Online Supplement for
Analytics branching and selection for the capacitated multi-item lot sizing problem with non-identical machines

Tao Wu[†], Zhe Liang[‡,*], Canrong Zhang[§]
[†] Advanced Analytics Department, Dow Chemical, 48642 Midland, MI, USA, danielwu9999@gmail.com
[‡,*] School of Economics & Management, Tongji University, Shanghai 200092, China, liangzhe@tongji.edu.cn
[§] Logistics Engineering and Simulation Laboratory, Graduate School at Shenzhen, Tsinghua University, Shenzhen 518055, China, crzhang@sz.tsinghua.edu.cn

This online supplement complements the paper titled "analytics branching and selection for the capacitated multi-item lot sizing problem with non-identical machines" published in *INFORMS Journal on Computing*. It covers some long proofs and an illustrative example which would have interrupted the flow of the paper but nevertheless provide more insights to the interested reader.

Key words: Production planning and scheduling; Column generation; Dantzig–Wolfe decomposition; Lot sizing; Non-identical parallel machines; Analytics branching and selection.

History:

1. An illustrative example for the analytics branching and selection method

Using an illustrative example, we describe the procedure of building the generalized linear models for generating insightful information and present an iteration of the analytics branching and selection method.

1.1 Generalized linear models for generating insightful information

We next describe the modeling procedure that has 5 main steps:

- **Step I: determine parameter settings of the problems ($\mathbb{P}$) that need to be solved.**
- **Step II: generate a number of groups of problems ($\mathbb{SP}$) that share similar parameter settings with problems $\mathbb{P}$ but have smaller sizes.** We must generate a large number of small-size problems to create a sufficient amount of data to build the

---

[*] Corresponding author

generalized linear model. However, here we describe the modeling procedure using six small-size problems ($\mathbb{SP}1$, $\mathbb{SP}2$, ..., $\mathbb{SP}6$), including three problems with a size J1-T3-M2 and the other three with a size J1-T2-M3.

- **Step III: obtain solution values $\hat{Y}$, $\bar{Y}$, ξ, and $\dot{Y}^1_{hybrid}$ for all small-size problems $\mathbb{SP}$.** We assume that the solution values $\hat{Y}$, $\bar{Y}$, ξ, and $\dot{Y}^1_{hybrid}$ for the six small-size problems $\mathbb{SP}$ are given in Tables 1 - 4.

Table 1 **The optimal solution values $\hat{Y}$**

	J1-T3-M2				J1-T2-M3		
$\hat{Y}^m_{jt}$	$\mathbb{SP}1$	$\mathbb{SP}2$	$\mathbb{SP}3$	$\hat{Y}^m_{jt}$	$\mathbb{SP}4$	$\mathbb{SP}5$	$\mathbb{SP}6$
$\hat{Y}^1_{11}$	1	0	0	$\hat{Y}^1_{11}$	1	0	0
$\hat{Y}^2_{11}$	0	1	1	$\hat{Y}^2_{11}$	0	1	0
$\hat{Y}^1_{12}$	0	1	1	$\hat{Y}^3_{11}$	0	0	1
$\hat{Y}^2_{12}$	1	0	0	$\hat{Y}^1_{12}$	1	0	0
$\hat{Y}^1_{13}$	0	1	0	$\hat{Y}^2_{12}$	1	0	0
$\hat{Y}^2_{13}$	1	0	1	$\hat{Y}^3_{12}$	0	1	1

