Mixed-Integer Nonlinear Optimisation: Multi-Term Underestimators

Ruth Misener
r.misener@imperial.ac.uk

Computational Optimisation Group
Centre for Process Systems Engineering
Department of Computing

Imperial College London

computational optimisation group

Centre for Process Systems Engineering

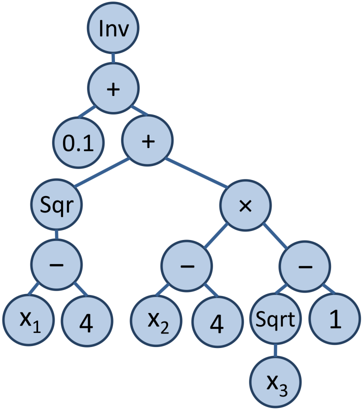
2017/05/11

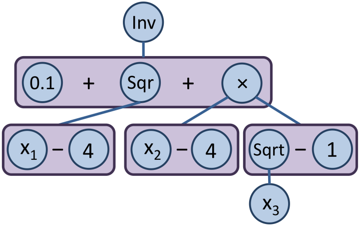
1. Factorable Programming Approach
   - Special Nonconvex Model Forms
   - Difficulties with Factorable Programming Approaches

2. Relaxations for Multi-Term Expressions
   - Reformulation Linearization Technique
   - Edge-Concavity

Introducing Auxilliary Variables[1]

$$1/\left(0.1 + (x_1 - 4)^2 + (x_2 - 4) \times (\sqrt{x_3} - 1)\right)$$





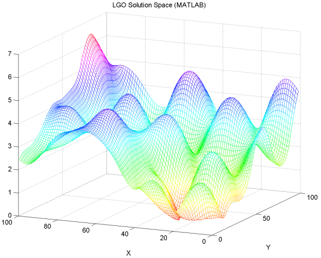
[1]Smith & Pantelides, *Comput Chem Eng*, 1999

Special Nonconvex Model Forms

$$\min_{\mathbf{x}\in D} \; f(\mathbf{x})$$

with: $D := \left\{ \mathbf{x} : \begin{array}{c} g_j(\mathbf{x}) \leq 0, \; j = 1, \ldots, m \\ \mathbf{x}^{\mathrm{L}} \leq \mathbf{x} \leq \mathbf{x}^{\mathrm{U}} \end{array} \right\}$

f continuous on D



- Separable Programs: f, g_j are separable functions, $\sum_k f_k(x_k)$
- Bilinear Programs: f, g_j contain bilinear terms, $x_k \times x_\ell$
- Geometric Programs: f, g_j are posynomial functions, $\sum_k c_k x_1^{a_{1k}} \cdots x_n^{a_{nk}}$, $c_k, a_{1k}, \ldots, a_{nk} \in \mathbb{R}$
- Concave Programs: f is concave on D
  - Global minimum attained at a certain extreme point of D
- Difference of convex (D.C.) Programs: f, g_j can be decomposed as $p(\mathbf{x}) - q(\mathbf{x})$, with p, q convex on D
  - Every twice continuously differentiable function is D.C.
  - The nonconvex programs $\min_{\mathbf{x}\in D} f(\mathbf{x})$ can be converted into D.C. form

Base Terms

Term Type	Mathematical Form	Relaxation Strategies
Constant	c	–
Bilinear / Quadratic	$x_1 \cdot x_2$	McCormick Hull Outer Approximation
Multivariate Signomial	$x_1^{a_1} \cdot x_2^{a_2} \cdots x_n^{a_n}$	Edge-Concave Exponential Transformation Fractional Outer Approximation
Absolute Value	$\lvert a \cdot x + b \rvert$	Equivalent MILP Representation
Exponential	$e^{a \cdot x + b}$	Secant Line Outer Approximation

Base Terms[1] [con't]

Term Type	Mathematical Form	Relaxation Strategies
Linear	$a \cdot x + b$	–
Logarithmic	$\log(a \cdot x + b)$	Secant Line Outer Approximation
Univariate Signomial	x^a	Odd Degree Monomial Secant Line Outer Approximation
Composite		
Exponent	$(a_1 \cdot x + b_1) \cdot e^{a_2 \cdot x + b_2}$	Secant Line
Fractional	$\frac{a_1 \cdot x + b_1}{a_2 \cdot x + b_2}$	Outer Approximation
Logarithm	$(a_1\, x + b_1) \log(a_2\, x + b_2)$	

[1] ANTIGONE: Algorithms for coNTinuous / Integer Global Optimization of Nonlinear Equations, Misener, Floudas, *Journal of Global Optimization*; **59**: 503 - 526, 2014

Composite Terms

Regularly appear in standard libraries: GLOBALLib (*e.g.*, `arki0018`, `chem`, `ex6_1_1`, `ex6_1_2`, `ex14_1_8`, `filter`); MINLPLib (`contvar`); AMPL Book Library (`nltrans`); PrincetonLib (*e.g.*, `bigbank`, `s367`, `s377`).

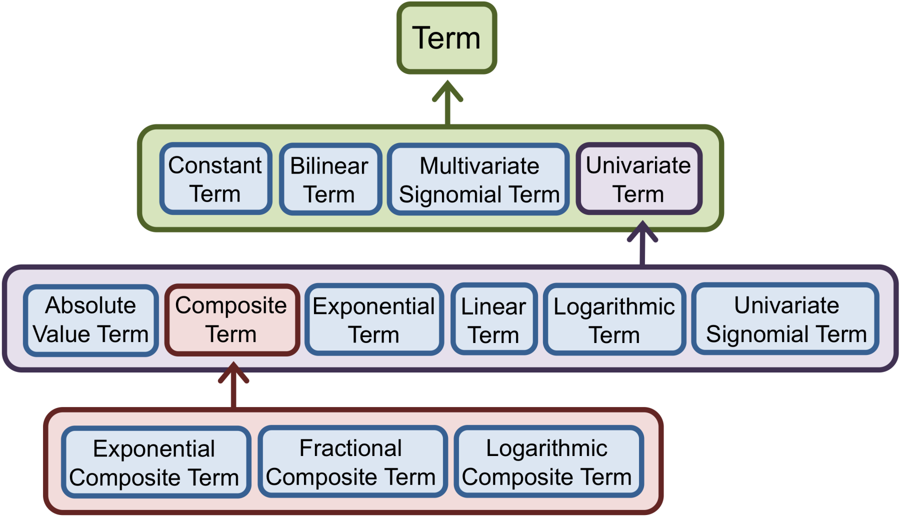
ANTIGONE knows analytic expressions for the second derivatives and uses interval arithmetic on $d^2 f/dx^2$ to determine convexity/concavity regions:

