Enhancing Computational Techniques for Stochastic Linear Programs

Udom Janjarassuk [1], Jeff Linderoth [2]

INFORMS Annual Meeting, 08

October 2, 2008

[1]Lehigh University – udj2@lehigh.edu
[2]University of Wisconsin-Madison – linderot@cs.wisc.edu

Introduction

Focus on solving large-scale two stages stochastic linear programs

- Use L-shaped decomposition method with trust-region enhanced
- Utilize computational grid for parallel computation

Main discussion of the talk

- Warm start for decomposition

Motivation

- Small SP is very easy, large SP is hard
- Use information from small problems to solve large problem

Outline

Two-stage Stochastic Linear Programs

$$\begin{aligned} z^* = \min_{x \geq 0} \quad \{f(x) \quad &:= \quad c^T x + \mathcal{Q}(x, \xi)\} \\ s.t. \qquad Ax \quad &= \quad b \end{aligned} \tag{1a}$$

where $\mathcal{Q}(x,\xi) = \mathbb{E}_\xi[Q(x,\xi)]$ and $Q(x,\xi)$ is the value of the optimal solution of the second-stage recourse problem

$$\begin{aligned} Q(x,\xi) = \quad \min_{y \geq 0} \qquad & q(\omega)^T y \\ s.t. \quad T(\omega)x + W(\omega)y \quad &= \quad h(\omega) \end{aligned} \tag{1b}$$

where ξ is a random vector and ω is a random event.

Two-stage SP - Extensive Form

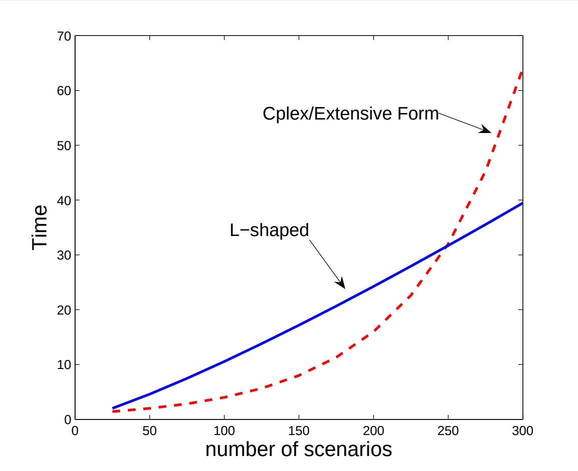
$$
\begin{array}{llrcll}
min & c^T x + \sum_{k=1}^{K} p_k q_k^T y_k & & & & \\
s.t. & & Ax & = & b & \\
 & & T_k x + W y_k & = & h_k, & \forall k = 1,..,K \\
 & & x & \geq & 0 & \\
 & & y_k & \geq & 0, & \forall k = 1,..,K
\end{array} \quad (2)
$$

where

- K is the total number of possible scenarios
- p_k is the probability associated with scenario k

We refer this as the *Deterministic Equivalent Problem.*
This formulation can be solved using LP solver.

Deterministic Equivalent vs Decomposition



Warm Start

Key Ideas

- To provide a good starting point
- To obtain cuts in order to tighten the lower bound

Scenario Partitioning

Given a large SP with K scenarios

- Partition the set of scenario into P subsets, each of size $N_p, p = 1..P$
- Form the DE problems using scenarios from each subset
- Solve each DE problem using LP solver
- Obtain solution and generate optimality cuts from each DE problem
- Use the average solution as a starting point to solve the original problem
- Modify the cuts from each DE problem according to its probability in order to fit the original problem
- Cut aggregation can also be done if necessary

