Linear Transformation-based Methods for Non-convex MIQPs

Eric Newby and Montaz Ali

School of Computer Science and Applied Mathematics,
& Transnet Center of System Engineering,
University of the Witwatersrand, South Africa

SANUM2016, University of Stellenbosch , March 22-24, 2016

Outline

We consider Non-convex Problem

We consider Non-convex Problem

$$
\begin{aligned}
\min_{x} \quad & h(x) = \frac{1}{2}x^T Hx + g^T x \qquad (1)\\
\text{s.t.} \quad & Ax \leq b,\\
& Dx = e,\\
& l \leq x \leq u,\\
& x = \left(x_c^T, x_d^T\right)^T \in \mathbb{R}^{n_c} \times \mathbb{Z}^{n_d},\\
& H \text{ is indefinite}
\end{aligned}
$$

The matrix H has the form

$$H = \begin{bmatrix} H_{cc} & H_{cd} \\ H_{cd}^T & H_{dd} \end{bmatrix},$$

$H_{cc} \in \mathcal{S}^{n_c}$, $H_{dd} \in \mathcal{S}^{n_d}$ and $H_{cd} \in \mathbb{R}^{(n_c, n_d)}$

The matrix H has the form

$$H = \begin{bmatrix} H_{cc} & H_{cd} \\ H_{cd}^T & H_{dd} \end{bmatrix},$$

$H_{cc} \in \mathcal{S}^{n_c}$, $H_{dd} \in \mathcal{S}^{n_d}$ and $H_{cd} \in \mathbb{R}^{(n_c,n_d)}$

1. The n_cth principal leading submatrix H_{cc} is positive definite.

The matrix H has the form

$$H = \begin{bmatrix} H_{cc} & H_{cd} \\ H_{cd}^T & H_{dd} \end{bmatrix},$$

$H_{cc} \in \mathcal{S}^{n_c}$, $H_{dd} \in \mathcal{S}^{n_d}$ and $H_{cd} \in \mathbb{R}^{(n_c, n_d)}$

1. The n_cth principal leading submatrix H_{cc} is positive definite.
2. The n_cth principal leading submatrix H_{cc} is invertible.

The matrix H has the form

$$H = \begin{bmatrix} H_{cc} & H_{cd} \\ H_{cd}^T & H_{dd} \end{bmatrix},$$

$H_{cc} \in \mathcal{S}^{n_c}$, $H_{dd} \in \mathcal{S}^{n_d}$ and $H_{cd} \in \mathbb{R}^{(n_c, n_d)}$

1. The n_cth principal leading submatrix H_{cc} is positive definite.
2. The n_cth principal leading submatrix H_{cc} is invertible.
3. The n_cth principal leading submatrix H_{cc} is singular.