$$\text{Exponential Composite} \quad \begin{cases} f_e(x) & = (a_1 \cdot x + b_1) \cdot e^{a_2 \cdot x + b_2} \\ \Downarrow & \\ \frac{d^2 f_e}{dx^2} & = (a_1 \cdot a_2 \cdot a_2 \cdot x + 2 \cdot a_1 \cdot a_2 + a_2 \cdot a_2 \cdot b_1) \cdot e^{a_2 \cdot x + b_2} \end{cases}$$

$$\text{Fractional Composite} \quad \begin{cases} f_f(x) & = \frac{a_1 \cdot x + b_1}{a_2 \cdot x + b_2} \\ \Downarrow & \\ \frac{d^2 f_f}{dx^2} & = -2 \cdot a_2 \cdot \frac{a_1 \cdot b_2 - a_2 \cdot b_1}{(a_2 \cdot x + b_2)^3} \end{cases}$$

$$\text{Logarithmic Composite} \quad \begin{cases} f_\ell(x) & = (a_1 \cdot x + b_1) \cdot \log(a_2 \cdot x + b_2) \\ \Downarrow & \\ \frac{d^2 f_\ell}{dx^2} & = \frac{a_1 \cdot a_2 \cdot (a_2 \cdot x + b_2) + a_1 \cdot a_2 \cdot b_2 - a_2 \cdot a_2 \cdot b_1}{(a_2 \cdot x + b_2)^2} \end{cases}$$

Combining Relaxations[2]



[2] ANTIGONE: Algorithms for coNTinuous / Integer Global Optimization of Nonlinear Equations, Misener, Floudas, *Journal of Global Optimization*; **59**: 503 - 526, 2014

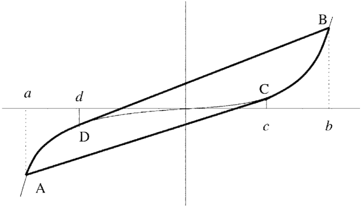
Specialising Terms throughout a Branch & Bound Tree

Absolute Value

Consider the term $|x_i|$. In the most general case, the framework introduces variables x_i^+, x_i^-, and y_i to model the MILP representation of $|x_i|$:

$$|x_i| \implies \begin{cases} |x_i| = x_i^+ + x_i^-; & x_i^- \leq |x_i|^U \cdot (1 - y_i); \\ x_i = x_i^+ - x_i^-; & x_i^+ \leq |x_i|^U \cdot \quad y_i; \\ x_i^+;\ x_i^- \geq 0;\ y_i \in \{0, 1\}. & \end{cases}$$

Odd Powers (Liberti & Pantelides, *J Glob Optim*, 2003):

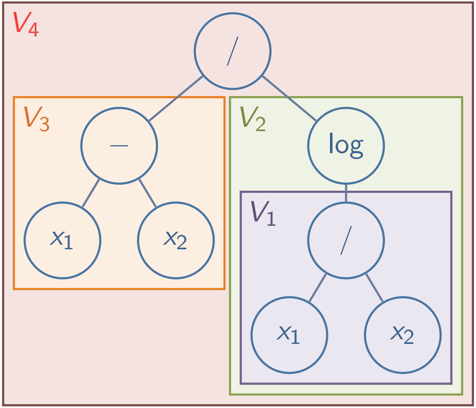


1. Factorable Programming Approach
   - Special Nonconvex Model Forms
   - Difficulties with Factorable Programming Approaches

2. Relaxations for Multi-Term Expressions
   - Reformulation Linearization Technique
   - Edge-Concavity

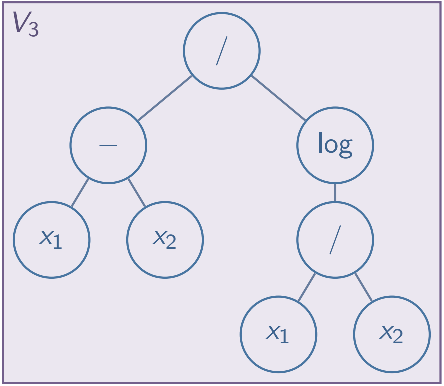
How can we make this simple? M Mistry

- $\Delta T_{\mathrm{LMTD}} = \frac{\mathbf{\Delta T_1} - \mathbf{\Delta T_2}}{\log(\mathbf{\Delta T_1}/\mathbf{\Delta T_2})}$
- LMTD is often approximated to avoid numerical difficulties
  - Chen approximation $\left[\mathbf{\Delta T_1}\ \mathbf{\Delta T_2}\ \frac{\mathbf{\Delta T_1}+\mathbf{\Delta T_2}}{2}\right]^{\frac{1}{3}}$
  - adding a small ε to parameters



How can we make this simple? M Mistry

- $\Delta T_{\mathrm{LMTD}} = \frac{\mathbf{\Delta T_1} - \mathbf{\Delta T_2}}{\log(\mathbf{\Delta T_1}/\mathbf{\Delta T_2})}$
- LMTD is often approximated to avoid numerical difficulties
  - Chen approximation $\left[\mathbf{\Delta T_1}\ \mathbf{\Delta T_2}\ \frac{\mathbf{\Delta T_1}+\mathbf{\Delta T_2}}{2}\right]^{\frac{1}{3}}$
  - adding a small ε to parameters