Generating Cuts from Warm Start

In L-shaped method, optimality cut is generated by

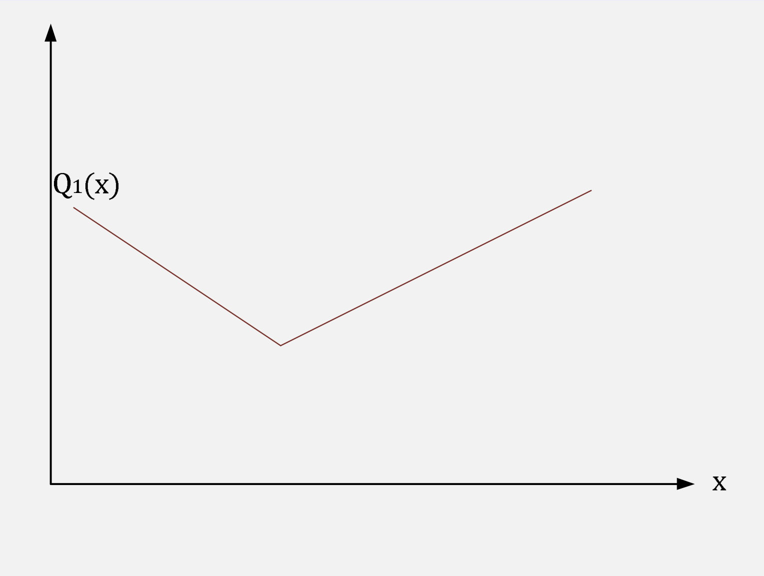
$$\theta \geq \sum_{k=1}^{K} \left[p_k \pi_k^T (h_k - T_k x) \right].$$

- where π_k is the optimal dual multiplier associated with scenario k

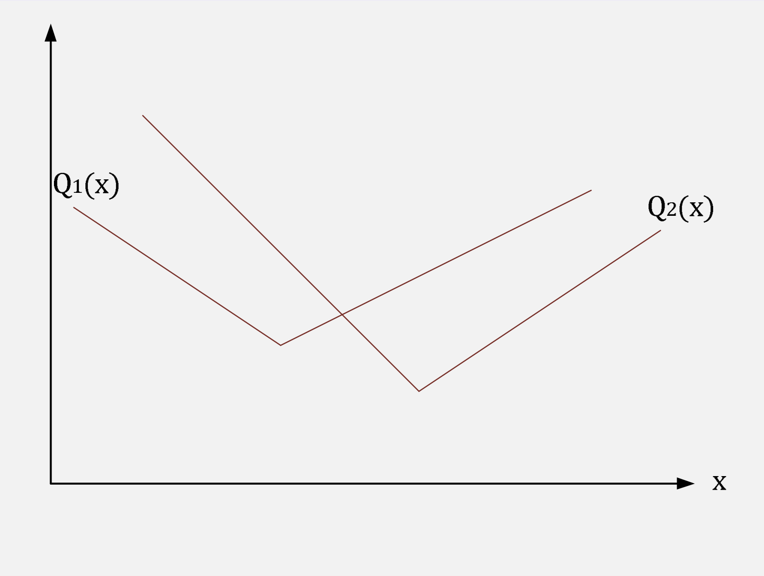
In the Multicut L-shaped version, we have