We consider Linear Transformation of Cases 1 and 3

We consider Linear Transformation of Cases 1 and 3

$$V = \begin{bmatrix} U_{cc} & U_{cd} \\ 0 & U_{dd} \end{bmatrix}, \tag{2}$$

U_{cc} and U_{dd} are arbitrary invertible matrices

We consider Linear Transformation of Cases 1 and 3

$$V = \begin{bmatrix} U_{cc} & U_{cd} \\ 0 & U_{dd} \end{bmatrix}, \qquad (2)$$

U_{cc} and U_{dd} are arbitrary invertible matrices

U_{cd} is an arbitrary matrix

We consider Linear Transformation of Cases 1 and 3

$$V = \begin{bmatrix} U_{cc} & U_{cd} \\ 0 & U_{dd} \end{bmatrix}, \tag{2}$$

U_{cc} and U_{dd} are arbitrary invertible matrices

U_{cd} is an arbitrary matrix

Any matrix V with the above form is invertible

We consider Linear Transformation of Cases 1 and 3

$$V = \begin{bmatrix} U_{cc} & U_{cd} \\ 0 & U_{dd} \end{bmatrix}, \tag{2}$$

U_{cc} and U_{dd} are arbitrary invertible matrices

U_{cd} is an arbitrary matrix

Any matrix V with the above form is invertible

Let U_{dd} be the unimodular matrix

Let $x = Vy$ problem (1) is equivalent to

$$\begin{aligned}
\min_{y} \quad & h(Vy) = \frac{1}{2} y^T V^T H V y + g^T V y \\
\text{s.t.} \quad & AVy \leq b, \\
& DVy = e, \\
& l \leq Vy \leq u, \\
& y = \left[y_c^T, \, y_d^T \right]^T, \\
& U_{dd} y_d \in \mathbb{Z}^{n_d}, \\
& U_{cc} y_c + U_{cd} y_d \in \mathbb{R}^{n_c}.
\end{aligned} \tag{3}$$

$U_{dd} y_d \in \mathbb{Z}^{n_d}$; we therefore restrict U_{dd} to be some unimodular matrix.

$U_{dd} y_d \in \mathbb{Z}^{n_d}$; we therefore restrict U_{dd} to be some unimodular matrix.

$$U_{dd} y_d \in \mathbb{Z}^{n_d} \Leftrightarrow y_d \in \mathbb{Z}^{n_d}.$$

$U_{dd} y_d \in \mathbb{Z}^{n_d}$; we therefore restrict U_{dd} to be some unimodular matrix.

$$U_{dd} y_d \in \mathbb{Z}^{n_d} \Leftrightarrow y_d \in \mathbb{Z}^{n_d}.$$

$$U_{cc} y_c + U_{cd} y_d \in \mathbb{R}^{n_c} \Leftrightarrow y_c \in \mathbb{R}^{n_c}.$$

$U_{dd}y_d \in \mathbb{Z}^{n_d}$; we therefore restrict U_{dd} to be some unimodular matrix.

$$U_{dd}y_d \in \mathbb{Z}^{n_d} \Leftrightarrow y_d \in \mathbb{Z}^{n_d}.$$

$$U_{cc}y_c + U_{cd}y_d \in \mathbb{R}^{n_c} \Leftrightarrow y_c \in \mathbb{R}^{n_c}.$$

Problem (3) now takes the following form:

$$\begin{aligned} \min_{y} \quad & h(Vy) = \frac{1}{2}y^T V^T H V y + g^T V y \\ \text{s.t.} \quad & AVy \leq b, \\ & DVy = e, \\ & l \leq Vy \leq u, \\ & y = \left[y_c^T, \, y_d^T\right]^T \in \mathbb{R}^{n_c} \times \mathbb{Z}^{n_d}. \end{aligned}$$

$$\begin{aligned}
\min_{y} \quad & h(Vy) = \frac{1}{2} y^T V^T H V y + g^T V y \\
\text{s.t.} \quad & AVy \le b, \\
& DVy = e, \\
& l \le Vy \le u, \\
& y = \left[y_c^T, \, y_d^T \right]^T \in \mathbb{R}^{n_c} \times \mathbb{Z}^{n_d}.
\end{aligned} \tag{4}$$

$$\begin{aligned}
y^T V^T H V y = & y_c^T U_{cc}^T H_{cc} U_{cc} y_c + 2 y_d^T \left(U_{cd}^T H_{cc} U_{cc} + U_{dd}^T H_{cd}^T U_{cc} \right) y_c \\
& + y_d^T \left(U_{cd}^T H_{cc} U_{cd} + U_{cd}^T H_{cd} U_{dd} + U_{dd}^T H_{cd}^T U_{cd} \right. \\
& \left. + U_{dd}^T H_{dd} U_{dd} \right) y_d .
\end{aligned} \tag{5}$$

Lower bounding problem at each B&B tree under estimates each of bilinear term (one variable and two constraints)

