A new stabilization method for column generation

Artur Pessoa

Eduardo Uchoa

Marcus Poggi de Aragão

Column generation may present convergence problems

When this happen ?

- The clearer factor that affects convergence speed is quite obvious: the size of the Master LP!

Column generation may present convergence problems

For example:

- CVRP instance E101 has 100 clients:
  - If the capacity is such that one needs 14 vehicles, convergence happens in less than 50 iterations;
  - If the capacity is increased such that 7 vehicle are needed, convergence requires more than 300 iterations;
  - If the capacity is such that 3 vehicles suffice, > 1000 iterations.

Column generation may present convergence problems

Column generation is a technique for solving Master LPs with a huge number of columns. But its behaviour still depends on “how huge” are those LPs.

- There are many more columns in the instance with 3 vehicles than in the instance with 14 vehicles.

Column generation may present convergence problems

More constrained instance => “few” columns in the Master LP => fast convergence.

Less constrained instance => “many” columns in the Master LP => slow convergence.

Column generation may present convergence problems

When this happen ?

- However, there are other more misterious factors also affecting the speed of convergence:

  - Too much degeneracy in the optimal solutions of the restricted Master LPs solved during the column generation is bad.

Column generation may present convergence problems

When this happen ?

- However, there are other more misterious factors also affecting the speed of convergence:

  - “Symmetry in the columns” (many almost equivalent columns with the same cost) is bad.

Dual Stabilization

Several dual stabilization techniques to improve convergence are used since the seventies by the non-differentiable optimization community. (Column generation is a method for solving a kind of piecewise linear concave maximization problem).

Du Merle, Villeneuve, Desrosiers and Hansen (1999) proposed a dual stabilization method specifically devised for column generation.

Dual Stabilization

MVDH99 stabilization is based on some observations:

- The columns that are part of the optimal basis are only generated in the last iterations, when the dual variables are already close their optimal values.
- Dual variables may oscilatte wildly in the first iterations, leading to “extreme columns”, with no chance of being part of the optimal basis.

Dual Stabilization

MVDH99 stabilization is based on some observations:

- So, one should try to generate columns using dual variables near the <u>stability center</u>, the current best guess for the optimal values of the dual variables.

Dual Stabilization

MVDH99 stabilization uses a <u>stabilizing function</u> penalizing (in the Master LP) dual solutions much away from the stability center.

- The stability center is changed along the method, until it converges to an optimal dual solution.
- The stabilizing function is also changed along the method, until it converges to a null function.

MVDH99 Stabilization

Let *P* be a feasible and bounded Master LP and *D* its dual

$$(P) \quad \begin{array}{ll} \min & cx \\ s.a & Ax = b \\ & x \geq 0 \end{array} \qquad \Bigg| \qquad (D) \quad \begin{array}{ll} \max & b\pi \\ s.a. & \pi A \leq c \end{array}$$

MVDH99 Stabilization

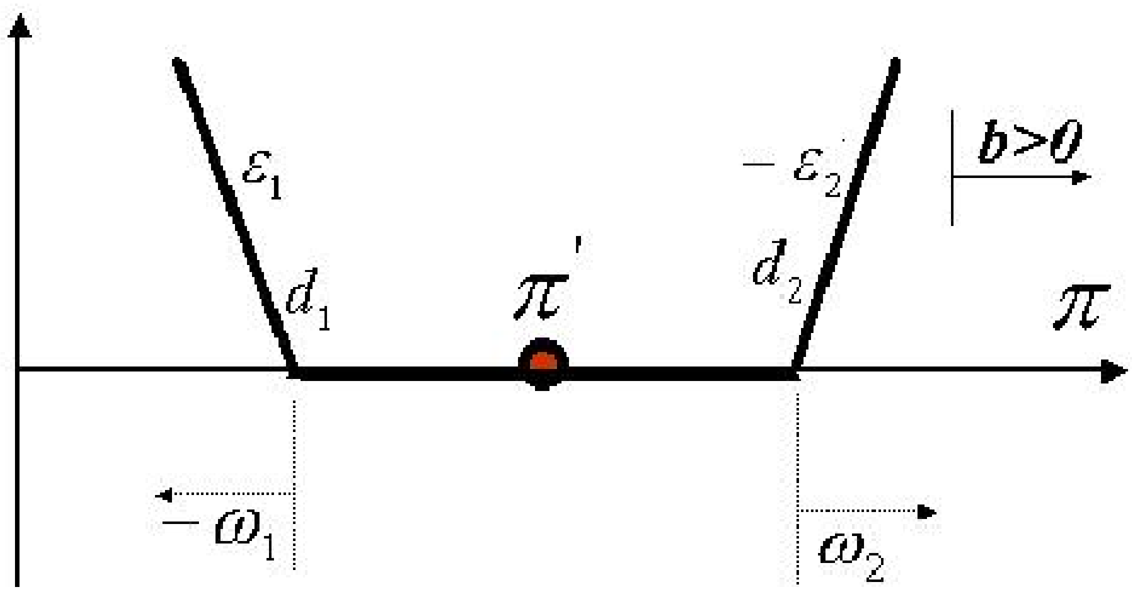
The stabilized primal P_e contains additional artificial columns modelling the stabilizing function

$$
\begin{array}{lllllll|l}
 & \min & cx & -d_1 y_1 & +d_2 y_2 & & & \\
(P_e) & s.a & Ax & -y_1 & +y_2 & = & b & \pi \\
 & & & y_1 & & \leq & \epsilon_1 & -\omega_1 \leq 0 \\
 & & & & y_2 & \leq & \epsilon_2 & -\omega_2 \leq 0 \\
 & & \multicolumn{3}{c}{x \geq 0, \;\; y_1 \geq 0, \;\; y_2 \geq 0} & & &
\end{array}
$$

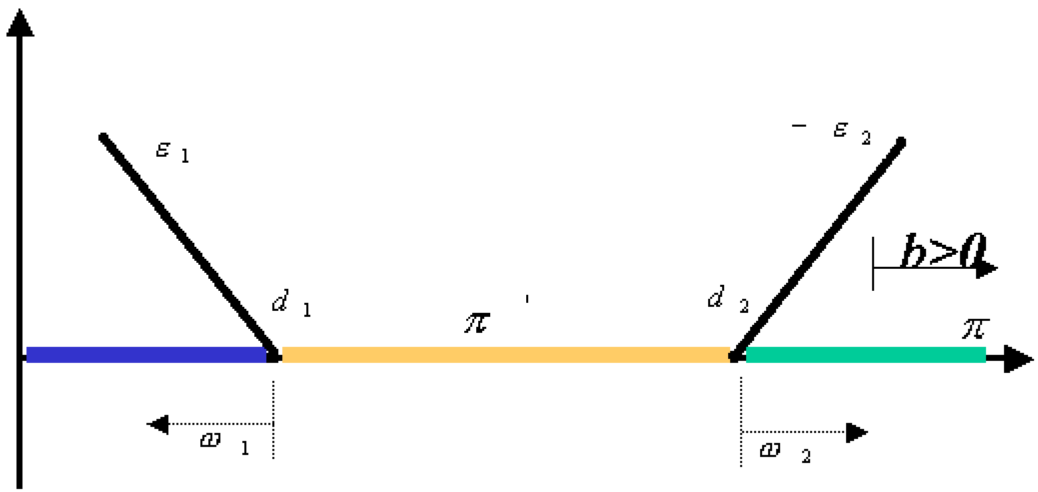
MVDH99 Stabilization

The corresponding dual problem D_e is:

$$
\begin{array}{llr}
 & \max & b\pi - \omega_1\epsilon_1 - \omega_2\epsilon_2 \\
(D_e) & s.a. & \pi A \le c \\
 & & d_1 - \omega_1 \le \pi \le d_2 + \omega_2 \\
 & & \omega_1 \ge 0, \omega_2 \ge 0
\end{array}
$$



There are rules for changing the stabilizing function along the iterations



-If π_i is out of the interval (i.e. it incurs a penalty), recenter (change the current stability center to it) and increase the interval.

-If π_i is within the interval, recenter, reduce interval and reduce penalties ε.

- ...

MVDH99 Stabilization

Recent implementations (Bem Amor, Frangioni and Desrosiers 2007) recommend using 5-piecewise linear stabilizing functions.

Drawbacks:

- Even more parameters to be calibrated.
- Increase of the size of the restricted Master LP (4 additional artificial variables by row)

The newly proposed stabilization method

Still based on the concept of keeping a stability center, but it has the following potential advantages:

- No need to change the Master LP
- Very simple, a single parameter to be calibrated
- Nice theoretical properties

Assumption

One obtains a valid Lagrangean lower bound $L(\pi)$ every time the pricing problem is solved with dual vector π.

