cdot), \mathbf{s}_{\mathbf{z}}^{r,cv}(\cdot), \mathbf{s}_{\mathbf{z}}^{r,cc}(\cdot)$



end



16

4(b). Compute composite relaxation of $\boldsymbol{\psi}_k$ (a valid relaxation of \boldsymbol{z}_k^{J+1}) and the respective subgradients thereof.

Let $\mathbf{u}_{\phi}, \mathbf{o}_{\phi}$ be composite relaxations of ϕ on $X \times P$. The functions $\mathbf{u}_{\phi}, \mathbf{o}_{\phi} : \mathbb{R}^{n_{\chi}} \times \mathbb{R}^{n_{\chi}} \times P \to \mathbb{R}^{n_{\chi}}$ will be defined as:

 $\overline{\mathbf{u}}_{\phi}(\mathbf{z}^{cv}, \mathbf{z}^{cc}, \mathbf{p}) \equiv \max\left\{\mathbf{z}^{cv}, \mathbf{u}_{\phi}(\mathbf{z}^{cv}, \mathbf{z}^{cc}, \mathbf{p})\right\}$ $\overline{\mathbf{o}}_{\phi}(\mathbf{z}^{cv}, \mathbf{z}^{cc}, \mathbf{p}) \equiv \min\left\{\mathbf{z}^{cc}, \mathbf{o}_{\phi}(\mathbf{z}^{cv}, \mathbf{z}^{cc}, \mathbf{p})\right\}$

 $\forall (\mathbf{z}^{cv}, \mathbf{z}^{cc}, \mathbf{p}) \in \mathbb{R}^{n_x} \times \mathbb{R}^{n_x} \times P$

Relaxation at next step is composite relaxation of GS result and prior relaxation.



19: return $\mathbf{z}_{k}^{r,cc}(\cdot), \mathbf{z}_{k}^{r,cc}(\cdot), \mathbf{s}_{\mathbf{z}}^{r,cv}(\cdot), \mathbf{s}_{\mathbf{z}}^{r,cv}(\cdot)$

4(b). Compute composite relaxation of $\boldsymbol{\psi}_k$ (a valid relaxation of \boldsymbol{z}_k^{J+1}) and the respective subgradients thereof.





19: return $\mathbf{z}_{k}^{r,cc}(\cdot), \mathbf{z}_{k}^{r,cc}(\cdot), \mathbf{s}_{\mathbf{z}}^{r,cv}(\cdot), \mathbf{s}_{\mathbf{z}}^{r,cv}(\cdot)$



2. Compute affine function between current relaxations and respective subgradients





 $19: \qquad \text{return} \, \mathbf{z}_k^{r,cc}(\cdot\,), \mathbf{z}_k^{r,cc}(\cdot\,), \mathbf{s}_{\mathbf{z}}^{r,cv}(\cdot\,), \mathbf{s}_{\mathbf{z}}^{r,cc}(\cdot\,) \\$

Next iteration...





Relaxation Algorithm

- 1. Compute relaxation of initial conditions using standard McCormick arithmetic.
- 2. Performance a relaxation in a block sequential fashion to compute relaxations and subgradients at next time step and repeat.
- 3. Use relaxations and subgradients available at discrete time points to construct relaxations of the objective and constraints. *Interpolating if necessary.*

1: procedure IVPBOUND(X, P,h, x₀, r)