Lower bounding problem at each B&B tree under estimates each of bilinear term (one variable and two constraints)

$$
\begin{aligned}
y^T V^T H V y = & y_c^T U_{cc}^T H_{cc} U_{cc} y_c + 2 y_d^T \left(U_{cd}^T H_{cc} U_{cc} + U_{dd}^T H_{cd}^T U_{cc} \right) y_c \\
& + y_d^T \left(U_{cd}^T H_{cc} U_{cd} + U_{cd}^T H_{cd} U_{dd} + U_{dd}^T H_{cd}^T U_{cd} \right. \\
& \left. + U_{dd}^T H_{dd} U_{dd} \right) y_d .
\end{aligned}
$$

We set $\left(U_{cd}^T H_{cc} U_{cc} + U_{dd}^T H_{cd}^T U_{cc}\right) = 0$ **and see if** U_{cc} **and** U_{dd} **have desired properties**

We set $\left(U_{cd}^T H_{cc} U_{cc} + U_{dd}^T H_{cd}^T U_{cc}\right) = 0$ **and see if** U_{cc} **and** U_{dd} **have desired properties**

We know that U_{cc} is invertible so we have

$$H_{cc} U_{cd} = -H_{cd} U_{dd}. \tag{6}$$

We set $\left(U_{cd}^T H_{cc} U_{cc} + U_{dd}^T H_{cd}^T U_{cc}\right) = 0$ **and see if** U_{cc} **and** U_{dd} **have desired properties**

We know that U_{cc} is invertible so we have

$$H_{cc} U_{cd} = -H_{cd} U_{dd}. \tag{6}$$

$$U_{cd} = -H_{cc}^{-1} H_{cd} U_{dd}.$$

Calculation of U_{dd}

$$\underset{(U_{dd})_i}{\operatorname{argmin}} \left\{ \max_{x_d} \left[\left(U_{dd}^{-1} \right)_i x_d : x \in \Omega_q \right] - \min_{x_d} \left[\left(U_{dd}^{-1} \right)_i x_d : x \in \Omega_q \right] \right\} \quad (7)$$

$$\begin{aligned} \text{s.t.} \quad & \left(U_{dd}^{-1} \right)_{i,i} = \pm 1, \\ & \left(U_{dd}^{-1} \right)_{i,j} = 0, \quad j = 1, \ldots, i-1, \\ & \left(U_{dd}^{-1} \right)_{i,j} \in \mathbb{Z}, \quad j = i+1, \ldots, n_d. \end{aligned}$$

Calculation of U_{dd}

$$\underset{(U_{dd})_i}{\operatorname{argmin}} \left\{ \max_{x_d} \left[\left(U_{dd}^{-1}\right)_i x_d : x \in \Omega_q \right] - \min_{x_d} \left[\left(U_{dd}^{-1}\right)_i x_d : x \in \Omega_q \right] \right\} \quad (7)$$

$$\begin{aligned} \text{s.t.} \quad & \left(U_{dd}^{-1}\right)_{i,i} = \pm 1, \\ & \left(U_{dd}^{-1}\right)_{i,j} = 0, \quad j = 1, \ldots, i-1, \\ & \left(U_{dd}^{-1}\right)_{i,j} \in \mathbb{Z}, \quad j = i+1, \ldots, n_d. \end{aligned}$$

Calculation of U_{cc}

- H_{cc} is Hermitian it is diagonalisable. Let U_{cc} be the diagonalising matrix of H_{cc}.
- The columns of U_{cc} are the normalizing eigenvectors of H_{cc}.