$$\theta_k \geq p_k \pi_k^T (h_k - T_k x) \quad \forall k \in 1, ..., K.$$

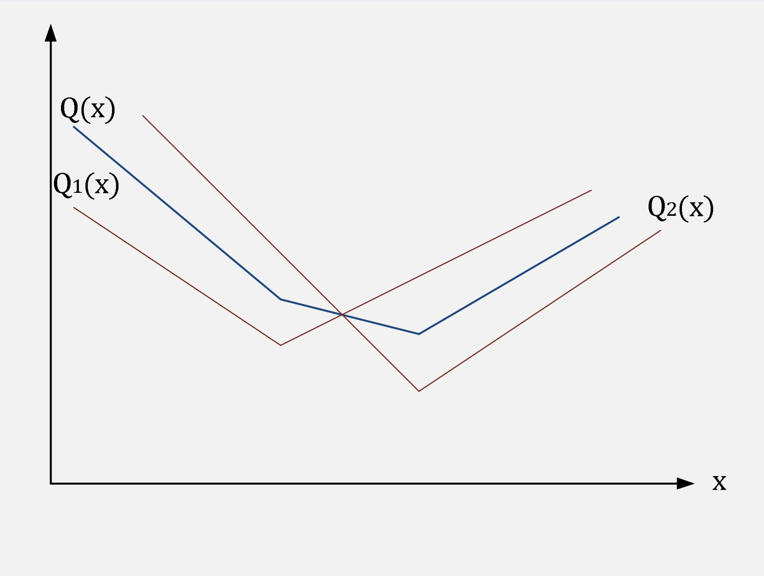
Generating Cuts from Warm Start



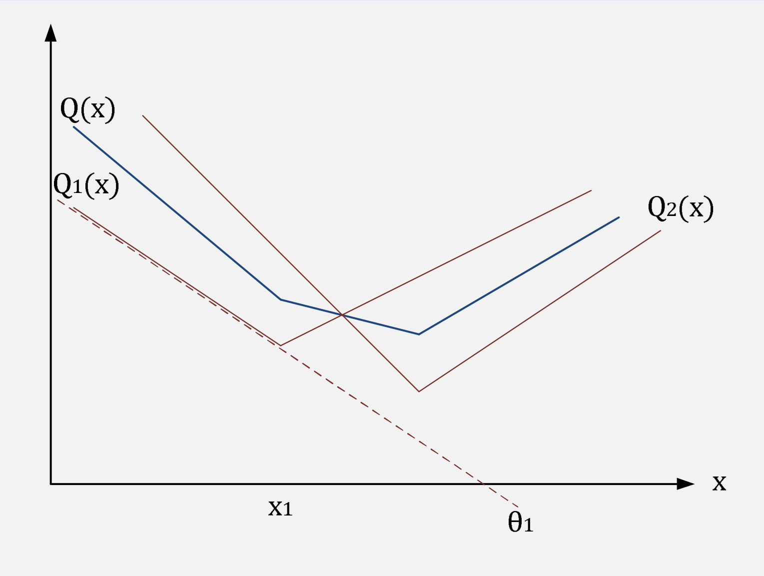
Generating Cuts from Warm Start(Cont'd)



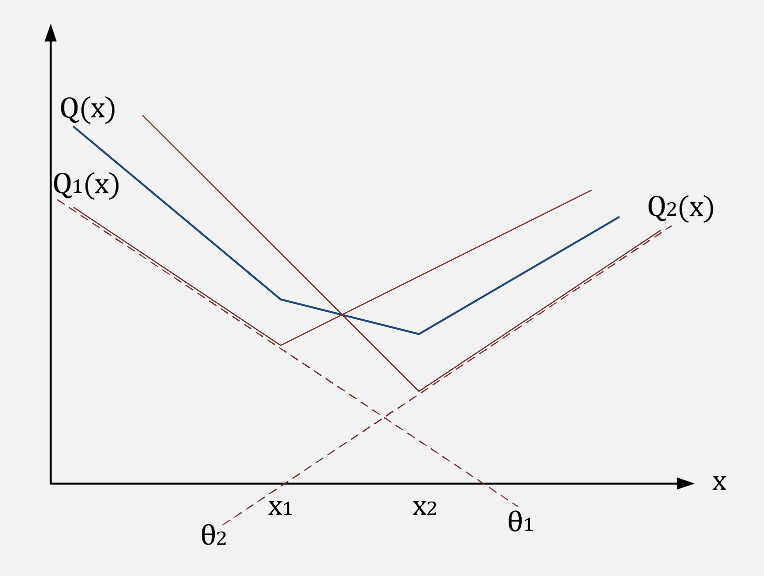
Generating Cuts from Warm Start(Cont'd)