- This is always the case when the Master LP arises from a Dantzig-Wolfe decomposition.

The newly proposed method

Input : parameter α, $0 < \alpha \leq 1$, and value ε

$\bar{\pi} \leftarrow 0;$

Do

Solve the restricted Master LP, obtaining the value Z_{RM} and the dual vector π_{RM};

(If you want, remove some non-basic columns);

$\pi_{ST} \leftarrow \alpha\bar{\pi} + (1-\alpha)\pi_{RM}$;

Solve the pricing with vector π_{ST}, obtaining column A_j;

If $L(\pi_{ST}) > L(\bar{\pi})$ then $L(\bar{\pi}) \leftarrow L(\pi_{ST})$;

If A_j has negative reduced cost with respect to π_{RM}, add it to the restricted Master LP;

Until $Z_{RM} - L(\bar{\pi}) < \varepsilon$

The newly proposed method

The trick: the pricing problem is not solved with the dual solutions from the restricted Master LPs, but with other vectors, that are closer to the current stability center!

Is this method sound ?

Theorem

If the solution of the pricing with vector π_{ST} does not give a column with negative reduced cost with respect to π_{RM}, then

$$L(\pi_{ST}) \geq L(\overline{\pi}) + \alpha(Z_{RM} - L(\overline{\pi})).$$

Theorem

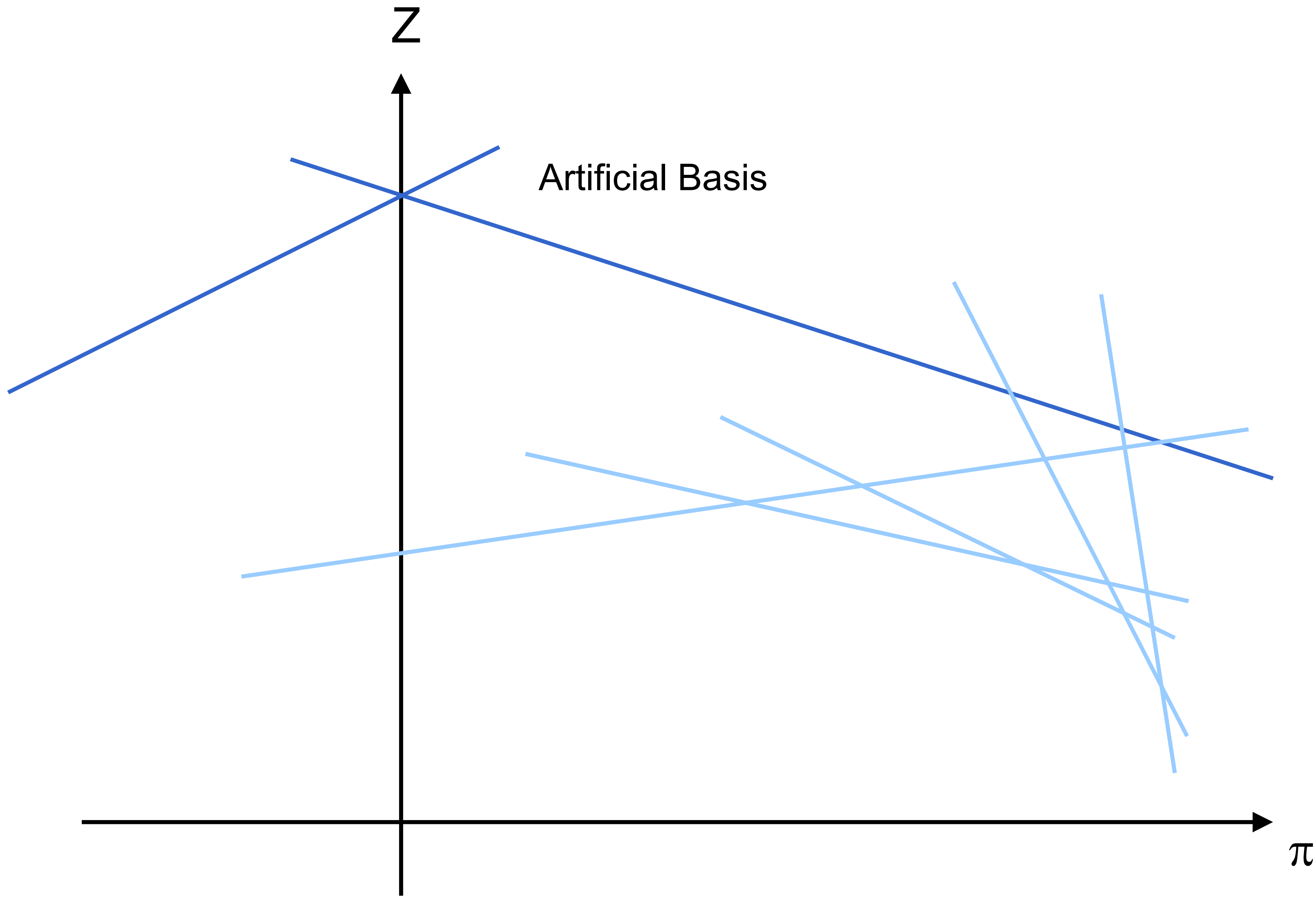
A misprice happens when the column generated by the stabilized pricing does not have negative reduced cost with respect to the "true" duals.

- The theorem says that a misprice is not a waste of time, quite to the contrary, it is guarantee that the gap $(Z_{RM} - L(\overline{\pi}))$ is reduced by at least a factor of 1/(1- α).

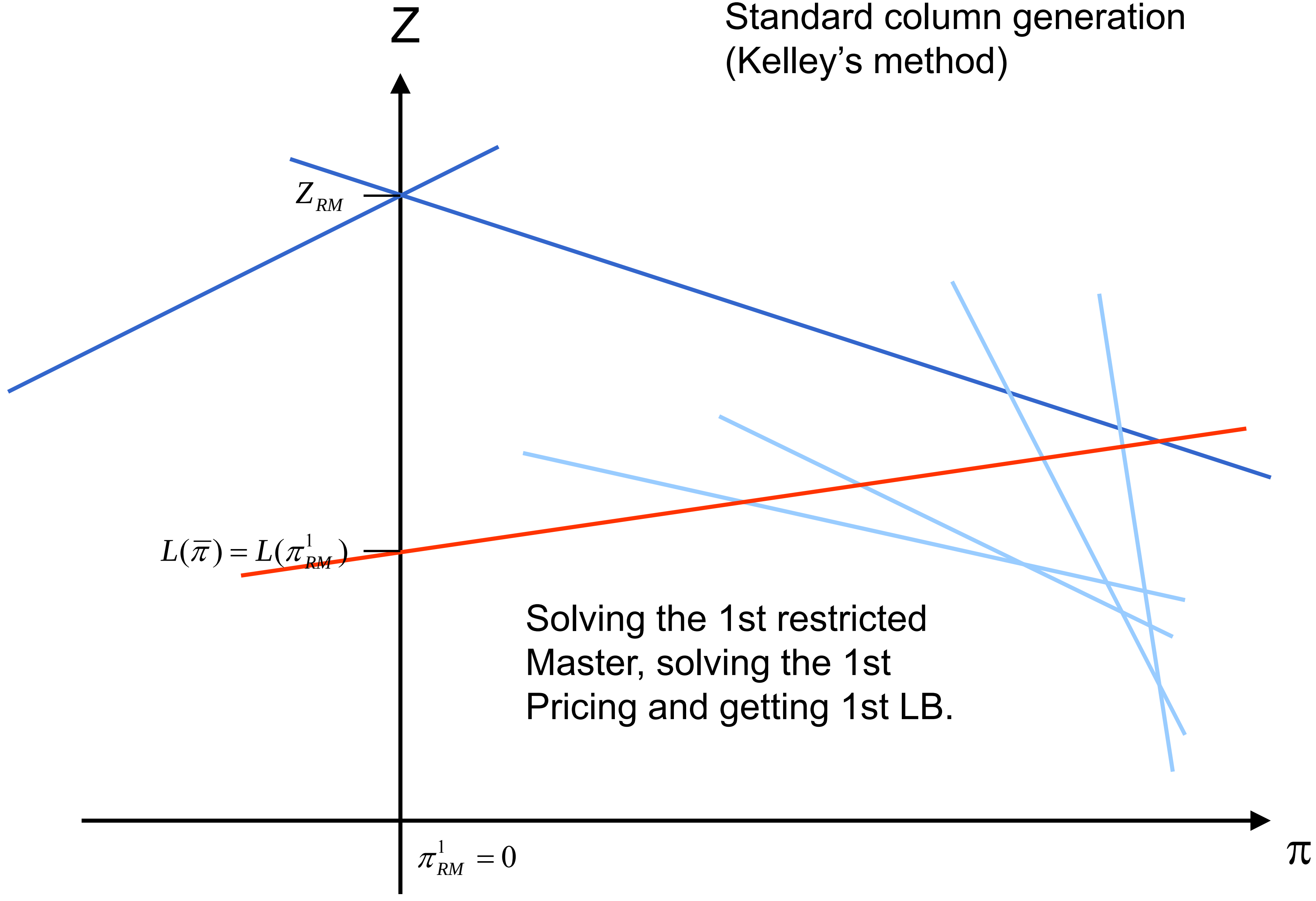
Theorem

Corollary: the method converges in a finite number of iterations.