$$
\begin{aligned}
y^TV^THVy = &y_c^TU_{cc}^TH_{cc}U_{cc}y_c + 2y_d^T\left(U_{cd}^TH_{cc}U_{cc} + U_{dd}^TH_{cd}^TU_{cc}\right)y_c \\
&+ y_d^T\left(U_{cd}^TH_{cc}U_{cd} + U_{cd}^TH_{cd}U_{dd} + U_{dd}^TH_{cd}^TU_{cd}\right. \\
&\left.+U_{dd}^TH_{dd}U_{dd}\right)y_d.
\end{aligned}
$$

$$
\begin{aligned}
y^T V^T H V y = & y_c^T U_{cc}^T H_{cc} U_{cc} y_c + 2 y_d^T \left(U_{cd}^T H_{cc} U_{cc} + U_{dd}^T H_{cd}^T U_{cc} \right) y_c \\
& + y_d^T \left(U_{cd}^T H_{cc} U_{cd} + U_{cd}^T H_{cd} U_{dd} + U_{dd}^T H_{cd}^T U_{cd} \right. \\
& \left. + U_{dd}^T H_{dd} U_{dd} \right) y_d .
\end{aligned}
$$

$$
\Theta_{dd} = U_{cd}^T H_{cc} U_{cd} + U_{cd}^T H_{cd} U_{dd} + U_{dd}^T H_{cd}^T U_{cd} + U_{dd}^T H_{dd} U_{dd} .
$$

$$
\begin{aligned}
y^T V^T H V y = & y_c^T U_{cc}^T H_{cc} U_{cc} y_c + 2 y_d^T \left(U_{cd}^T H_{cc} U_{cc} + U_{dd}^T H_{cd}^T U_{cc} \right) y_c \\
& + y_d^T \left(U_{cd}^T H_{cc} U_{cd} + U_{cd}^T H_{cd} U_{dd} + U_{dd}^T H_{cd}^T U_{cd} \right. \\
& \left. + U_{dd}^T H_{dd} U_{dd} \right) y_d .
\end{aligned}
$$

$$
\Theta_{dd} = U_{cd}^T H_{cc} U_{cd} + U_{cd}^T H_{cd} U_{dd} + U_{dd}^T H_{cd}^T U_{cd} + U_{dd}^T H_{dd} U_{dd} .
$$

$$
U_{cd} = -H_{cc}^{-1} H_{cd} U_{dd} .
$$

$$y^T V^T H V y = y_c^T U_{cc}^T H_{cc} U_{cc} y_c + 2 y_d^T \left(U_{cd}^T H_{cc} U_{cc} + U_{dd}^T H_{cd}^T U_{cc} \right) y_c + y_d^T \left(U_{cd}^T H_{cc} U_{cd} + U_{cd}^T H_{cd} U_{dd} + U_{dd}^T H_{cd}^T U_{cd} + U_{dd}^T H_{dd} U_{dd} \right) y_d.$$

$$\Theta_{dd} = U_{cd}^T H_{cc} U_{cd} + U_{cd}^T H_{cd} U_{dd} + U_{dd}^T H_{cd}^T U_{cd} + U_{dd}^T H_{dd} U_{dd}.$$

$$U_{cd} = -H_{cc}^{-1} H_{cd} U_{dd}.$$

$$\Theta_{dd} = U_{dd}^T \left(H_{dd} - H_{cd}^T H_{cc}^{-1} H_{cd} \right) U_{dd}. \tag{8}$$

The Transformed Problem

$$
\begin{aligned}
\min_{y} \quad & h(Vy) = \frac{1}{2}\left(y_c^T \Theta_{cc} y_c + y_d^T \Theta_{dd} y_d\right) + g^T V y \\
\text{s.t.} \quad & AVy \leq b, \\
& DVy = e, \\
& l \leq Vy \leq u, \\
& y^L \leq y \leq y^U,
\end{aligned}
\tag{9}
$$

$$
\begin{aligned}
\min_{x} \quad & h(x) = \frac{1}{2} x^T H x + g^T x \\
\text{s.t.} \quad & Ax \leq b, \\
& Dx = e, \\
& l \leq x \leq u,
\end{aligned}
$$