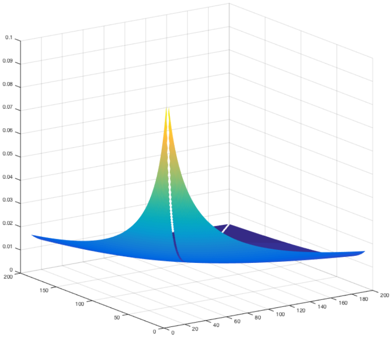


Use of the Reciprocal

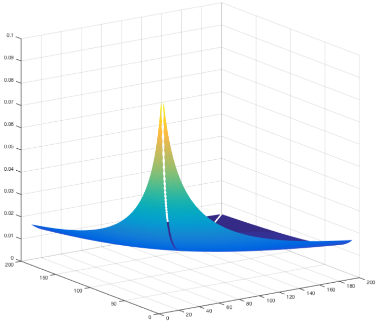
- The area:

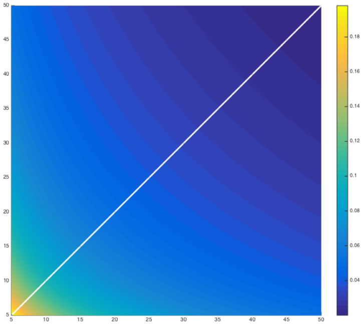
$$A = \frac{q \cdot U}{\Delta\, T_{LMTD}}$$

- We calculate the area using the reciprocal of LMTD.
- Approximation of the reciprocal doesn't introduce a layer of complexity in reasoning errors



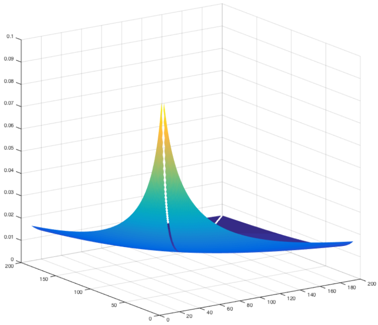
Limits M Mistry

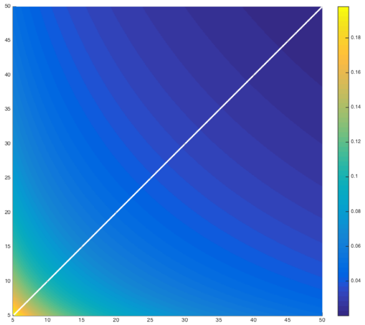




Limits

M Mistry





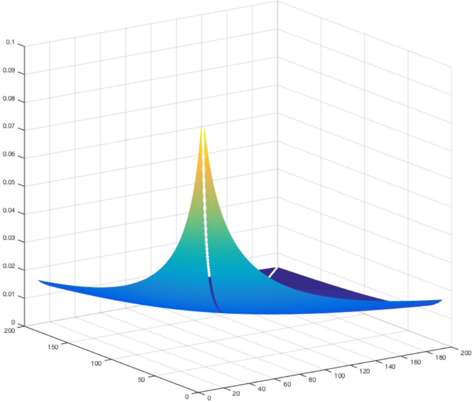
We should use polar coordinates

$$
\begin{aligned}
\lim_{(r,\theta)\to(r_c,{}^{\pi}/_{4})} p(r)\cdot\frac{q_1(\theta)}{q_2(\theta)} &= \left[\lim_{r\to r_c} p(r)\right]\left[\lim_{\theta\to{}^{\pi}/_{4}} \frac{q_1(\theta)}{q_2(\theta)}\right] \\
&= r_c^{-1}\cdot\sqrt{2} \\
&= \frac{1}{c\sqrt{2}}\cdot\sqrt{2} \\
&= \frac{1}{c}
\end{aligned}
$$

Limits

M Mistry

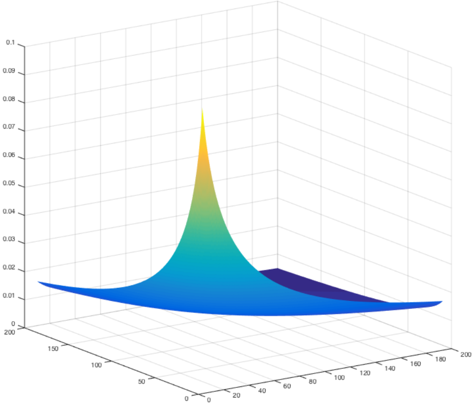
- the reciprocal of LMTD has a set of indeterminacies over the line "$x = y$"
- we showed that the limit exists [Zavala-Río. Femat. Santiesteban-Co, 2005]
- the same proof technique was extended to higher derivatives



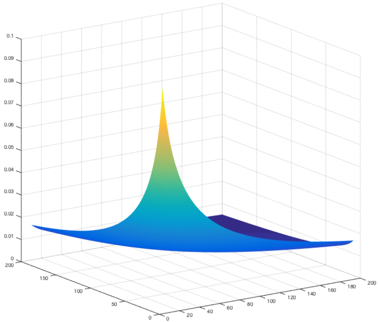
Limits

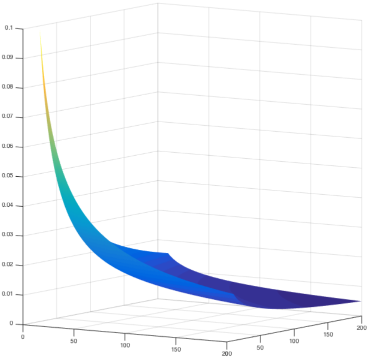
M Mistry

- the reciprocal of LMTD has a set of indeterminacies over the line "$x = y$"
- we showed that the limit exists [Zavala-Río. Femat. Santiesteban-Co, 2005]
- the same proof technique was extended to higher derivatives



Convexity[3]





This function is convex!