The value of ε can be calculated to assure that an optimal basis of the Master LP is achieved.

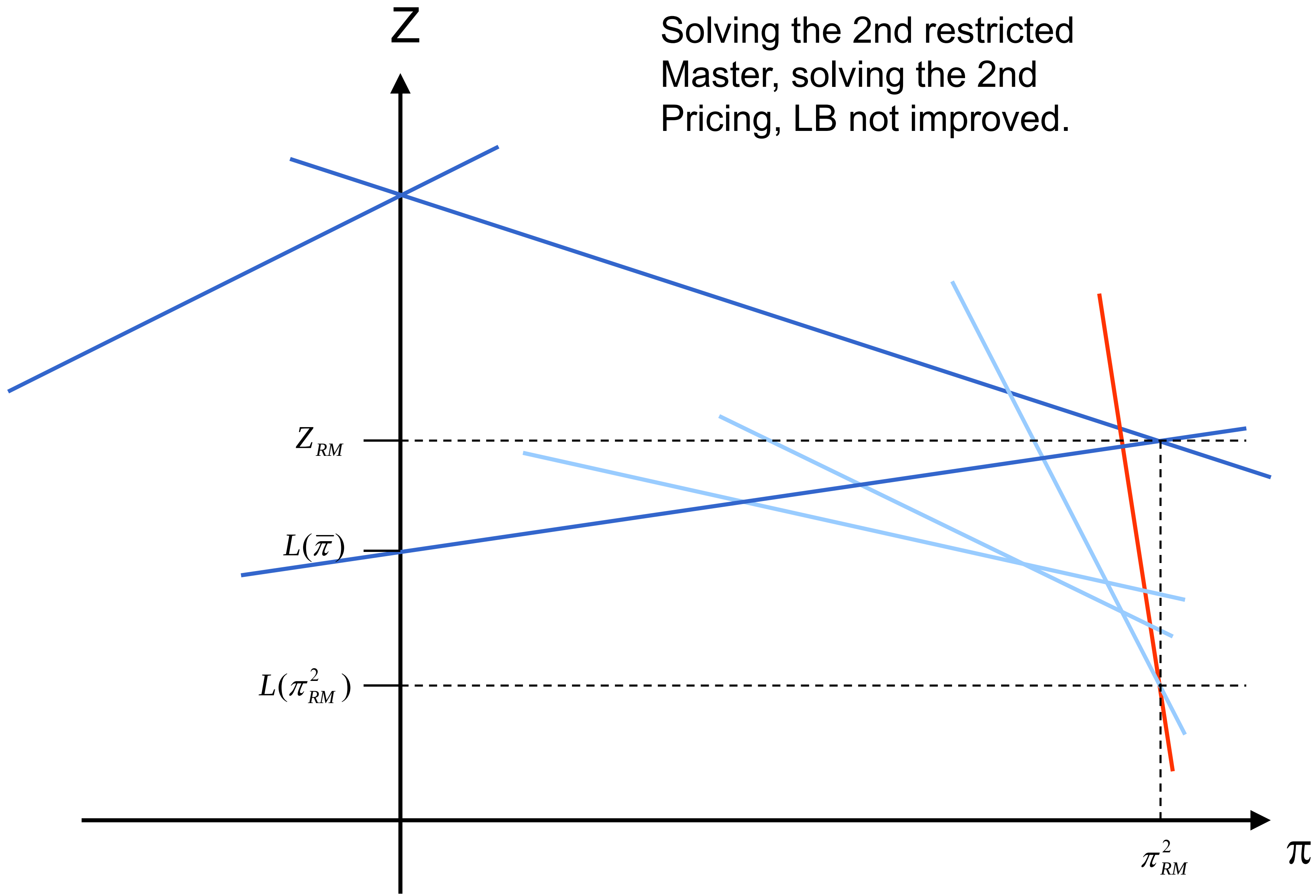

Z
Artificial Basis
π

Standard column generation
(Kelley's method)

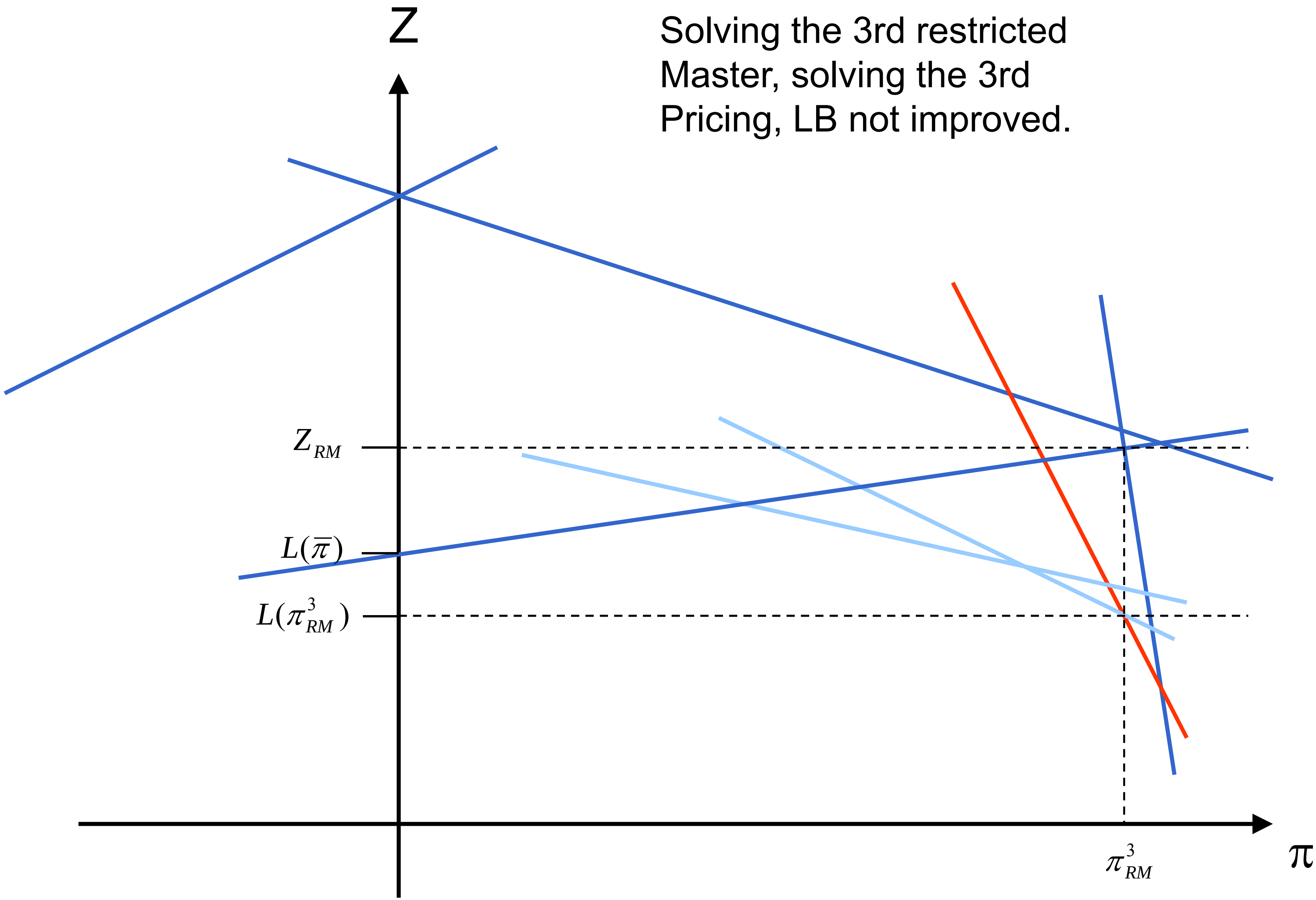


Solving the 1st restricted
Master, solving the 1st
Pricing and getting 1st LB.