- : $\mathbf{z}_{0}^{cv}, \mathbf{z}_{0}^{cc}, \mathbf{s}_{\mathbf{z}_{0}}^{cv}, \mathbf{s}_{\mathbf{z}_{0}}^{cv} \leftarrow \mathsf{McCormickRelax}(X, P, \mathbf{x}_{0}(\cdot))$ \triangleright Standard McCormick relaxation of \mathbf{x}_{0} on P
- $3: \qquad \textbf{z}_1^{cv}(\cdot), \textbf{z}_1^{cc}(\cdot), \textbf{s}_{\textbf{z}_1}^{cv}(\cdot), \textbf{s}_{\textbf{z}_1}^{cc}(\cdot) \leftarrow \text{BlockRelax}(\textbf{h}_1, X, P, \textbf{z}_0^{cv}, \textbf{z}_0^{cc}, \textbf{s}_{\textbf{z}_0}^{cv}, \textbf{s}_{\textbf{z}_0}^{cc}, \textbf{x}^L, \textbf{x}^U, \textbf{0}, \textbf{0}, r)$
- 4: **for** $i \leftarrow 2$ **to** K **do**
- 5: $\mathbf{z}_{i}^{cv}(\cdot), \mathbf{z}_{i}^{cc}(\cdot), \mathbf{s}_{\mathbf{z}_{i}}^{cv}(\cdot), \mathbf{s}_{\mathbf{z}_{i}}^{cc}(\cdot) \leftarrow \mathsf{BlockRelax}(\mathbf{h}_{i}, X, P, \mathbf{z}_{i-1}^{cv}, \mathbf{z}_{i-1}^{cc}, \mathbf{s}_{\mathbf{z}_{i-1}}^{cv}, \mathbf{z}_{i-2}^{cc}, \mathbf{z}_{i-2}^{cv}, \mathbf{z}_{i-2}^{cc}, \mathbf{s}_{\mathbf{z}_{i-2}}^{cv}, \mathbf{s}_{\mathbf{z}_{i-2}}^{cc}, \mathbf{r})$
- 6: end for
- 7: return $z^{cv}(\cdot), z^{cc}(\cdot), s_z^{cv}(\cdot), s_z^{cc}(\cdot)$

8: end procedure

21

Exhibits Partition Convergence

Partition Convergence¹³:

Consider a nested sequence of intervals $\{P_q\}, P_q \subset P, q \in \mathbb{N}$, such that $\{P_q\} \rightarrow [\overline{\mathbf{p}}, \overline{\mathbf{p}}]$ for some $\overline{\mathbf{p}} \in \mathbf{P}$. Let z_q^{cv}, z_q^{cc} be relaxations of z on P_q obtained using the prior algorithm. Let $\phi_q^{cv}(\cdot) = u_{\phi}(\mathbf{z}_k^{cv}(\cdot), \mathbf{z}_k^{cc}(\cdot), \cdot)$ be a convex relaxation of the objective function ϕ on P_q . Let $\hat{\phi}_q^{cv} = \min_{\mathbf{p} \in P_q} \phi_q^{cv}$, then $\lim_{q \to \infty} \hat{\phi}_q^{cv} = \phi_q^{cv}(\mathbf{z}(\overline{\mathbf{p}}), \overline{\mathbf{p}})$



Proof of Partition Convergence

- K = 1 case is trivially true... Proceed by contradiction.
- Suppose that for K > 1, $\lim_{q \to \infty} \hat{\phi}_q^{cv} \neq \phi_q^{cv}(\mathbf{z}(\overline{\mathbf{p}}), \overline{\mathbf{p}})$.

. . .