[3] Mistry M., Misener R. Optimising Heat Exchanger Network Synthesis using Convexity Properties of the Logarithmic Mean Temperature Difference. *Submitted*, 2016

RLT: Equality Equation/Variable

GloMIQO considers every product of variable x_i that participates nonlinearly in MIQCQP with linear equality equation m containing only continuous variables (*i.e.*, m such that $Q_m = 0;\ c_m = 0;\ b_m^{\mathrm{LO}} = b_m^{\mathrm{UP}} = b_m$):

$$\left(\sum_j a_{m,j} \cdot x_j - b_m \right) \cdot x_i = \sum_j a_{m,j} \cdot x_j \cdot x_i - b_m \cdot x_i = 0$$

If the RLT equation does not increase the number of nonlinear terms in MIQCQP and is not already present in the model formulation, GloMIQO adds it directly to the model.

RLT: Inequality Equation/Variable

GloMIQO considers the product of variable x_i that participates nonlinearly in **??** with linear equation m containing only continuous variables (*i.e.*, $Q_m = 0;\ c_m = 0$). For inequality constraint m, GloMIQO considers four products:

$$\begin{aligned}
\left(a_m \cdot x - b_m^{\mathrm{UP}}\right) \cdot \left(x_i - x_i^{\mathrm{LO}}\right) &\leq 0 \\
\left(a_m \cdot x - b_m^{\mathrm{UP}}\right) \cdot \left(x_i^{\mathrm{UP}} - x_i\right) &\leq 0 \\
\left(b_m^{\mathrm{LO}} - a_m \cdot x\right) \cdot \left(x_i - x_i^{\mathrm{LO}}\right) &\leq 0 \\
\left(b_m^{\mathrm{LO}} - a_m \cdot x\right) \cdot \left(x_i^{\mathrm{UP}} - x_i\right) &\leq 0
\end{aligned}$$

Added if they do not increase the number of bilinear terms.

RLT: Equation/Equation

GloMIQO considers products of two linear equations m, n that contain exclusively continuous variables (*i.e.*, $Q_m = Q_n = 0$ and $c_m = c_n = 0$):

$$\begin{aligned}
-1 \cdot \left(a_m \cdot x - b_m^{\mathrm{UP}}\right) \cdot \left(a_n \cdot x - b_n^{\mathrm{UP}}\right) &\leq 0 \\
\left(a_m \cdot x - b_m^{\mathrm{UP}}\right) \cdot \left(a_n \cdot x - b_n^{\mathrm{LO}}\right) &\leq 0 \\
\left(a_m \cdot x - b_m^{\mathrm{LO}}\right) \cdot \left(a_n \cdot x - b_n^{\mathrm{UP}}\right) &\leq 0 \\
-1 \cdot \left(a_m \cdot x - b_m^{\mathrm{LO}}\right) \cdot \left(a_n \cdot x - b_n^{\mathrm{LO}}\right) &\leq 0
\end{aligned}$$

Added if they do not increase the number of bilinear terms.

The Hartree-Fock Problem instance `beryllium`

$$
\begin{aligned}
\min \quad & -15.73426 \cdot c_{12}^2 - 15.73426 \cdot c_{11}^2 + 0.5721648 \cdot c_{12} \cdot c_{22} \cdot c_{21}^2 \\
& +1.56814504 \cdot c_{12}^2 \cdot c_{11} \cdot c_{21} + 1.56814504 \cdot c_{11}^2 \cdot c_{12} \cdot c_{22} \\
& -7.7290488 \cdot c_{11} \cdot c_{21} - 7.7290488 \cdot c_{12} \cdot c_{22} - 4.204318 \cdot c_{21}^2 \\
& -4.204318 \cdot c_{22}^2 + 2.2988306 \cdot c_{11}^4 + 4.5976612 \cdot c_{11}^2 \cdot c_{12}^2 \\
& -1.329488452 \cdot c_{11} \cdot c_{21} \cdot c_{12} \cdot c_{22} + 0.8353663 \cdot c_{21}^2 \cdot c_{22}^2 + 0.41768315 \cdot c_{21}^4 \\
& +0.41768315 \cdot c_{22}^4 + 2.124875442 \cdot c_{11}^2 \cdot c_{22}^2 \\
& +2.124875442 \cdot c_{12}^2 \cdot c_{21}^2 + 1.460131216 \cdot c_{12}^2 \cdot c_{22}^2 + 0.5721648 \cdot c_{11} \cdot c_{21}^3 \\
& +0.5721648 \cdot c_{12} \cdot c_{22}^3 + 0.5721648 \cdot c_{11} \cdot c_{21} \cdot c_{22}^2 \\
& +1.56814504 \cdot c_{12}^3 \cdot c_{22} + 1.460131216 \cdot c_{11}^2 \cdot c_{21}^2 + 1.56814504 \cdot c_{11}^3 \cdot c_{21} \\
& +2.2988306 \cdot c_{12}^4 \\
\text{s.t.} \quad & c_{11}^2 + c_{21}^2 + 2 \cdot 0.259517 \cdot c_{11} \cdot c_{21} = 1 \\
& c_{21}^2 + c_{22}^2 + 2 \cdot 0.259517 \cdot c_{12} \cdot c_{22} = 1 \\
& c_{11} \cdot c_{12} + c_{21} \cdot c_{22} + 0.259517 \cdot (c_{11} \cdot c_{22} + c_{21} \cdot c_{12}) = 0 \\
& c_{11};\ c_{12};\ c_{21};\ c_{22} \in [-2, 1]
\end{aligned}
$$

19 signomial terms and a root node relaxation of -1.141×10^2.

Edge-Concavity – A Vertex Polyhedral Envelope!

Definition (Tardella, 2003)

We say that a function f is concave along a direction $\mathbf{d}$ on a convex set S, if f is concave on all the sets:

$$S_d(\mathbf{x}) = \{\mathbf{y} \in S : \mathbf{y} = \mathbf{x} + \lambda\mathbf{d}, \lambda \in \mathbb{R}\}, \mathbf{x} \in S.$$

When P is a polyhedron and f is concave along all directions parallel to all edges of P we say that f is edge-concave on P.

Tardella (2003): Edge-concavity implies a vertex polyhedral envelope.