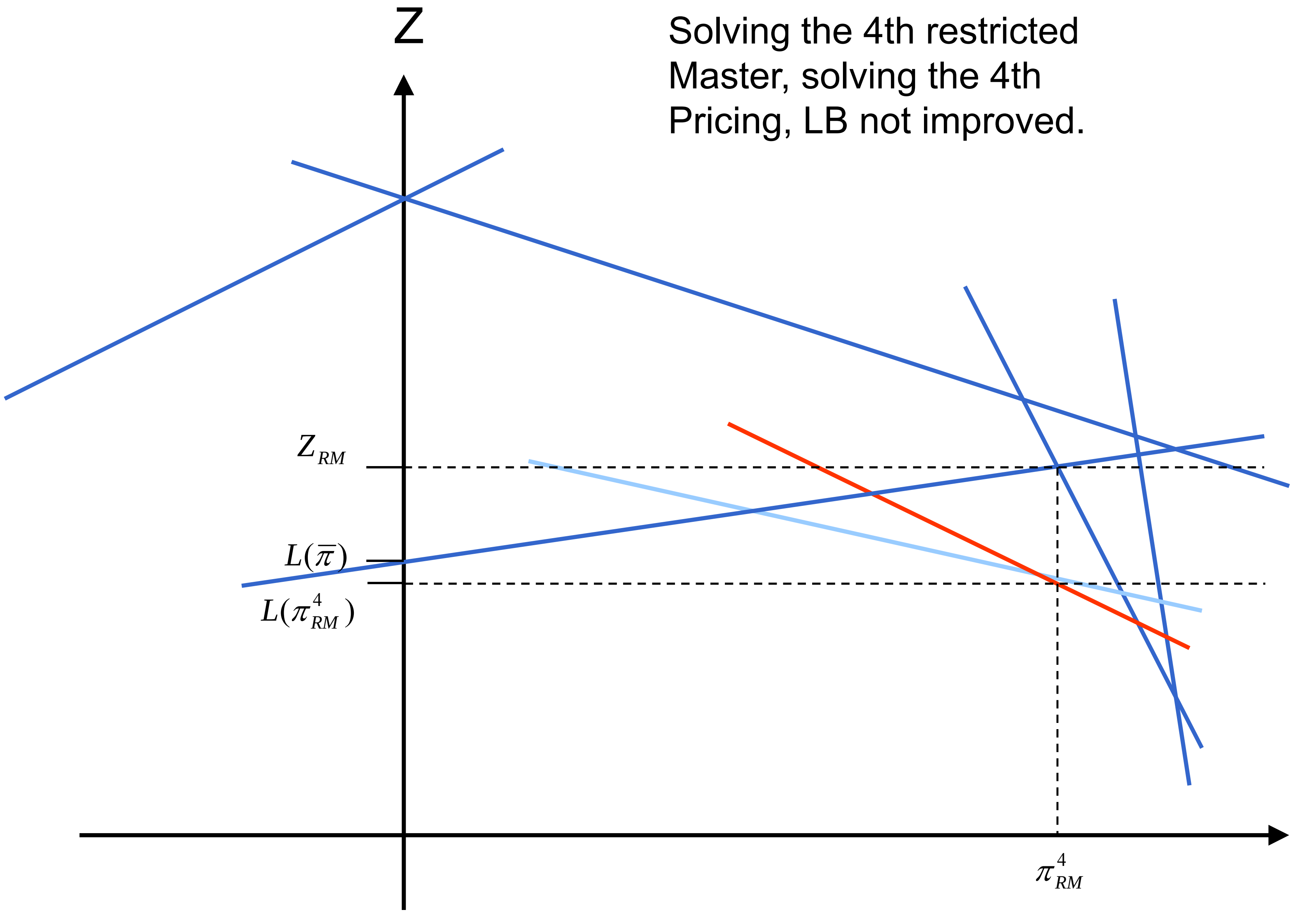
Solving the 2nd restricted Master, solving the 2nd Pricing, LB not improved.



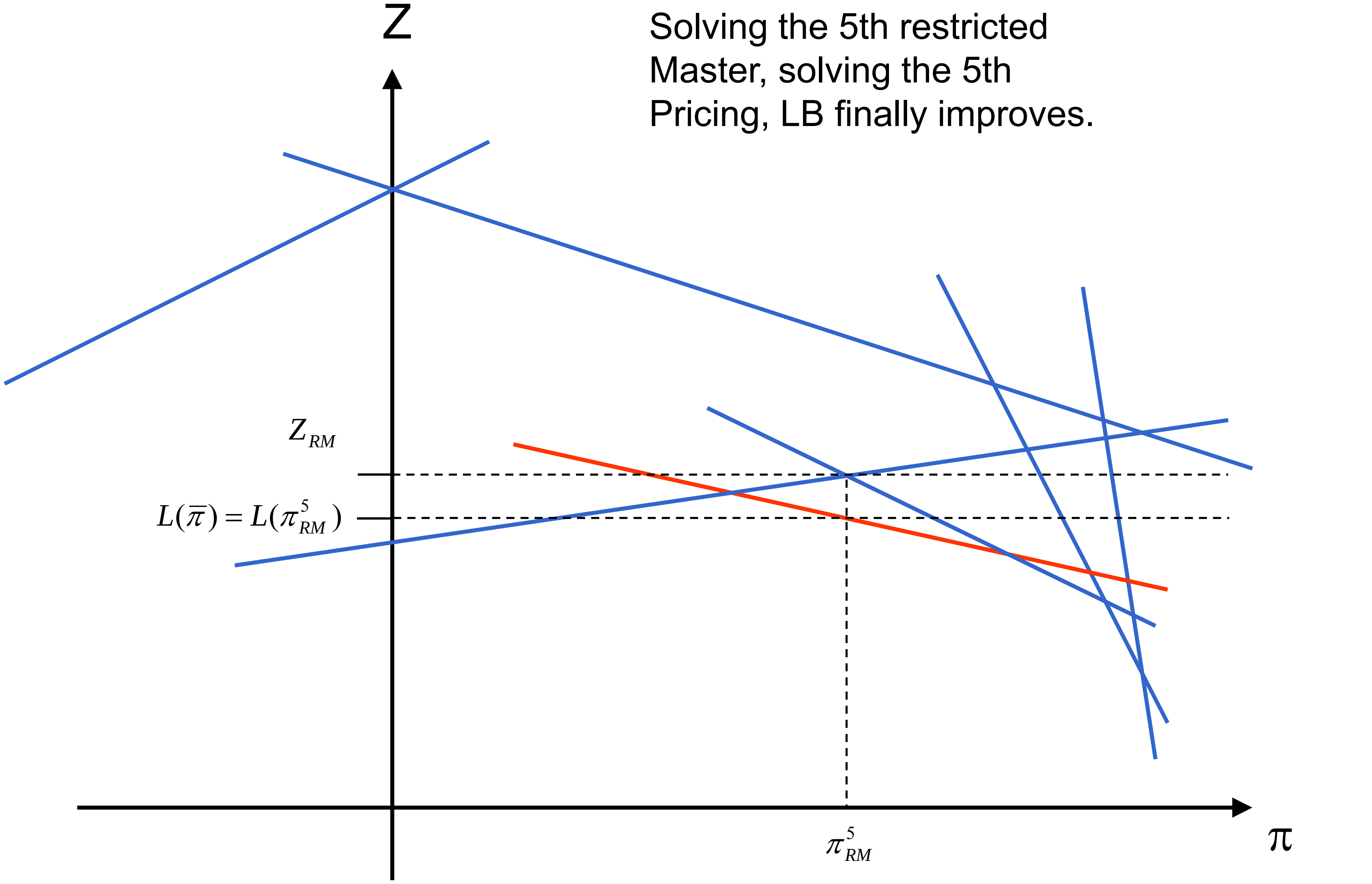
Solving the 3rd restricted Master, solving the 3rd Pricing, LB not improved.