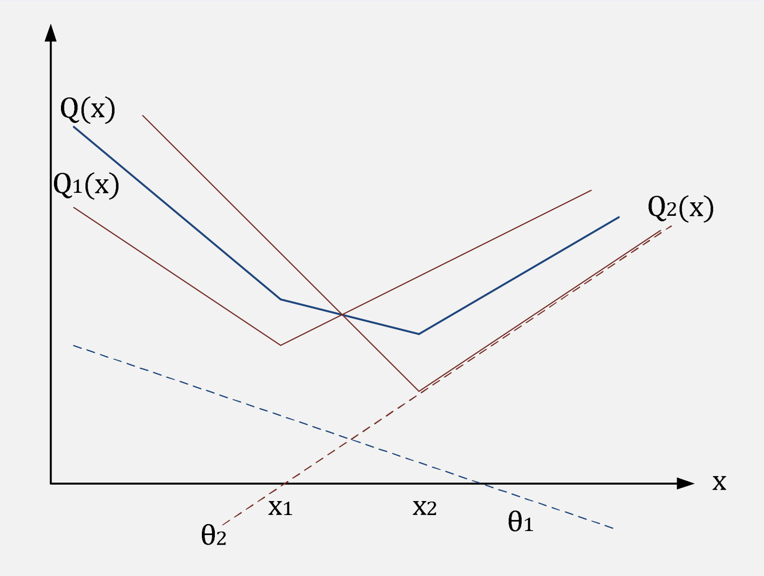
Generating Cuts from Warm Start(Cont'd)



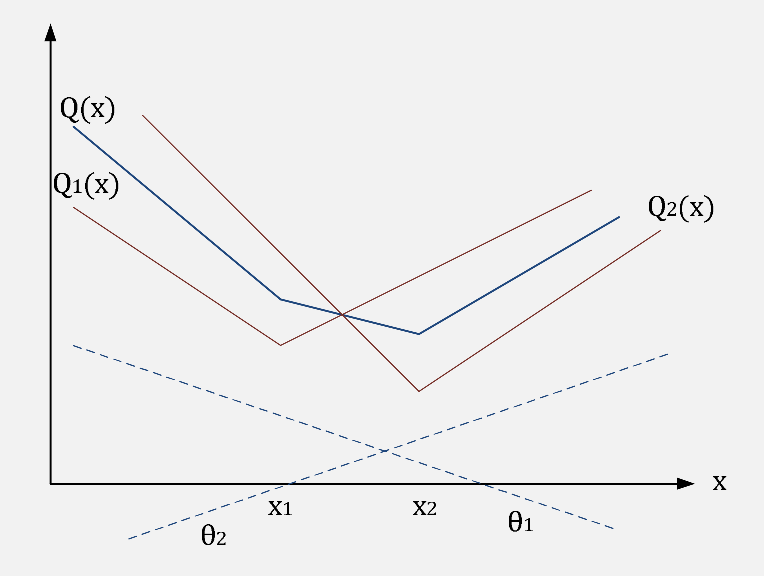
Generating Cuts from Warm Start(Cont'd)



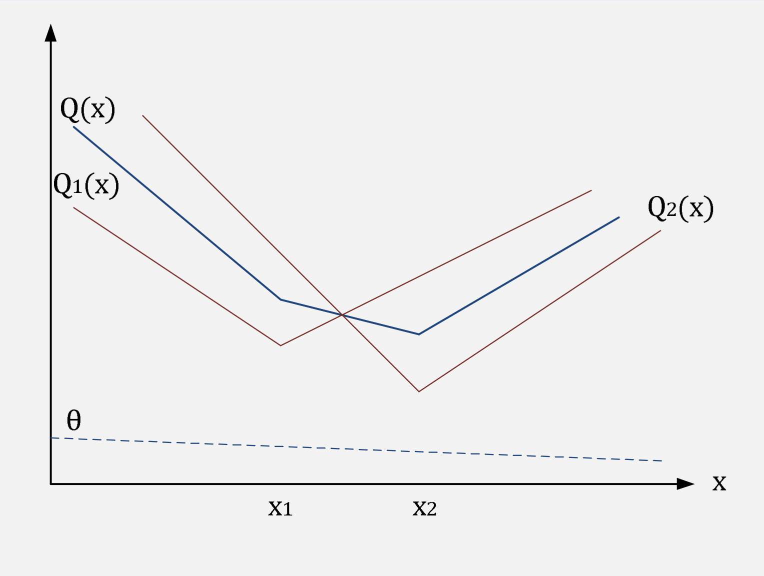
Generating Cuts from Warm Start(Cont'd)