What if P is a box?, i.e. $P = \{\mathbf{x} \in \mathbb{R}^n : \mathbf{x}^L \leq \mathbf{x} \leq \mathbf{x}^U\}$

Let $f \in \mathcal{C}^2$ and let $H_f(\mathbf{x})$ denote the Hessian matrix of f at $\mathbf{x}$. Then f is edge-concave on P if $\mathbf{d}_i^T H_f(\mathbf{x})\mathbf{d}_i \leq 0 \ \forall\ i = 1, \ldots, k, x \in P$. In particular, when P is a box, f is componentwise concave if and only if $f_{x_i x_i}(\mathbf{x}) \leq 0\ \forall i = 1, \ldots, n$ and $x \in P$.

How is edge-concavity useful?

Generalized polynomials & fractional functions (Tardella, 2008)

Let $P = \{\mathbf{x} \in \mathrm{R}^n : \mathbf{x}^L \leq \mathbf{x} \leq \mathbf{x}^U\}$ be a box and consider a function $f : P \mapsto \mathrm{R}^n$ of the form:

$$f(\mathbf{x}) = \frac{\prod_{j=1}^{k} f_j(x_j)}{\prod_{j=k+1}^{n} f_j(x_j)},$$

and, for $i = 1, \ldots, n$, let the functions $f_{-i}(\mathbf{x})$ be obtained by setting $f_i(x_i) \equiv 1$, i.e. by deleting the factor $f_i(x_i)$.

Assume that for every $i = 1, \ldots, k$ either f_i is concave and $f_{-i}(\mathbf{x}) \geq 0$ on P, or f_i is convex and $f_{-i}(\mathbf{x}) \leq 0$ on P. Furthermore, assume that for every $i = k+1, \ldots, n$ either f_i is nonpositive convex and $f_{-i}(\mathbf{x}) \geq 0$ on P, or f_i is nonnegative concave and $f_{-i}(\mathbf{x}) \leq 0$ on P. Then $f(\mathbf{x})$ has a vertex polyhedral convex envelope on P.

Sum Decomposable

Sum Decomposable The convex envelope of a sum of functions is equal to the sum of the convex envelopes of the functions.

Tardella, 2008

Let V_p be the set of vertices on a polytope P, define edge-concave functions f, $g \mapsto \Re$ with facet representations $\{f_i : i \in I\}$, $\{g_j : j \in J\}$ defining the convex hull of f, g, respectively, and let $F_i = \{x \in P : \text{conv}_{V_P}(f)(x) = \ f_i(x)\}$, $i \in I$ and $G_j = \{x \in P : \text{conv}_{V_P}(g)(x) = g_j(x)\}$, $j \in J$ denote the linearity domains of $\text{conv}_{V_P}(f)$ and $\text{conv}_{V_P}(g)$ (*i.e.*, the sets F_i and G_i are polyhedra composed of facet-defining hyperplanes f_i and g_j). The following are equivalent:

1. $\text{conv}_{V_P}(f) + \text{conv}_{V_P}(g)$ is vertex polyhedral;
2. $\text{conv}_{V_P}(f) + \text{conv}_{V_P}(g) = \text{conv}_{V_P}(f + g)$;
3. $F_i \cap G_j$ has all vertices in V_P $\forall\ i \in I$, $j \in J$.

Almost separable

Tardella, (2008)

For the specific case of *almost separable* function $h(x, y, z) = f(x, y, z) + g(x, y, z) = \hat{f}(x, y) + \hat{g}(x, z)$ defined on $V = X \times Y \times Z$ where X, Y, Z are the vertex sets of polytopes, then the three conditions listed above are further equivalent to:

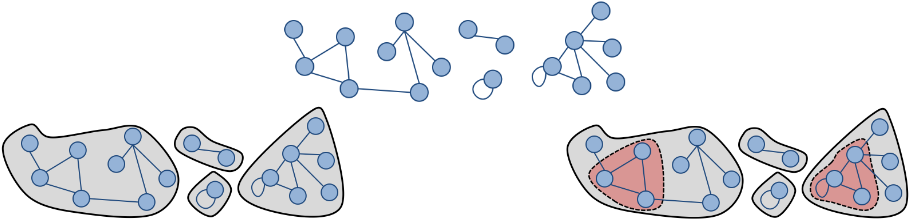
4. $F_i^X \cap G_j^X$ has all vertices in X for all linearity domains F_i of $\mathrm{conv}_{V_P}(f)$ and G_i of $\mathrm{conv}_{V_P}(g)$.

Low Dimensional Quadratic Aggregation

$$f(x_i, x_j, x_k) = \alpha_1 \cdot x_i \cdot x_i + \alpha_2 \cdot x_i \cdot x_j + \alpha_3 \cdot x_i \cdot x_k + \alpha_4 \cdot x_i + \alpha_5 \cdot x_j \cdot x_j + \alpha_6 \cdot x_j \cdot x_k + \alpha_7 \cdot x_j + \alpha_8 \cdot x_k \cdot x_k + \alpha_9 \cdot x_k$$

Where $\alpha_1, \ldots, \alpha_9$ are scalars and $\alpha_1, \alpha_5, \alpha_8$ are non-positive scalars.

What data structures to use for quadratics? (ex8_1_4)



$$\begin{array}{lll} \min & 12 \cdot x_1^2 - 6 \cdot x_1 \cdot x_2 + 6 \cdot x_2^2 - 6.3 \cdot x_3^2 + x_4^2 & \\ \text{s.t.} & -x_1 \cdot x_1 + x_3 & = 0 \\ & -x_3 \cdot x_1 + x_4 & = 0 \\ & x_1 \cdot x_4 - x_3 \cdot x_3 & = 0 \\ & x \in \mathbb{R}^C & \end{array}$$

becomes:

$$\underbrace{12 \cdot x_1^2 - 6 \cdot x_1 \cdot x_2 + 6 \cdot x_2^2}_{MT1} + \underbrace{-6.3 \cdot x_3^2}_{MT2} + \underbrace{x_4^2}_{MT3}$$

Example: 3D Edge-Concave Aggregations [1/2]