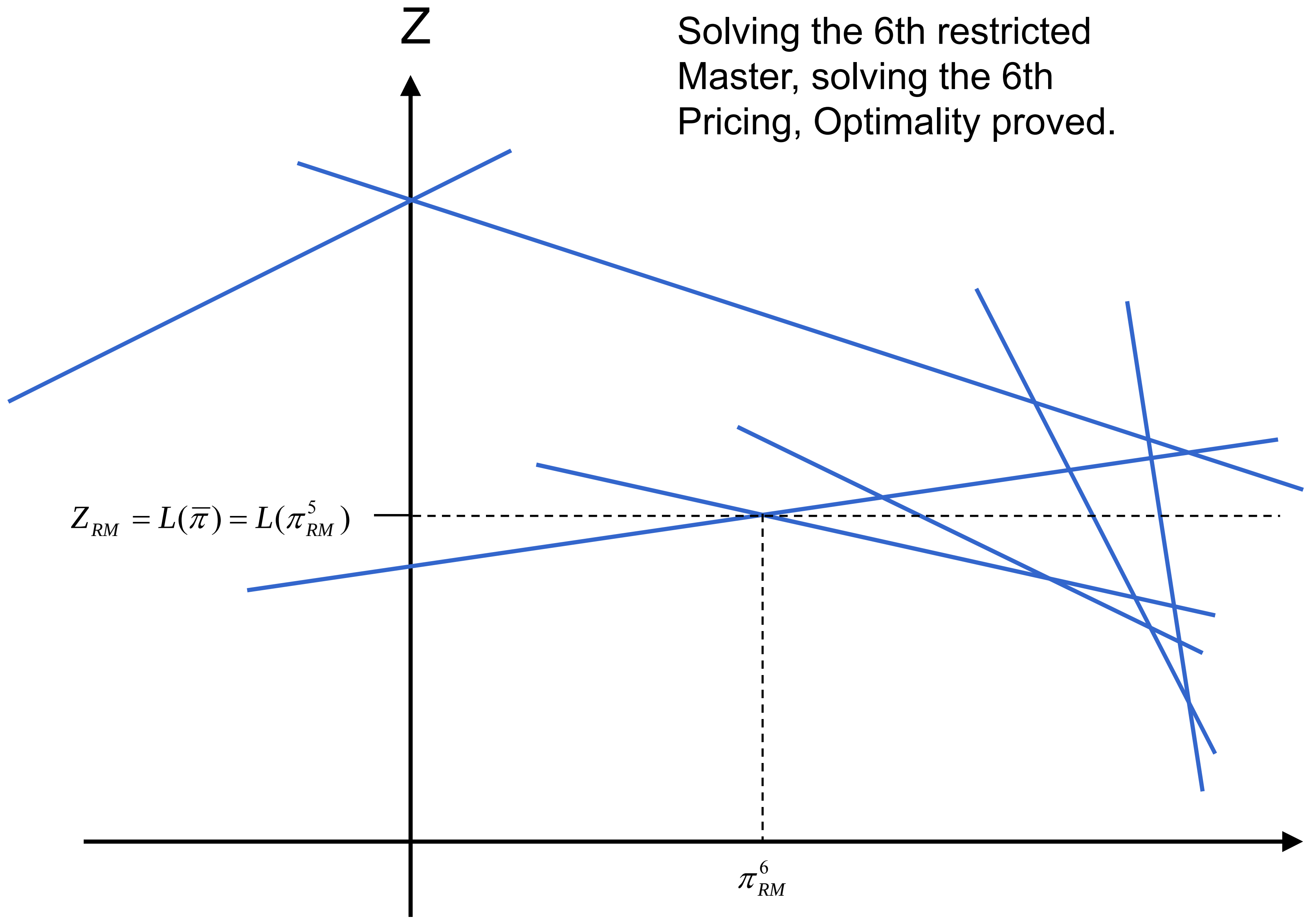


Solving the 4th restricted Master, solving the 4th Pricing, LB not improved.

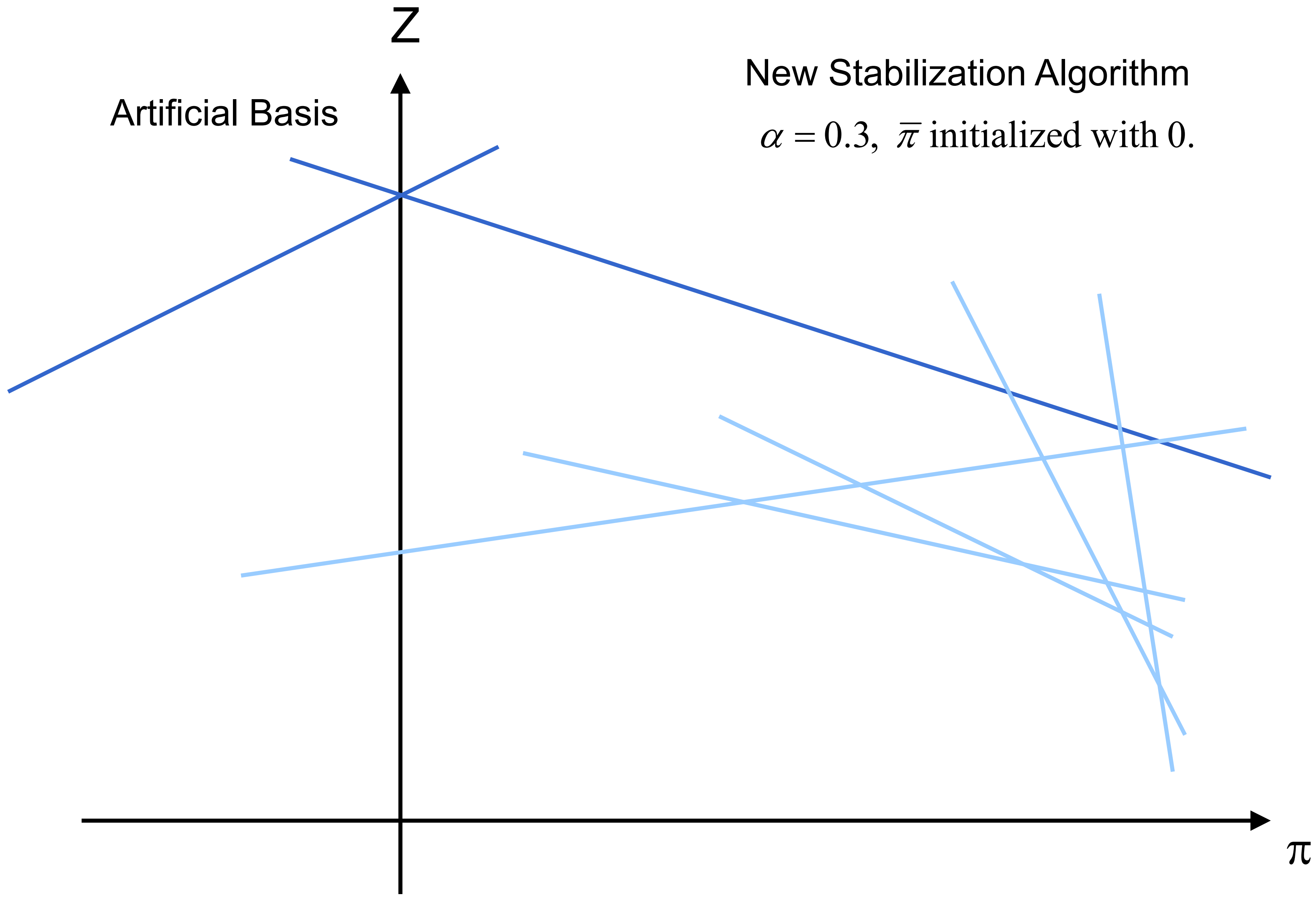


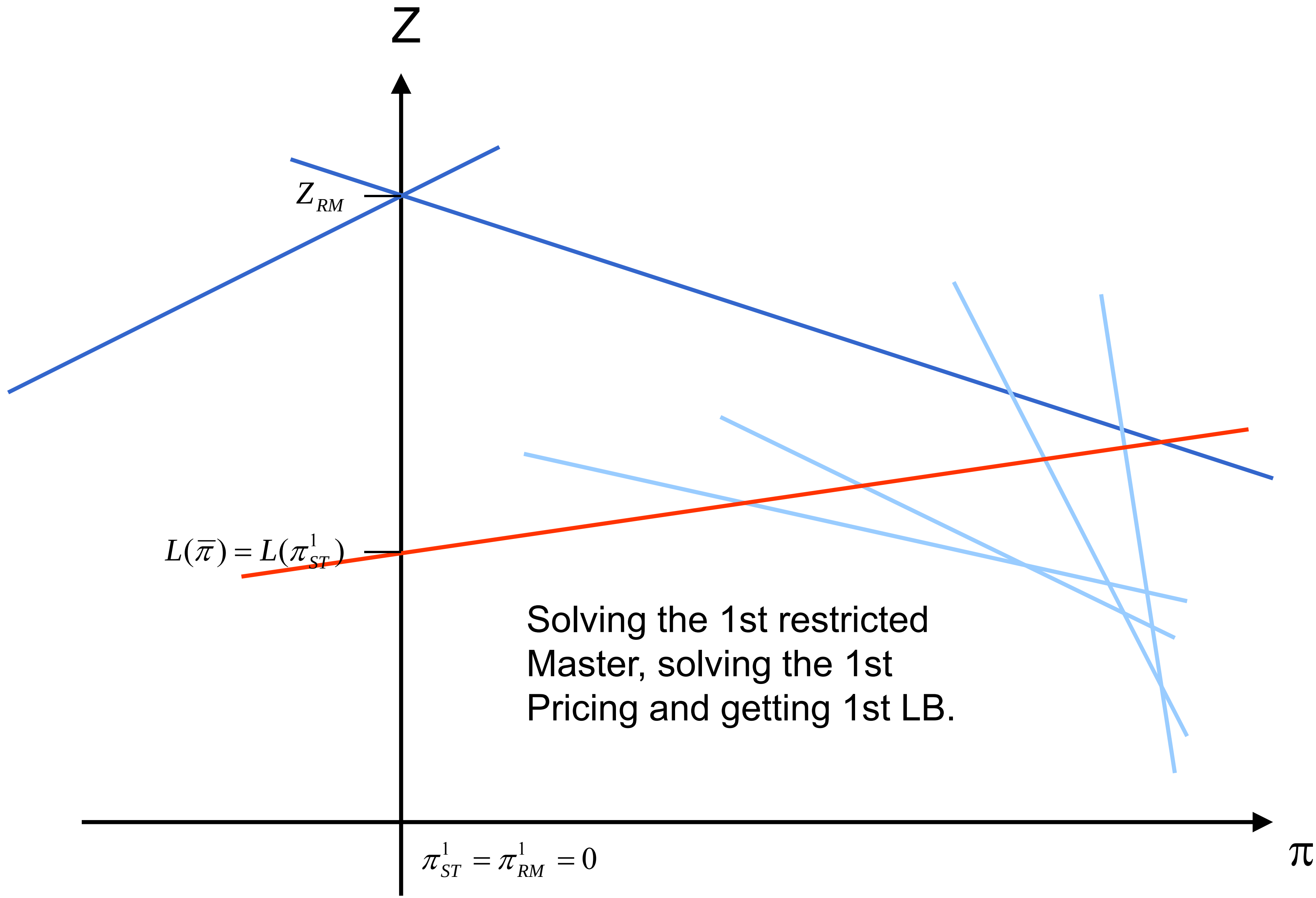
Solving the 5th restricted Master, solving the 5th Pricing, LB finally improves.