$$\begin{aligned} \min_{x} \quad & h(Vy) = \frac{1}{2}\left(y_c^T \Theta_{cc} y_c + y_d^T \Theta_{dd} y_d\right) + g^T Vy \\ \text{s.t.} \quad & AVy \leq b,\ DVy = e,\ l \leq Vy \leq u,\ U_{dd} y_d = z \\ & y = \left[y_c^T, y_d^T\right]^T \in \mathbb{R}^{n_c} \times \mathbb{R}^{n_d},\ z \in \mathbb{Z}^{n_d}. \end{aligned} \tag{10}$$

$$\Theta_{dd} = U_{dd}^T \left(H_{dd} - H_{cd}^T H_{cc}^{-1} H_{cd}\right) U_{dd}$$

Calculation of U_{dd}

$$\operatorname*{argmin}_{(U_{dd})_i} \left\{ \max_{x_d} \left[\left(U_{dd}^{-1}\right)_i x_d : x \in \Omega_q \right] - \min_{x_d} \left[\left(U_{dd}^{-1}\right)_i x_d : x \in \Omega_q \right] \right\} \tag{11}$$

$$\begin{aligned}
\text{s.t.} \quad & \left(U_{dd}^{-1}\right)_{i,i} = \pm 1, \\
& \left(U_{dd}^{-1}\right)_{i,j} = 0, \quad j = 1, \ldots, i-1, \\
& \left(U_{dd}^{-1}\right)_{i,j} \in \mathbb{Z}, \quad j = i+1, \ldots, n_d.
\end{aligned}$$

Consider the convexification of the following non-convex MIQP

$$\begin{aligned}
\min_{x} \quad & h(x) = \frac{1}{2}x^T Hx + g^T x \\
\text{s.t.} \quad & Ax \leq b, \\
& Dx = e, \\
& l \leq x \leq u, \\
& H_{cc} \quad \textcolor{red}{\text{Positive Definite}}
\end{aligned}$$

Consider the convexification of the following non-convex MIQP

$$\begin{aligned} \min_x \quad & h(x) = \frac{1}{2}x^T Hx + g^T x \\ \text{s.t.} \quad & Ax \leq b, \\ & Dx = e, \\ & l \leq x \leq u, \\ & H_{cc} \quad \text{Positive Definite} \end{aligned}$$

Billionnet et al.(2012), Mathematical Programming 131, 381–401

Denote Convexification of above Problem as the Mixed Integer Quadratic Convex Reformulation (MIQCR)

Consider the Convexification of the following non-convex MIQP, H_{cc} Positive Definite

$$
\begin{aligned}
\min_{y} \quad & h(Vy) = \frac{1}{2}\left(y_c^T \Theta_{cc} y_c + y_d^T \Theta_{dd} y_d\right) + g^T Vy \\
\text{s.t.} \quad & AVy \leq b, \\
& DVy = e, \\
& l \leq Vy \leq u, \\
& y^L \leq y \leq y^U,
\end{aligned}
$$

Denote Convexification of above Problem as the Mixed Integer Quadratic Transformation and Convex Reformulation (MIQTCR)

Pörn et al, Comput. Chem. Eng (1999)
The non-convex terms of the transformed problem are bilinear terms involving only the integer variables.

$$V = \begin{bmatrix} U_{cc} & U_{cd} \\ 0_{n_d,n_c} & \widetilde{U}_{dd} U_{dd} \end{bmatrix}.$$

$$\begin{aligned} \min_{y} \quad & h(Vy) = \frac{1}{2}\left(y_c^T \Theta_{cc} y_c + y_d^T \Theta_{dd} y_d\right) + g^T V y \\ \text{s.t.} \quad & AVy \le b, \\ & DVy = e, \\ & l \le Vy \le u, \\ & y^L \le y \le y^U \end{aligned} \tag{12}$$

Convex Reformulation by Pörn et al (1999):

Convex Reformulation by Pörn et al (1999):

Applied to Our Transformed Problem (MIQTBC)