Consider:

$$f(x_i, x_j, x_k) = 0.5 \cdot x_i x_j - 0.9 \cdot x_i x_k - x_j x_k$$
$$x_i \in [-10, 0]\,;\; x_j \in [4, 10]\,;\; x_k \in [7, 10]$$

with McCormick relaxation:

$$f(x_i, x_j, x_k) \geq \begin{cases} -7 \cdot x_i - 15 \cdot x_j + 5 \cdot x_k - 30 \\ -7 \cdot x_i - 12 \cdot x_j - 1 \cdot x_k \\ -4.3 \cdot x_i - 15 \cdot x_j - 4 \cdot x_k + 60 \\ -4.3 \cdot x_i - 12 \cdot x_j - 10 \cdot x_k + 90 \\ -4 \cdot x_i - 10 \cdot x_j + 5 \cdot x_k - 50 \\ -4 \cdot x_i - 7 \cdot x_j - 1 \cdot x_k - 20 \\ -1.3 \cdot x_i - 10 \cdot x_j - 4 \cdot x_k + 40 \\ -1.3 \cdot x_i - 7 \cdot x_j - 10 \cdot x_k + 70 \end{cases}$$

Example: 3D Edge-Concave Aggregations [2/2]

Consider:

$$\begin{aligned} f(x_i, x_j, x_k) = & 0.5 \cdot x_i x_j - 0.9 \cdot x_i x_k - x_j x_k \\ & x_i \in [-10, 0]\,;\ x_j \in [4, 10]\,;\ x_k \in [7, 10] \end{aligned}$$

and facets of the convex envelope determined through the Meyer & Floudas (*Math Program*, 2005) algorithm:

$$f(x_i, x_j, x_k) \geq \begin{cases} -7 \cdot x_i - 15 \cdot x_j + 5 \cdot x_k - 30 \\ -1.3 \cdot x_i - 7 \cdot x_j - 10 \cdot x_k + 70 \\ -5.2 \cdot x_i - 12 \cdot x_j - 1 \cdot x_k + 18 \\ -3.1 \cdot x_i - 10 \cdot x_j - 4 \cdot x_k + 40 \\ -4 \cdot x_i - 10 \cdot x_j - 1 \cdot x_k + 10 \\ -4.3 \cdot x_i - 12 \cdot x_j - 4 \cdot x_k + 48 \end{cases}$$

Standard Pooling Network p-Formulation

Objective $\max_{x_{il}, y_{lj}, z_{ij}, p_{lk}} \sum_{(l,j)\in T_Y} d_j \cdot y_{lj} + \sum_{(i,j)\in T_Z} d_j \cdot z_{ij} - \sum_{(i,l)\in T_X} \gamma_i \cdot x_{il} - \sum_{(i,j)\in T_Z} \gamma_i \cdot z_{ij}$

Feed Avail $\left[A_i^L \le \sum_{l:(i,l)\in T_X} x_{il} + \sum_{j:(i,j)\in T_Z} z_{ij} \le A_i^U \quad \forall\ i \right.$

Pool Capacity $\left[S_l^L \le \sum_{i:(i,l)\in T_X} x_{il} \le S_l^U \quad \forall\ l \right.$

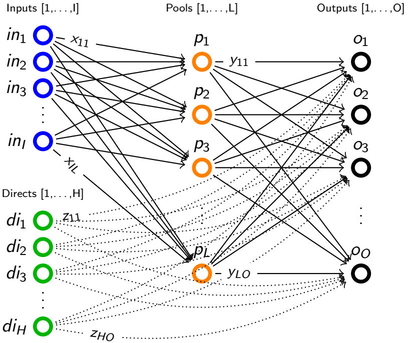
Product Demand $\left[D_j^L \le \sum_{l:(l,j)\in T_Y} y_{lj} + \sum_{i:(i,j)\in T_Z} z_{ij} \le D_j^U \quad \forall\ j \right.$

Material Balance $\left[\sum_{i:(i,l)\in T_X} x_{il} - \sum_{j:(l,j)\in T_Y} y_{lj} = 0 \quad \forall\ l \right.$

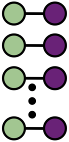
Quality Balance $\left[\sum_{i:(i,l)\in T_X} C_{ik} x_{il} = p_{lk} \sum_{j:(l,j)\in T_Y} y_{lj} \quad \forall\ l, k \right.$

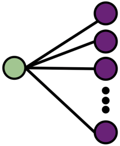
Product Quality
$$\left[\sum_{l:(l,j)\in T_Y} p_{lk} y_{lj} + \sum_{i:(i,j)\in T_Z} C_{ik} z_{ij} \right. \begin{cases} \ge P_{jk}^L \left(\sum_{l:(l,j)\in T_Y} y_{lj} + \sum_{i:(i,j)\in T_Z} z_{ij} \right) \\ \le P_{jk}^U \left(\sum_{l:(l,j)\in T_Y} y_{lj} + \sum_{i:(i,j)\in T_Z} z_{ij} \right) \end{cases} \quad \forall\ j, k$$

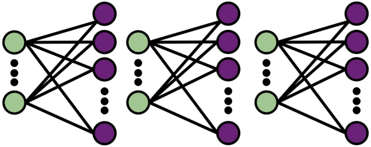
Bounds $[x_{il},\ y_{lj},\ z_{ij} \ge 0\ \forall\ i,\ l,\ j$



Process networks equations typically sum decomposable

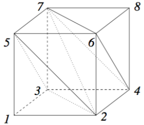
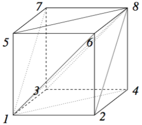
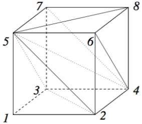
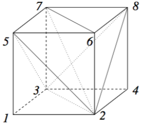
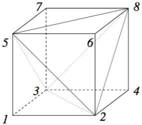
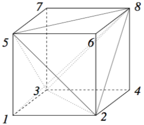






Typical process networks blending problems with intermediate nodes. Bilinear terms arise from multiplying qualities p (light green) by flowrates f (dark purple). Individual process networks equations as depicted on left have no cycles and therefore no dominant cuts. Figure on right graphs the collection of bilinear terms in a process networks problem; the graph is bipartite with one disjoint vertex set per pool-like structure.