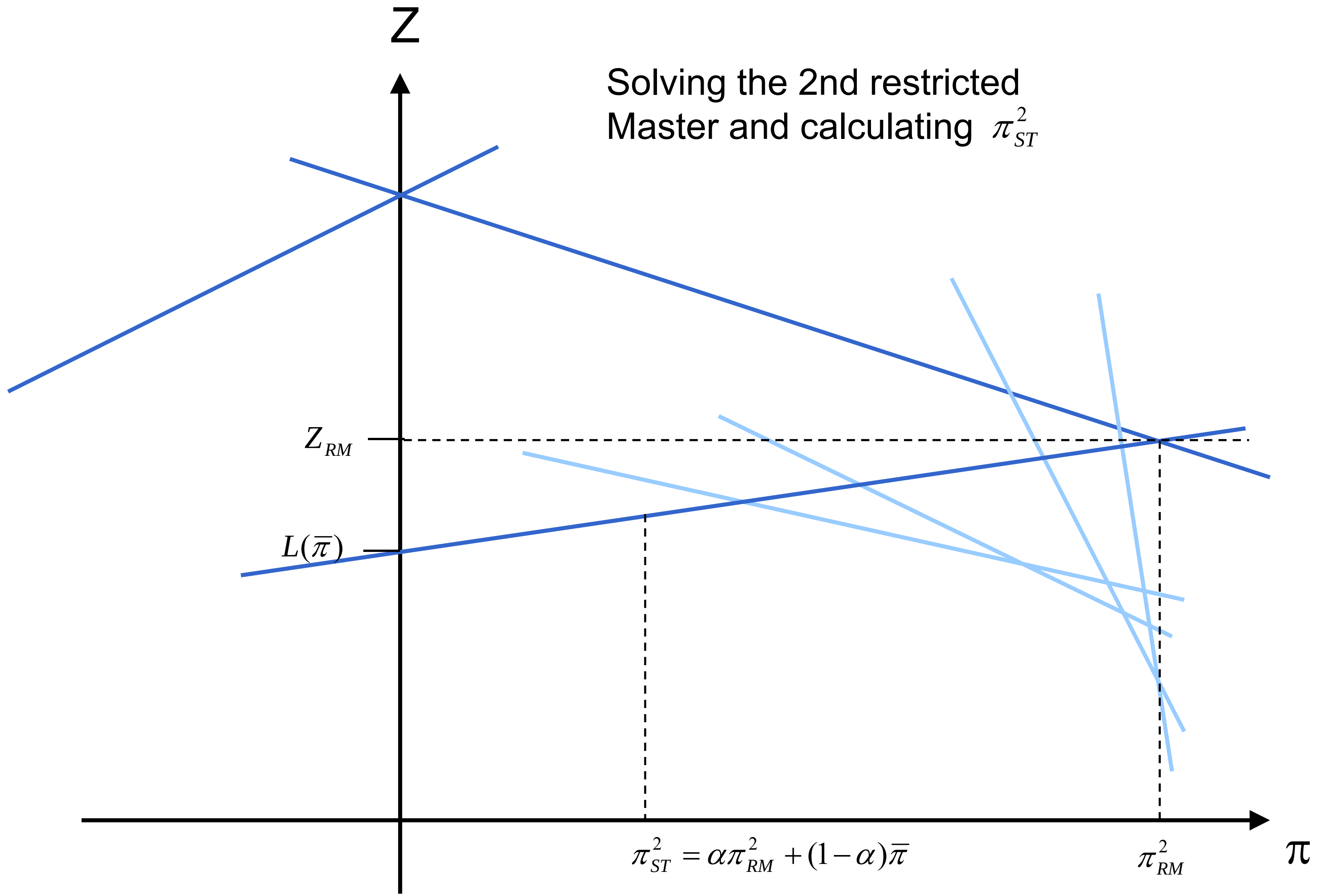
Z
Solving the 6th restricted Master, solving the 6th Pricing, Optimality proved.
$Z_{RM} = L(\overline{\pi}) = L(\pi^5_{RM})$
π^6_{RM}
π

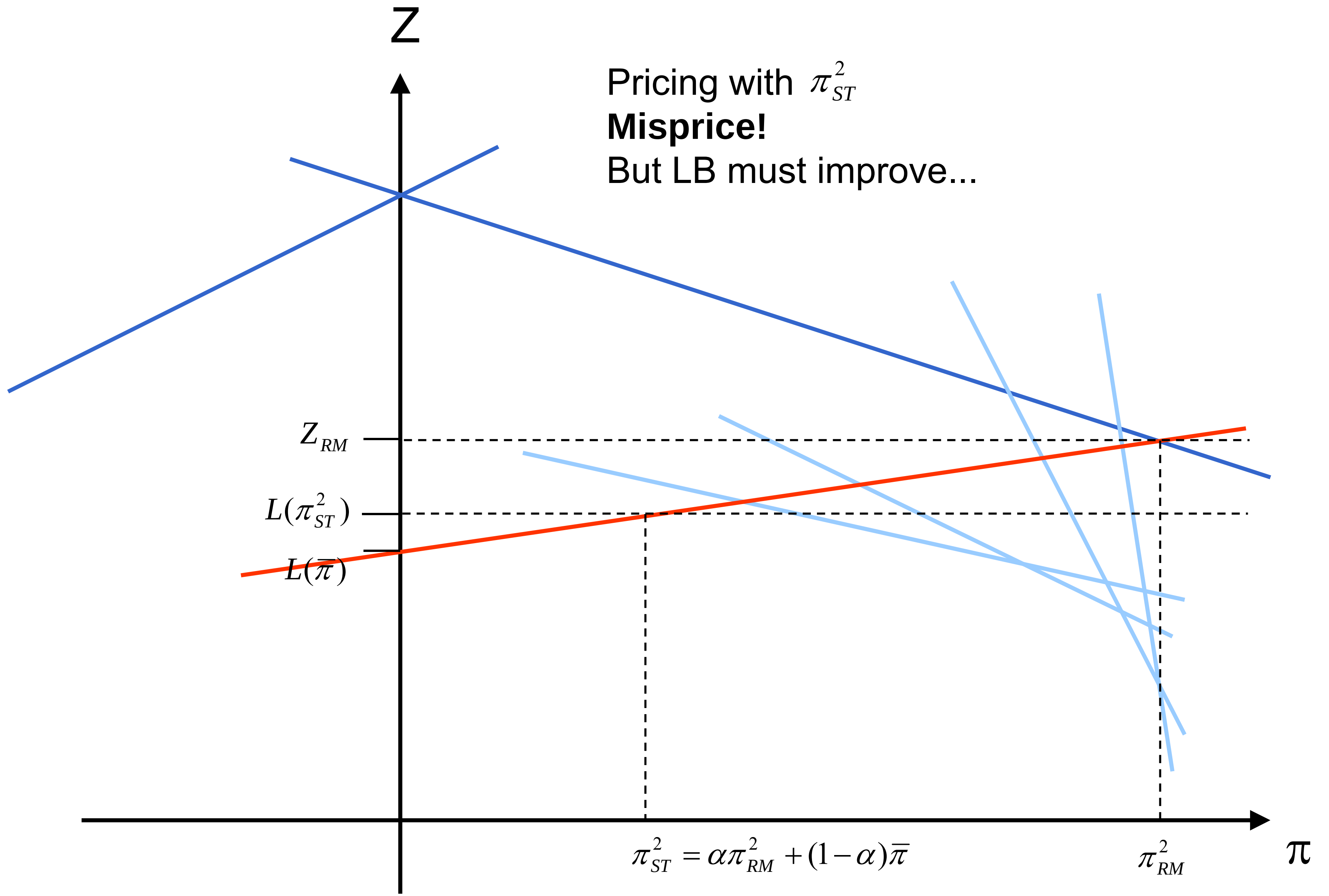

Z
New Stabilization Algorithm
$\alpha = 0.3$, $\bar{\pi}$ initialized with 0.
Artificial Basis
π