Convexification Results in a Convex MINLP – Not a Convex MIQP

Results obtained using MINLP solver: Couenne 0.3.2 on the NEOS server

Type 1. Bound constraints: $-2 \leq x_i \leq 2$

Type 1. Bound constraints: $-2 \leq x_i \leq 2$

Type 2. Sparse linear inequality constraints: matrix A had sparse block diagonal structure

Type 1. Bound constraints: $-2 \leq x_i \leq 2$

Type 2. Sparse linear inequality constraints: matrix A had sparse block diagonal structure

Type 3. Dense linear inequality constraints: matrix A was dense.

Type 1. Bound constraints: $-2 \leq x_i \leq 2$

Type 2. Sparse linear inequality constraints: matrix A had sparse block diagonal structure

Type 3. Dense linear inequality constraints: matrix A was dense.

- MIQCR (Convex MIQP)
- MIQTCR (Convex MIQP)
- MIQTBC (Convex MINLP)

- Solver: Couenne 0.3.2 on the NEOS server

n	MIQCR	MIQTCR	MIQTBC
4	5.412	4.313	1.330
6	42.082	20.522	6.456
8	47.235	49.611	19.410
10	110.43	192.12	151.96
12	301.37	451.29	475.54
14	1032.1	1688.3	2012.5

Table: The time taken to solve problems using Couenne for Constraints Type 2

n	MIQCR	MIQTCR	MIQTBC
4	3.094	1.714	0.657
6	15.83	10.15	12.45
8	99.32	255.03	68.34
10	5352.3	3687.3	1958.6

Table: The time taken to solve problems using Couenne for Constraints Type 3

- We have developed a B&B algorithm for solving this type of MIQPs
- Reduce Bilinear Terms in the during Linear Transformation

$$
\begin{aligned}
y^T V^T H V y = & y_c^T U_{cc}^T H_{cc} U_{cc} y_c + 2 y_d^T \left(U_{cd}^T H_{cc} U_{cc} + U_{dd}^T H_{cd}^T U_{cc} \right) y_c \\
& + y_d^T \left(U_{cd}^T H_{cc} U_{cd} + U_{cd}^T H_{cd} U_{dd} + U_{dd}^T H_{cd}^T U_{cd} \right. \\
& \left. + U_{dd}^T H_{dd} U_{dd} \right) y_d .
\end{aligned}
$$

Transformation uses the following form of Hessian

$$\Theta = \Theta^{(1)} + \Theta^{(2)},$$
$$\Theta = \begin{bmatrix} \Theta_{cc}^{(1)} & 0 \\ 0 & \Theta_{dd}^{(1)} \end{bmatrix} + \begin{bmatrix} \Theta_{cc}^{(2)} & \Theta_{cd}^{(2)} \\ \Theta_{cd}^{(2)T} & \Theta_{dd}^{(2)} \end{bmatrix} \tag{13}$$

$\Theta_{cc}^{(1)}$, $\Theta_{cc}^{(2)}$, $\Theta_{dd}^{(2)}$ are diagonal; $\Theta^{(2)}$ is PD.

Theorem

$\exists$ U_{cc} such that U_{cc} diagonalises H_{cc} and Θ can be written in the following form

$$\Theta = \begin{bmatrix} \Theta_{cc}^{(1)} & 0 \\ 0 & \Theta_{dd}^{(1)} \end{bmatrix} + \Theta^{(2)}, \tag{14}$$

where $\Theta^{(2)}$ is positive definite.

$$
\begin{aligned}
y^T V^T H V y = & y_c^T U_{cc}^T H_{cc} U_{cc} y_c + 2 y_d^T \left(U_{cd}^T H_{cc} U_{cc} + U_{dd}^T H_{cd}^T U_{cc} \right) y_c \\
& + y_d^T \left(U_{cd}^T H_{cc} U_{cd} + U_{cd}^T H_{cd} U_{dd} + U_{dd}^T H_{cd}^T U_{cd} \right. \\
& \left. + U_{dd}^T H_{dd} U_{dd} \right) y_d .
\end{aligned}
$$