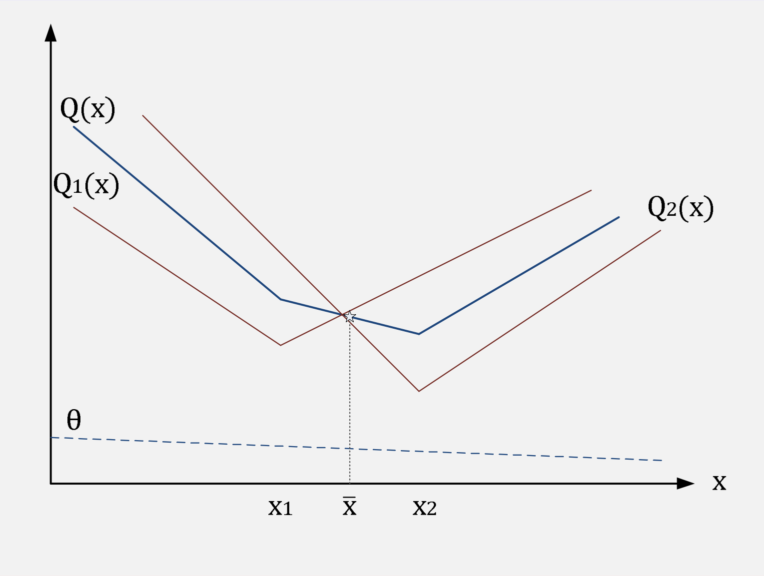
Generating Cuts from Warm Start(Cont'd)



Generating Cuts from Warm Start(Cont'd)



Generating Cuts from Warm Start(Cont'd)



Computational Results

Setting:

- Test on 19 problems from literatures
- Use sampling technique to generate large sample problems, sample size vary from 1,000 to 20,000 based on difficulty
- Each sample problem is solved by using decomposition method on computational grid with 54 processors
- Warm start using scenario partitioning with P = 200
- Cuts are aggregated within each subset
- Results are based on the average of 10 trials

Qualities of Solutions

Instance	N	HqOpt%	AvgOpt%	1stOpt%
20term	10000	0.04	0.00	0.03
4node_32768	10000	0.00	1.08	2.67
AIRL2	20000	0.00	0.02	0.03
assets_sm	20000	0.00	0.00	0.00
biweekly_lg	10000	0.00	0.00	0.00
electric_lg	10000	0.00	0.00	0.00
gbd	20000	0.00	0.06	0.23
LandS	20000	0.00	0.00	0.03
PGP1	20000	7.70	0.14	0.84
phone	20000	0.00	0.00	0.00
product_sm	10000	0.00	0.00	0.00
semi4	1000	0.24	177.83	259.06
snip4x9	10000	0.11	1.88	3.33
snip7x5	10000	0.14	1.33	2.06
ssn	5000	2.50	19.09	76.61
stocfor2	20000	0.00	0.00	0.07
storm	10000	0.00	0.00	0.00
weekly_lg	2000	0.00	0.00	0.00
weekly_md	2000	0.00	0.00	0.00

Table: Average percentage from optimal by evaluating different solutions

Optimality Gaps from Warm Start

Instance	N	UB	LB	OptGap%
20term	10000	254370.3	254290.4	0.031
4node_32768	10000	451.7	446.7	1.113
AIRL2	20000	269680.5	269569.4	0.041
assets_sm	20000	-723.9	-723.9	0.000
biweekly_lg	10000	-4211.8	-4213.8	0.047
electric_lg	10000	-7539.1	-7539.1	0.000
gbd	20000	1654.2	1651.3	0.174
LandS	20000	225.7	225.6	0.027
PGP1	20000	439.6	436.7	0.664
phone	20000	36.9	36.9	0.000
product_sm	10000	-34165.9	-34165.9	0.000
semi4	1000	314.5	90.9	197.568
snip4x9	10000	10.8	9.8	9.412
snip7x5	10000	81.3	77.5	4.739
ssn	5000	11.5	1.9	99.501
stocfor2	20000	-39772.2	-39806.4	0.086
storm	10000	15498030.0	15497399.0	0.004
weekly_lg	2000	-12502.5	-12502.5	0.000
weekly_md	2000	-6149.4	-6149.4	0.000