Z
Z_{RM}
$L(\overline{\pi}) = L(\pi^1_{ST})$
Solving the 1st restricted
Master, solving the 1st
Pricing and getting 1st LB.
$\pi^1_{ST} = \pi^1_{RM} = 0$
π

Solving the 2nd restricted Master and calculating π_{ST}^2





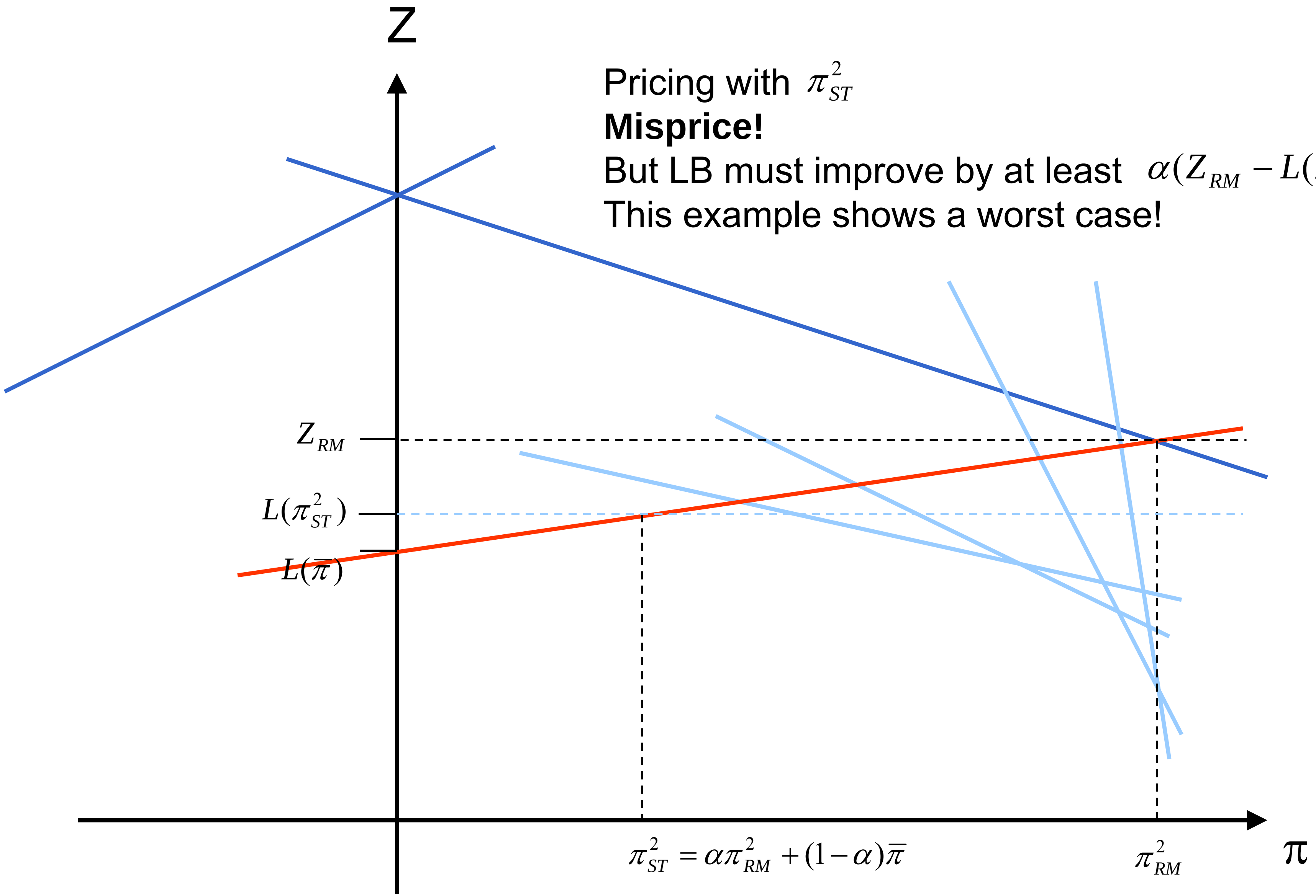
Z
Pricing with π^2_{ST}
Misprice!
But LB must improve...
Z_{RM}
$L(\pi^2_{ST})$
$L(\overline{\pi})$
$\pi^2_{ST} = \alpha\pi^2_{RM} + (1-\alpha)\overline{\pi}$
π^2_{RM}
π

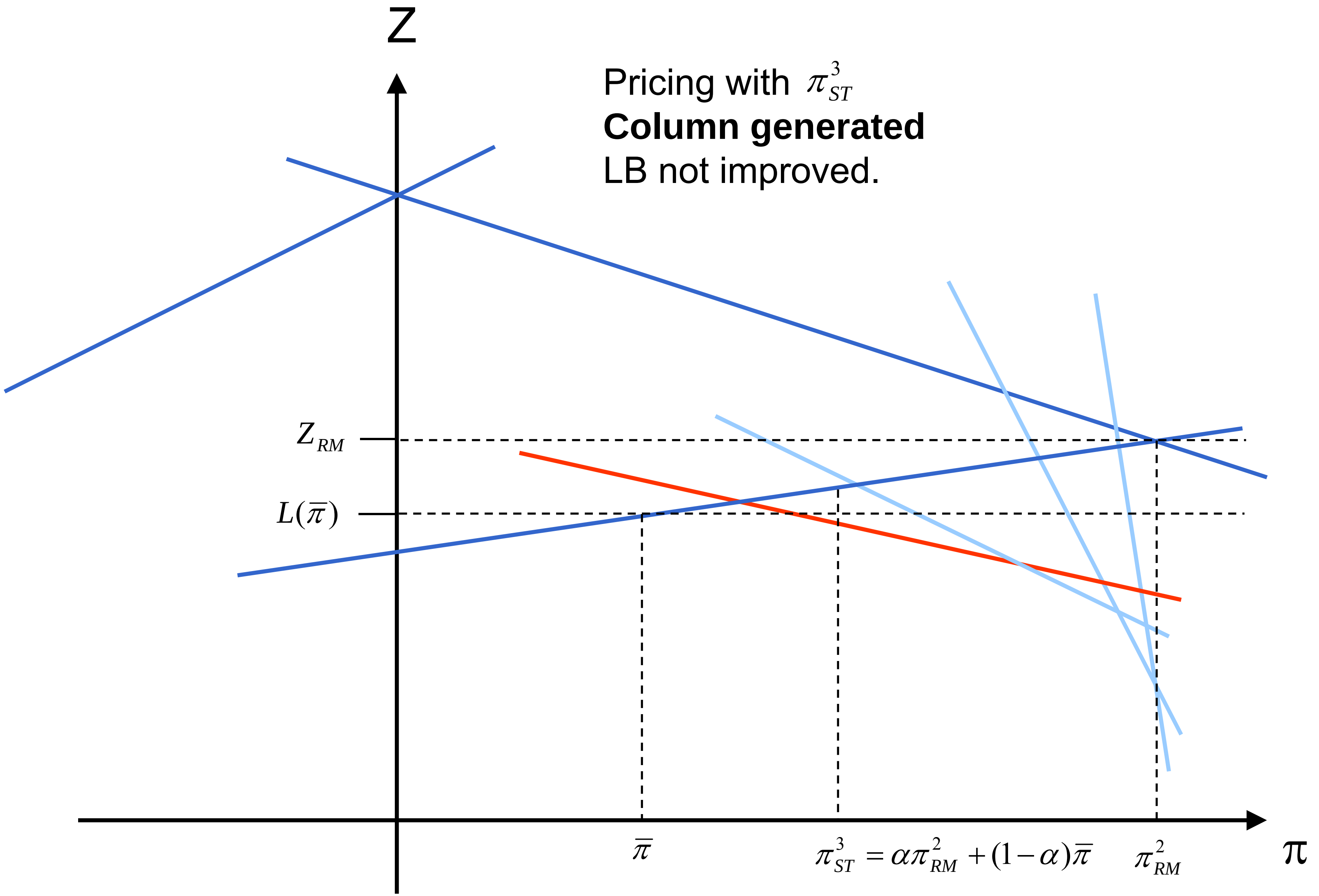
Pricing with π_{ST}^{2}

Misprice!