Choose V such that the Hessian $y^T V^T H V y$ $(=\Theta)$ is

$$
\begin{aligned}
\Theta &= \Theta^{(1)} + \Theta^{(2)}, \\
\Theta &= \begin{bmatrix} \Theta_{cc}^{(1)} & 0 \\ 0 & \Theta_{dd}^{(1)} \end{bmatrix} + \begin{bmatrix} \Theta_{cc}^{(2)} & \Theta_{cd}^{(2)} \\ \Theta_{cd}^{(2)T} & \Theta_{dd}^{(2)} \end{bmatrix} \qquad (15)
\end{aligned}
$$

- $U_{cc}^T(H_{cc}U_{cd} + H_{cd}U_{dd})$ must must be small to make $\Theta^{(2)}$ PD.
- $U_{dd} = I_{n_d}$
- Find U_{cc} that diagonalize H_{cc} by setting U_{cd}=0.
- An algorithm for calculating U_{cc} is given

Figure: Performance profile when H_{cc} Singular using B&B for $n_c = n_d$.

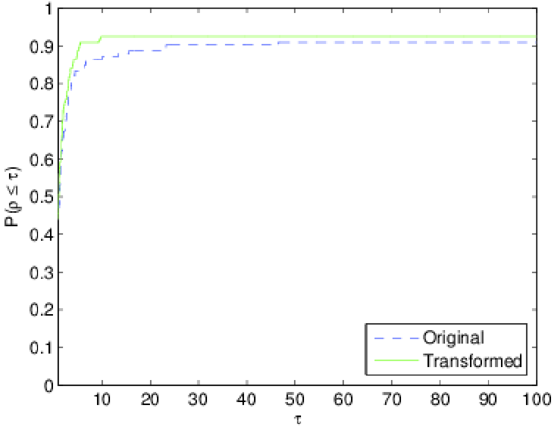


Figure: Performance profile when H_{cc} Singular using B&B for $n_c > n_d$.

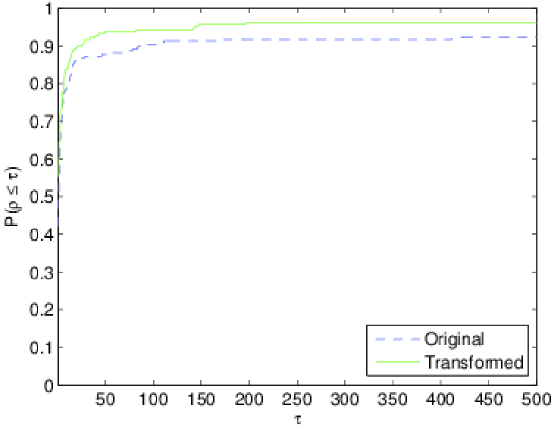
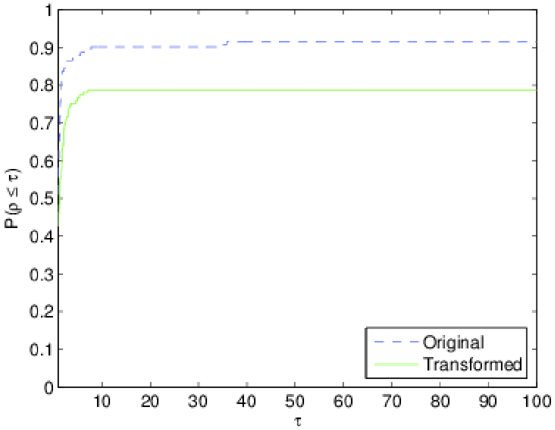


Figure: Performance profile when H_{cc} Singular using B&B for $n_c < n_d$.



Thank You!