Table: Average optimality gaps after warm starting

Instance	N	Time w.o.WS	Time w. WS	WS Time(each)
20term	10000	1628.3	403.6	0.83
4node_32768	10000	273.8	137.9	0.83
AIRL2	20000	129.4	84.9	0.02
assets_sm	20000	96.6	37.3	0.04
biweekly_lg	10000	299.6	59.8	1.59
electric_lg	10000	771.6	106.5	0.31
gbd	20000	188.8	144.3	0.02
LandS	20000	138.8	97.0	0.01
PGP1	20000	174.1	149.2	0.02
phone	20000	149.5	108.4	0.23
product_sm	10000	716.5	127.8	0.32
semi4	1000	1863.3	1088.4	1.79
snip4x9	10000	445.4	184.6	0.29
snip7x5	10000	267.9	147.7	0.40
ssn	5000	606.2	253.1	0.37
stocfor2	20000	430.2	94.1	0.36
storm	10000	486.7	207.4	3.16
weekly_lg	2000	1083.2	39.1	0.36
weekly_md	2000	493.2	71.1	0.18

Table: Decomposition time and warm starting time in second

Performance Profiles

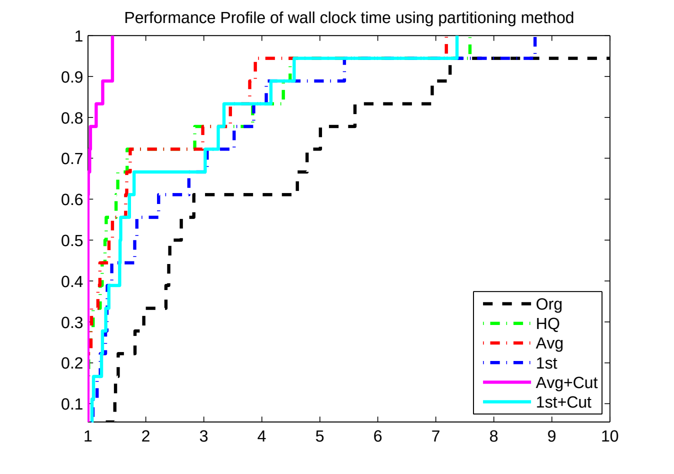


Figure: Performance profile of wall clock time

Performance Profiles

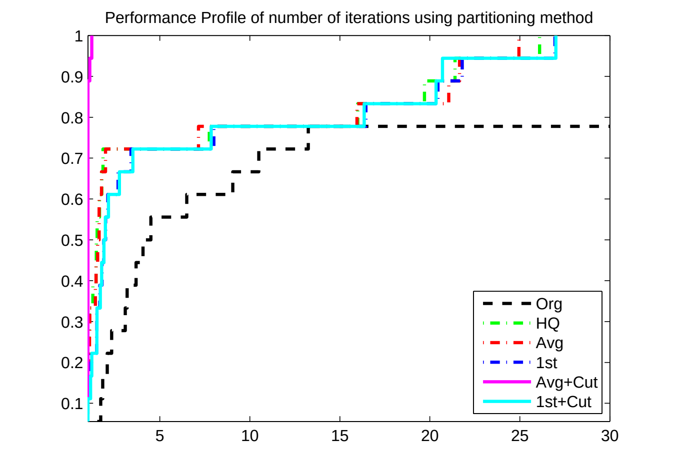


Figure: Performance profile of number of iterations

Varying the Size of Partitions

- Test only 3 problems that have large optimality gaps
- Size of partition vary based on problem difficulties
- Use the average solution to obtain the upper bound
- Solve the master problem with cuts from warm start to obtain the lower bound