- u_{ϕ} continuous and exhibits partition convergence as it is constructed using a generalized McCormick relaxation framework¹⁴. Then
 - Either $\lim_{q\to\infty} \mathbf{z}_{k,q}^{c\nu} \neq \mathbf{z}_{k,q}(\overline{\mathbf{p}})$ or $\lim_{q\to\infty} \mathbf{z}_{k,q}^{c\nu} \neq \mathbf{z}_{k,q}(\overline{\mathbf{p}})$ which implies
 - Either $\lim_{q\to\infty} \mathbf{z}_{k-1,q}^{cv} \neq \mathbf{z}_{k-1,q}(\overline{\mathbf{p}})$ or $\lim_{q\to\infty} \mathbf{z}_{k-1,q}^{cv} \neq \mathbf{z}_{k-1,q}(\overline{\mathbf{p}})$ which implies
 - Either $\lim_{q\to\infty} \mathbf{z}_{k-2,q}^{c\nu} \neq \mathbf{z}_{k-2,q}(\overline{\mathbf{p}})$ or $\lim_{q\to\infty} \mathbf{z}_{k-2,q}^{c\nu} \neq \mathbf{z}_{k-2,q}(\overline{\mathbf{p}})$ which implies
 - Either $\lim_{q\to\infty} \mathbf{z}_{1,q}^{cv} \neq \mathbf{z}_{1,q}(\overline{\mathbf{p}})$ or $\lim_{q\to\infty} \mathbf{z}_{1,q}^{cv} \neq \mathbf{z}_{1,q}(\overline{\mathbf{p}})$ which is contradiction.





Implementation

- Branch and bound algorithm as an extension to the EAGO global optimizer available at <u>https://github.com/PSORLab/EAGODifferential.jl</u>
- IntervalArithmetic.jl for validated interval calculations.
- Relaxations from McCormick submodule of EAGO.jl.
- All simulations run on single thread of Intel Xeon E3-1270 v5 3.60/4.00GHz processor with 16GM ECC RAM, Ubuntu 18.04LTS using Julia v1.1.0¹⁵. Intel MKL 2019 (Update 2) for BLAS/LAPACK.



https://github.com/PSORLab/EAGO.jl



A 1D Example

• Consider the 1D pODE-IVP:

 $\frac{dx}{dt}(p,t) = -x^{2} + p$ $x_{o}(p) = 9$ $t \in [0,1], \qquad x \in X = [0.1,9]$ $p \in P = [-1,1]$

- The BDF method enclosure depends heavily on the chosen step size whereas the AM method bounds do not.
- In either case, applying 5 iterations of a corresponding parametric interval method further tightens the relaxations.



Kinetic Problem

Problem Statement

- Fit the rate constants (k_{2f}, k_{3f}, k₄) of oxygen addition to cyclohexadienyl radicals to data. ¹⁶
- First addressed by global by Singer et al.¹⁷
- Explicit Euler form solved by by Mitsos¹⁸
- Implicit Euler form addressed in Stuber¹⁹

pODE IVP:

$$\begin{aligned} \dot{x}_A &= k_1 x_Z x_Y - c_{O2} \left(k_{2f} + k_{3f} \right) x_A + \left(k_{2f} / K_2 \right) x_D + \left(k_{3f} / K_3 \right) x_B - k_5 x_A^2, \\ \dot{x}_B &= k_{3f} c_{O2} x_A - \left(k_{3f} / K_3 + k_4 \right) x_B, \quad \dot{x}_D = k_{2f} c_{O2} x_A - \left(k_{2f} / K_2 \right) x_D, \\ \dot{x}_Y &= -k_{1s} x_Y x_Z, \qquad \dot{x}_Z = -k_1 x_Y x_Z, \\ \dot{x}(t=0) &= (0,0,0,0.4,140) \end{aligned}$$

Objective Variables

$$f^* = \min_{p \in P} \sum_{i=1}^{n} (I^i - I^i_{data})^2$$

s.t. $I^i = x^i_A + \frac{2}{21} x^i_B + \frac{2}{21} x^i_D$

Decision Variables

$$\boldsymbol{p}=(k_{2f},k_{3f},k_4)$$

State Variables

$$\boldsymbol{x} = (x_A, x_B, x_D, x_Y, x_Z)$$

Constants

$$k_1, k_{1s}, k_5, K_2, K_3, c_{O2}, \Delta t, n$$

16. J. W. Taylor, et al. Direct measurement of the fast, reversible addition of oxygen to cyclohexadienyl radicals in nonpolar solvents, *Phys. Chem. A*, 108 (2004), pp. 7193–7203.
17. A. B. Singer et al., Global dynamic optimization for parameter estimation in chemical kinetics A. B. Singer et al., *J. Phys. Chem. A*, 110 (2006), pp. 971–976
18. Mitsos et al. McCormick-based relaxations of algorithms. *SIAM Journal on Optimization*, SIAM (2009) 20, 73-601
19. Stuber, M.D. et al. Convex and concave relaxations of implicit functions. *Optimization Methods and Software* (2015), 30, 424-460