But LB must improve by at least $\alpha(Z_{RM} - L(\overline{\pi}))$

This example shows a worst case!



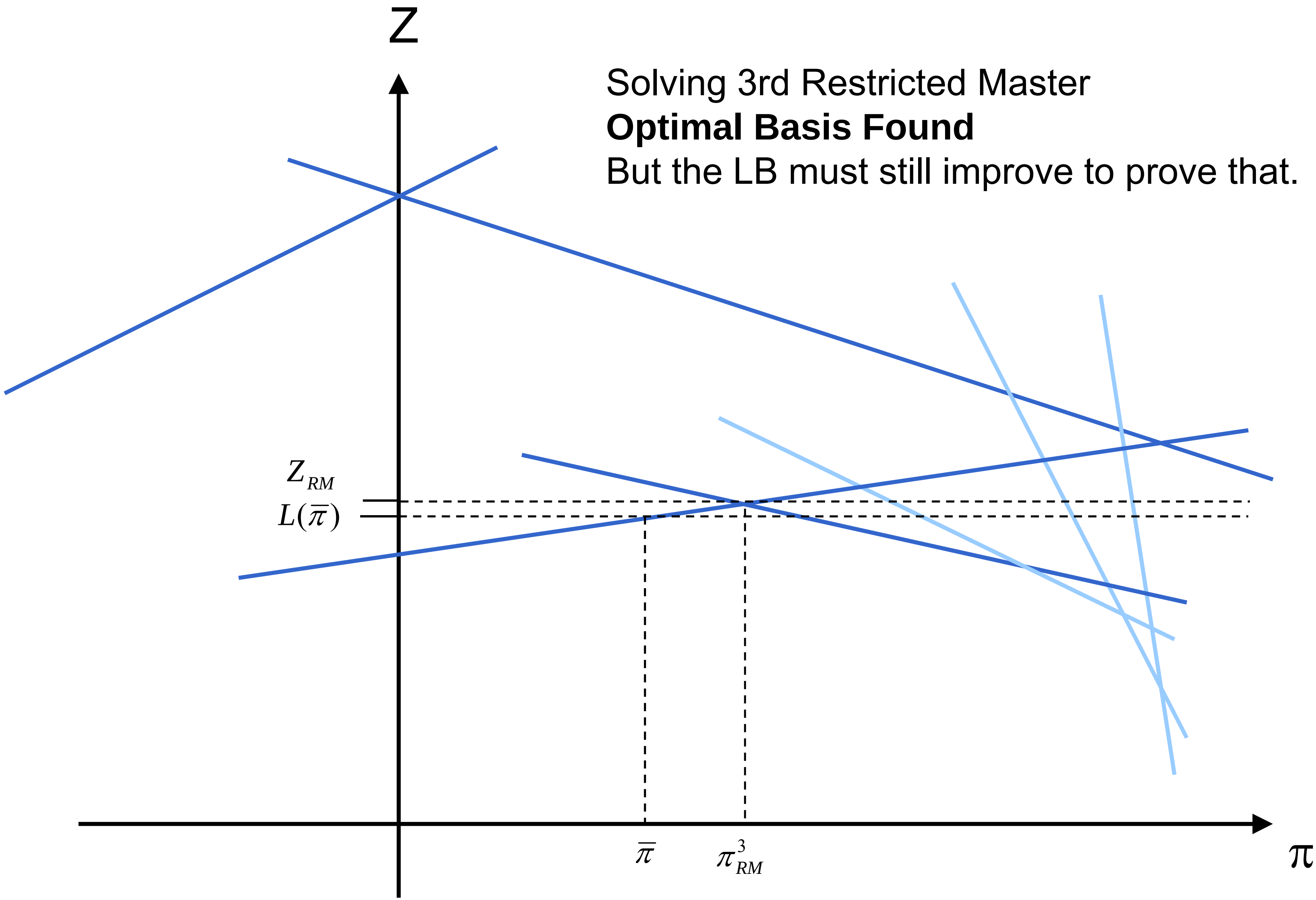


Z
Pricing with π_{ST}^{3}
Column generated
LB not improved.
Z_{RM}
$L(\bar{\pi})$
$\bar{\pi}$
$\pi_{ST}^{3} = \alpha\pi_{RM}^{2} + (1-\alpha)\bar{\pi}$
π_{RM}^{2}
π

Solving 3rd Restricted Master

Optimal Basis Found

But the LB must still improve to prove that.



Experiments

Single Machine Weighted Tardiness problem $(1||\Sigma w_j T_j)$

Given a set of *n* jobs, where each job *j* has a:

- Processing time (p_j),
- Due date (d_j),
- Weight (w_j),

Sequence the jobs minimizing $\Sigma w_j T_j$, where $T_j = \max\{0, C_j - d_j\}$ is the tardiness of j with respect to its completion time C_j.

Parallel Identical Machines Weighted Tardiness problem ($P||\Sigma w_j T_j$)

Given a set of *n* jobs, where each job *j* has a:

- Processing time (p_j),
- Due date (d_j),
- Weight (w_j),

and m identical machines,

Sequence the jobs in the available machines minimizing $\Sigma w_j T_j$, where $T_j = \max\{0, C_j - d_j\}$ is the tardiness of j with respect to its completion time C_j.

BCP for scheduling problems

Pessoa, Uchoa, Poggi de Aragão and Freitas (2008) proposed a BCP for those kinds of scheduling problems (presentation tomorrow), by considering them as VRPs.

Column generation convergence without stabilization is poor:

- Huge number of columns in the Master LP
- Extreme degeneracy (when m=1, an optimal basis may have one variable with a non-zero value)
- Symmetry in the columns

Results over $1||\Sigma w_j T_j$ instances

- OR-Library benchmarks (375 instances with 40, 50 and 100 jobs).
- α fixed to 0.10

Methods:

A – Standard column generation

B – Stabilized column generation

C - Standard column generation + fixing by red. costs

D - Stabilized column generation + fixing by red. costs

E – Some iterations of the Volume algorithm to hot start the stabilization center + D

n	Met.	Time	Iter	St.ch	MisP	R.Arcs
40	A	367	819	-	-	1.5M
	B	55	116	83	38	1.5M
	C	45	322	-	-	30
	D	14	87	72	33	151
	E	12	72	2	1	3
50	A	1545	1766	-	-	3M
	B	147	160	95	38	3M
	C	138	586	-	-	241
	D	35	119	85	33	562
	E	28	103	25	5	180
100	D	673	338	146	40	5246
	E	387	267	23	18	4855

n	Met.	Time	Iter	St.ch	MisP	R.Arcs
40	A	367	819	-	-	1.5M
	B	55	116	83	38	1.5M
	C	45	322	-	-	30
	D	14	87	72	33	151
	E	12	72	2	1	3
50	A	1545	1766	-	-	3M
	B	147	160	95	38	3M
	C	138	586	-	-	241
	D	35	119	85	33	562
	E	28	103	25	5	180
100	D	673	338	146	40	5246
	E	387	267	23	18	4855

n	Met.	Time	Iter	St.ch	MisP	R.Arcs
40	A	367	819	-	-	1.5M
	B	55	116	83	38	1.5M
	C	45	322	-	-	30
	D	14	87	72	33	151
	E	12	72	2	1	3
50	A	1545	1766	-	-	3M
	B	147	160	95	38	3M
	C	138	586	-	-	241
	D	35	119	85	33	562
	E	28	103	25	5	180
100	D	673	338	146	40	5246
	E	387	267	23	18	4855

n	Met.	Time	Iter	St.ch	MisP	R.Arcs
40	A	367	819	-	-	1.5M
	B	55	116	83	38	1.5M
	C	45	322	-	-	30
	D	14	87	72	33	151
	E	12	72	2	1	3
50	A	1545	1766	-	-	3M
	B	147	160	95	38	3M
	C	138	586	-	-	241
	D	35	119	85	33	562
	E	28	103	25	5	180
100	D	673	338	146	40	5246
	E	387	267	23	18	4855

Conclusions

It is a promising method, that should be compared with the methods that use stabilization functions.

- Not done yet

Conclusions

There is room for complicating the method a little bit, in order to improve its practical behaviour:

- A simple idea is increasing the α parameter when one suspects that the current restricted Master LP solution is already close to the optimal.

Conclusions

Beyond column generation

- New pricing rule for the simplex method ?

Thank You!