Vary Partition Size – ssn

Partition Size	DE Time	Solve Time	Iteration
25	0.38	351.33	47.1
50	0.93	389.95	44.9
100	2.65	303.89	42.8
250	10.01	324.86	41.4
500	32.74	281.05	42.7

Table: Average solving time and number of iterations for ssn with N = 5,000

Upper and Lower Bounds after Warm Start – ssn

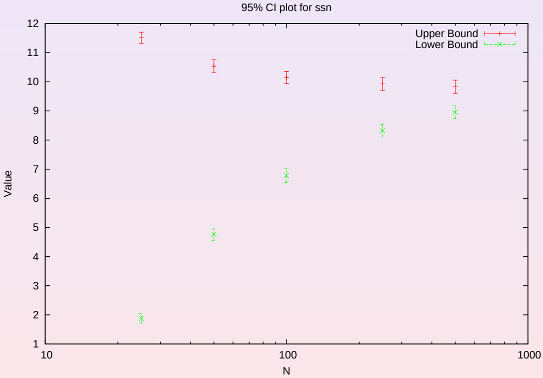


Figure: Upper and lower bounds for problem ssn at 95% confidence interval.

Vary Partition Size – snip7x5

Partition Size	DE Time	Solve Time	Iteration
25	0.13	195.36	28.3
50	0.40	205.86	28.2
100	1.28	183.95	29.9
200	4.30	171.75	27.9
400	12.56	154.55	25.3

Table: Average solving time and number of iterations for snip7x5 with N = 10,000

Upper and Lower Bounds after Warm Start – snip7x5

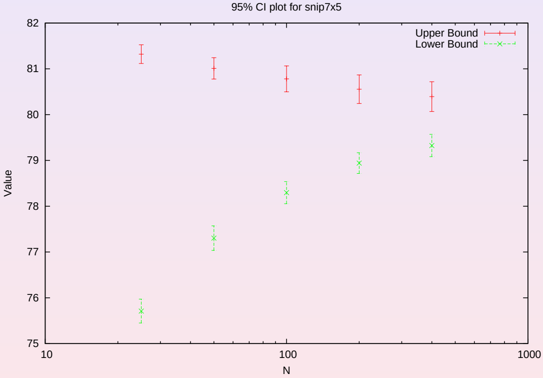


Figure: Upper and lower bounds for problem snip7x5 at 95% confidence interval

Vary Partition Size – semi4

Partition Size	DE Time	Solve Time	Iteration
1	0.27	736.99	80.2
2	0.55	746.92	76.6
4	1.28	713.05	76.9
8	3.25	722.62	72.3
16	8.30	692.32	66.2

Table: Average solving time and number of iterations for semi4 with N = 800

Upper and Lower Bounds after Warm Start – semi4

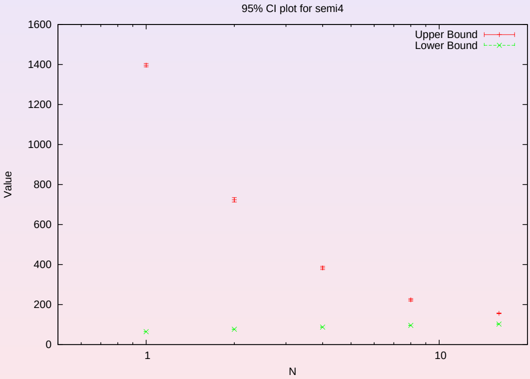


Figure: Upper and lower bounds for problem semi4 at 95% confidence interval.

Conclusions

- We use scenario partitioning method for warm start in solving large scale SP
- Our method provides a good starting point and also provides cuts that tighten the lower bound
- Computational time and number of iteration can be reduced significantly in most instances
- Changing the size of partition contributes small changes in performance
- Our method best suit in parallel environment