Implementation

Problem Statement

Fit the rate addition

Explicit I

- First add Affine relaxations used to compute lower bound (CPLEX 12.8).
- Implicit Upper-bound computed by integrating ODE at midpoint of active node then evaluating objective & constraints.

pODE IVP:

 $\dot{x}_{B} = k_{3f}$

- Duality-based bound tightening was performed. Two iterations of each PILMS $\dot{x}_A = k_1 x_2$ with was used after five iterations of a block sequential parametric interval method.
 - Absolute and relative convergence tolerances for the B&B algorithm of 10^{-2} and 10^{-5} , respectively.

16. J. W. Taylor, et al. D

 $\dot{x}(t=0)$

17. A. B. Singer et al., Glob

18. Mitsos et al. McCormick-based relaxations of algorithms. SIAM Journal on Optimization, SIAM (2009) 20, 73-601

19. Stuber, M.D. et al. Convex and concave relaxations of implicit functions. Optimization Methods and Software (2015), 30, 424-460

data

 $, x_{Z})$

 $_{02},\Delta t,n$

2004), pp. 7193–7203.

Kinetic Problem

Kinetic problem was solved subject to three discretization schemes for K = 100 and K = 200.

Solution Method	K	Iterations	Average time per iteration	Solution time	SSE at Solution
Implicit Euler	100	33987	45×10 ⁻³ s	29.7min	26947.246
	200	23,525	59×10 ⁻³ s	23.4min	16796.038
2-Step AM	100	62024	12×10 ⁻² s	>2 h	N/A*
	200	6068	22×10 ⁻² s	22.6min	13077.998
2-Step BDF	100	88408	81×10 ⁻³ s	>2 h	N/A*
	200	27600	26×10 ⁻² s	>2 h	N/A*
Explicit Euler	100	>300,000	23×10 ⁻⁴ s	>2 h	N/A
	200	>300,000	24×10 ⁻⁴ s	>2 h	N/A



28

A plug-flow reactor problem

Optimization Problem

$$\phi^* = \min_{p \in [0,1]} p$$

s.t. $\hat{x}_K (t = t_f, p) - 0.08 \le 0$

Method of Lines Discretization

$$\frac{d\hat{\mathbf{x}}}{dt}(p,t) = -\frac{\Delta\hat{\mathbf{x}}}{\Delta y} - Da(p)\hat{\mathbf{x}}$$
$$Da(p) = 0.1 + 0.3p$$

Problem Setup

- Start-up of a single species PFR
- Inlet concentration is 1
- Initial concentration is 0
- Number of timesteps of variable size = 30
- Number of spatial discretization points = 20
- The PFR is sparged and we can module the rate of reaction by changing airflow (*p*)



Future Directions

- Extension to provide enclosure of truncation error.
 - Interval forms have been addressed.
 - Generalization to relaxations and use in global optimization outstanding.

 Constructing bounds of PDE constrained systems via relaxation of Crank-Nicolson methods.



Courtesy of wikimedia commons (https://commons.wikimedia.org/wiki/File:HeatEquationCNApproximate.svg)



Acknowledgements

Collaborator(s):



Anne Le, PhD Student, UMass Amherst

PSOR Lab at UCONN (https://psor.uconn.edu/):





Funding:

- National Science Foundation, Award No.: 1560072, 1706343, 1932723*
- University of Connecticut

Any opinions, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the National Science Foundation.

Fin...

Any Questions?