Triangulation Types of 3-Cube (Meyer & Floudas, 2005)

Cell	Vertices	Figure	Vertices	Figure
	Type A		Type B	
1	1 2 3 5		1 2 4 8	
2	4 6 7 8		1 2 6 8	
3	2 3 5 7		1 5 6 8	
4	2 4 6 7		1 5 7 8	
5	2 5 6 7		1 3 7 8	
6	2 3 4 7		1 3 4 8	
	Type C		Type D	
1	1 2 3 5		1 2 3 5	
2	2 3 5 4		2 3 5 7	
3	2 4 5 6		2 5 6 7	
4	4 5 6 8		2 6 7 8	
5	4 5 7 8		2 3 4 8	
6	3 4 5 7		2 3 7 8	
	Type E		Type F	
1	1 2 3 5		1 2 3 8	
2	2 3 5 8		1 2 5 8	
3	2 3 4 8		1 3 5 8	
4	2 5 6 8		2 5 6 8	
5	3 5 7 8		3 5 7 8	
6			2 3 4 8	

3D Edge-Concave Aggregations: GLOBALLib Test Instances

Problem Name	# Cnt Vars	# Eqns	# Bln	Root Node Rlxn		Glob Opt.	Gap Clsd
				McC Only	McC + EC		
camshape100	200	201	198	-4.8321	-4.6583	-4.2842	0.32
camshape200	400	401	398	-4.9213	-4.8475	-4.2785	0.11
camshape400	800	801	798	-5.0645	-5.0243	-4.2757	0.06
dispatch	5	3	6	3153.30	3155.29	3155.29	1.00
st_iqpbk1	9	8	36	-1298.96	-1204.63	-621.49	0.14
st_iqpbk1	9	8	36	-2601.98	-2413.95	-1195.23	0.13

Important Principal - Balanced Matrices

Theorem (Crama, 1993): For every canonical pseudo-Boolean expression:

$$\psi(\mathbf{x}) = \sum_{T \in \Gamma} q_T \prod_{i \in T} x_i$$

where $\mathbf{x} \in [0,\, 1]^d$, $\Gamma \subseteq \{T : T \subseteq \{1,\, 2,\, \ldots,\, d\}\}$, and $q_T \neq 0 \ \forall\ T \in \Gamma$, the *standard extension* ψ^S of ψ:

$$\psi^S(\mathbf{x}) = \sum_{T \in \Gamma} q_T \begin{cases} \min\{x_i : i \in T\} & \text{if } q_T < 0 \\ \max\left\{0, 1 - |T| + \sum_{i \in T} x_i\right\} & \text{if } q_T > 0 \end{cases}$$

is equal to the unique *convex extension* ψ^C of ψ if and only if $M(\psi)$, the constraint matrix of the LP is balanced:

$$\begin{array}{lll} \min & \sum_{T \in \Gamma} q_T \cdot w_T & \\ \text{s.t.} & w_T - x_i \leq 0 & \text{for } T \in \Gamma,\, i \in T,\, q_T < 0, \\ & -w_T + \sum_{i \in T} x_i \leq |T| - 1 & \text{for } T \in \Gamma, \quad q_T > 0, \\ & 0 \leq w_T \leq 1 & \text{for } T \in \Gamma, \\ & 0 \leq x_i \leq 1 & \text{for } i \in \{1,\, 2,\, \ldots,\, d\} \end{array}$$

MIQCQP as an Undirected Graph - Balanced Matrix

Balanced Matrix Definition

A $\{0, \pm 1\}$-matrix A is *balanced* if for each submatrix B of A with exactly two nonzeros in each row and each column, the sum of all the components in B is divisible by 4.

Balanced Matrices - Consider matrices M_1 and M_2

The only submatrices of M_1 and M_2 with exactly 2 nonzeros in each row and each column are M_1 and M_2 themselves. M_1 is balanced because the sum of all the components is 0 and $0 \bmod (4) = 0$. But M_2 is *not* balanced because the sum of all the components is 6 and $6 \bmod (4) = 2$. M_2 is an unbalanced hole of length 6 (Conforti et al., 2001).

$$M_1 = \begin{pmatrix} 1 & 1 & 0 \\ 1 & 0 & -1 \\ 0 & -1 & -1 \end{pmatrix}; \quad M_2 = \begin{pmatrix} 1 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 1 \end{pmatrix}$$

How can we identify balanced matrices? [1/2]

Theorem

The triangulation representing the convex extension ψ^C of quadratic expression $\mathbf{x}^T \cdot Q_m \cdot \mathbf{x}$ for $m \in \{0, \ldots, M\}$ dominates the termwise relaxation ψ^S if and only if the triangulation representing the convex extension $\hat{\psi}^C$ of $\hat{\mathbf{x}}^T \cdot \hat{Q}_m \cdot \hat{\mathbf{x}}$ where $\hat{\mathbf{x}} \in [0, 1]^d$ and $\hat{Q}_{m,\,i,\,j} = Q_{m,\,i,\,j} \cdot (x_i^U - x_i^L) \cdot (x_j^U - x_j^L)$ dominates the standard extension $\hat{\psi}^S$.

How can we identify balanced matrices? [2/2]

Proof

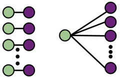
To begin, we show that $\mathbf{x}^T \cdot Q_m \cdot \mathbf{x}$ is equivalent to $\hat{\mathbf{x}}^T \cdot \hat{Q}_m \cdot \hat{\mathbf{x}}$ up to an affine shift. Using basic algebra:

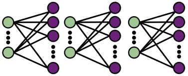
$$\begin{aligned}
\mathbf{x}^T Q_m \mathbf{x} &= \left(\mathbf{x} - \mathbf{x}^L\right)^T Q_m \left(\mathbf{x} - \mathbf{x}^L\right) + \left(\mathbf{x}^L\right)^T \left(Q_m + Q_m^T\right)\left(\mathbf{x} - \mathbf{x}^L\right) + \left(\mathbf{x}^L\right)^T Q_m \mathbf{x}^L \\
&= \hat{\mathbf{x}}^T \hat{Q}_m \hat{\mathbf{x}} + \left(\frac{\mathbf{x}^L}{\mathbf{x}^U - \mathbf{x}^L}\right)^T \left(\hat{Q}_m + \hat{Q}_m^T\right)\hat{\mathbf{x}} + \left(\frac{\mathbf{x}^L}{\mathbf{x}^U - \mathbf{x}^L}\right)^T \hat{Q}_m \frac{\mathbf{x}^L}{\mathbf{x}^U - \mathbf{x}^L} \\
&= \hat{\mathbf{x}}^T \hat{Q}_m \hat{\mathbf{x}} + \hat{a}_m \hat{\mathbf{x}} + \hat{b}_m
\end{aligned} \tag{1}$$

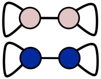
where: $\hat{\mathbf{x}} = \frac{\mathbf{x} - \mathbf{x}^L}{\mathbf{x}^U - \mathbf{x}^L} \in [0,\, 1]^d$, $\hat{a}_m = \left(\frac{\mathbf{x}^L}{\mathbf{x}^U - \mathbf{x}^L}\right)^T \cdot \left(\hat{Q}_m + \hat{Q}_m^T\right)$, and $\hat{b}_m = \left(\frac{\mathbf{x}^L}{\mathbf{x}^U - \mathbf{x}^L}\right)^T \cdot \hat{Q}_m \cdot \frac{\mathbf{x}^L}{\mathbf{x}^U - \mathbf{x}^L}$.

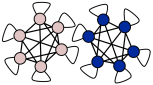
By Equation (1), $\mathbf{x}^T \cdot Q_m \cdot \mathbf{x}$ and $\hat{\mathbf{x}}^T \cdot \hat{Q}_m \cdot \hat{\mathbf{x}} + \hat{a}_m \cdot \hat{\mathbf{x}} + \hat{b}_m$ have identical triangulations. Linear and constant terms $\hat{a}_m \cdot \hat{\mathbf{x}} + \hat{b}_m$ have standard extensions equal to their convex extensions. Therefore, deviations between ψ^S and ψ^C are entirely due to a difference between $\hat{\psi}^S$ and $\hat{\psi}^C$. $\square$

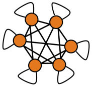
More Examples

Equation: Process Networks 

Bilinear Terms: Process Networks 

Equation: Geometry 

Terms: Geometry 

Terms: Assignment 

GloMIQO 2 analyses each equation and the collection of bilinear terms using an undirected graph representation; show the results of that analysis for several classes of MIQCQP

Complexity of generating edge-concave facets

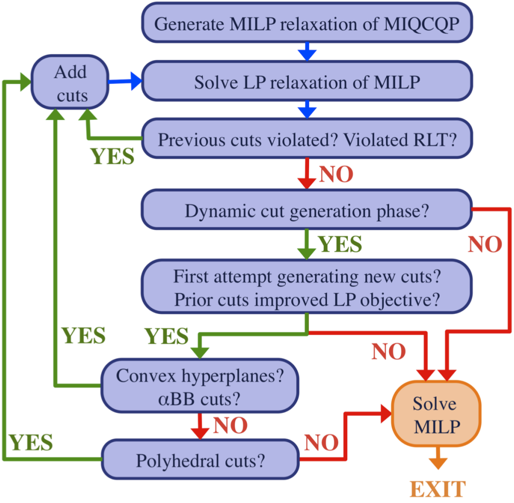
n Dim	2^n Vertices	$\binom{2^n}{n+1}$ Candidates	$n!$ Max Possible Facets
2	4	4.000×10^0	2.000×10^0
3	8	7.000×10^1	6.000×10^0
4	16	4.368×10^3	2.400×10^1
5	32	$\approx 9.062 \times 10^5$	1.200×10^2
6	64	$\approx 6.212 \times 10^8$	7.200×10^2
7	128	$\approx 1.430 \times 10^{12}$	5.040×10^3
8	256	$\approx 1.129 \times 10^{16}$	4.032×10^4
9	512	$\approx 3.123 \times 10^{20}$	$\approx 3.629 \times 10^5$
10	1024	$\approx 3.081 \times 10^{25}$	$\approx 3.629 \times 10^6$

Comparison of the GloMIQO 2 Cut Classes

Cut Class	Validity	Generation Complexity	Efficacy
RLT	Global‡	Evaluating an expression	Problem specific
Convexity	Global	Evaluating an expression	Convex hull
αBB	Global	Calculating min. eigenvalue	Tighter cut exists
Edge-Concave	Local	Generating n! facets	Convex hull

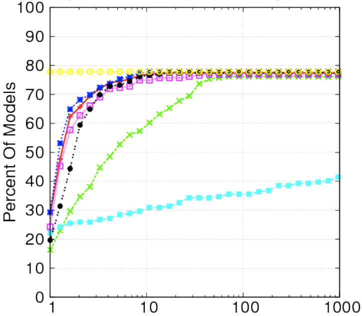
‡ If updated at each node

Idea: Let's Use Trade-Offs to our Advantage![6]

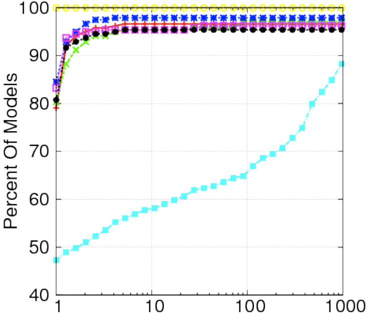


[6] Misener, Smadbeck, & Floudas, *Optim Met Softw*, 2015

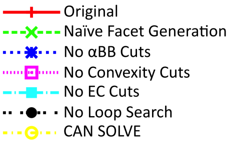
Computational Implications?



Time s: Log_2 Scale



Remaining Gap at 7200 s: Log_2 Scale



239 BoxQP/StQP/QCQP: Performance Profile illustrating the effect of knocking out each of the algorithmic components in GloMIQO 2