Table 2 **The LP-relaxed solution values $\bar{Y}$**

	J1-T3-M2							J1-T2-M3					
	LP-relaxed values			Bins of the LP-relaxed values				LP-relaxed values			Bins of the LP-relaxed values		
$\bar{Y}^m_{jt}$	$\mathbb{SP}1$	$\mathbb{SP}2$	$\mathbb{SP}3$	$\mathbb{SP}1$	$\mathbb{SP}2$	$\mathbb{SP}3$	$\bar{Y}^m_{jt}$	$\mathbb{SP}4$	$\mathbb{SP}5$	$\mathbb{SP}6$	$\mathbb{SP}4$	$\mathbb{SP}5$	$\mathbb{SP}6$
$\bar{Y}^1_{11}$	0.61	0.32	0.33	[0.6, 0.8)	[0.2, 0.4)	[0.2, 0.4)	$\bar{Y}^1_{11}$	0.71	0.13	0.00	[0.6, 0.8)	[0.0, 0.2)	[0.0, 0.2)
$\bar{Y}^2_{11}$	0.00	0.95	0.77	[0.0, 0.2)	[0.8, 1.0]	[0.6, 0.8)	$\bar{Y}^2_{11}$	0.24	0.96	0.11	[0.2, 0.4)	[0.8, 1.0]	[0.0, 0.2)
$\bar{Y}^1_{12}$	0.19	0.26	0.93	[0.0, 0.2)	[0.2, 0.4)	[0.8, 1.0]	$\bar{Y}^3_{11}$	0.00	0.00	0.78	[0.0, 0.2)	[0.0, 0.2)	[0.6, 0.8)
$\bar{Y}^2_{12}$	0.75	0.45	0.00	[0.6, 0.8)	[0.4, 0.6)	[0.0, 0.2)	$\bar{Y}^1_{12}$	0.63	0.00	0.00	[0.6, 0.8)	[0.0, 0.2)	[0.0, 0.2)
$\bar{Y}^1_{13}$	0.61	0.47	0.12	[0.6, 0.8)	[0.4, 0.6)	[0.0, 0.2)	$\bar{Y}^2_{12}$	0.00	0.22	0.44	[0.0, 0.2)	[0.2, 0.4)	[0.4, 0.6)
$\bar{Y}^2_{13}$	0.34	0.66	0.71	[0.2, 0.4)	[0.6, 0.8)	[0.6, 0.8)	$\bar{Y}^3_{12}$	0.00	0.11	0.69	[0.0, 0.2)	[0.0, 0.2)	[0.6, 0.8)

Table 3 **The solution values ξ**

	J1-T3-M2				J1-T2-M3		
ξ^m_{jt}	$\mathbb{SP}1$	$\mathbb{SP}2$	$\mathbb{SP}3$	ξ^m_{jt}	$\mathbb{SP}4$	$\mathbb{SP}5$	$\mathbb{SP}6$
ξ^1_{11}	1	2	2	ξ^1_{11}	1	2	2
ξ^2_{11}	2	1	1	ξ^2_{11}	2	1	3
ξ^1_{12}	2	2	1	ξ^3_{11}	3	3	1
ξ^2_{12}	1	1	2	ξ^1_{12}	1	2	3
ξ^1_{13}	1	2	1	ξ^2_{12}	3	1	1
ξ^2_{13}	2	1	2	ξ^3_{12}	2	3	2

Table 4 **The hybrid solution values $\dot{Y}^1_{hybrid}$ from the Dantzig-Wolfe and column generation methods**

J1-T3-M2				J1-T2-M3			
$\dot{Y}^{m1}_{jt_{hybrid}}$	$\mathbb{SP}1$	$\mathbb{SP}2$	$\mathbb{SP}3$	$\dot{Y}^{m1}_{11_{hybrid}}$	$\mathbb{SP}4$	$\mathbb{SP}5$	$\mathbb{SP}6$
$\dot{Y}^{11}_{11_{hybrid}}$	1	0	0	$\dot{Y}^{11}_{11_{hybrid}}$	1	0	0
$\dot{Y}^{21}_{11_{hybrid}}$	0	1	1	$\dot{Y}^{21}_{11_{hybrid}}$	0	1	0
$\dot{Y}^{11}_{12_{hybrid}}$	0	0	1	$\dot{Y}^{31}_{11_{hybrid}}$	0	0	1
$\dot{Y}^{21}_{12_{hybrid}}$	1	0	0	$\dot{Y}^{11}_{12_{hybrid}}$	1	1	0
$\dot{Y}^{11}_{13_{hybrid}}$	1	1	1	$\dot{Y}^{21}_{12_{hybrid}}$	0	0	1
$\dot{Y}^{21}_{13_{hybrid}}$	0	0	1	$\dot{Y}^{31}_{12_{hybrid}}$	0	0	0

As shown in Table 5, we calculate the average of $\hat{Y}$ under each scenario of $(\bar{Y}, \xi, \dot{Y}^1_{hybrid})$ for each small-size problem and let the average be $\hat{Y}^{avg}$. As a result, we get a number of observations $(\hat{Y}^{avg}, \bar{Y}, \xi$, and $\dot{Y}^1_{hybrid})$ given at the right side of Table 5.

Table 5 **The solution values $\hat{Y}$, $\bar{Y}$, ξ, $\dot{Y}^1_{hybrid}$, and the modeling data**

	The solution values $\hat{Y}$, $\bar{Y}$, ξ, and $\dot{Y}^1_{hybrid}$				The modeling data			
	$\hat{Y}$	$\bar{Y}$	ξ	$\dot{Y}_{hybrid}$	$\hat{Y}^{avg}$	$\bar{Y}$	ξ	$\dot{Y}_{hybrid}$
$\mathbb{SP}1$	1	[0.6, 0.8)	1	1	0.67	[0.6, 0.8)	1	1
	0	[0.0, 0.2)	2	0	0	[0.0, 0.2)	2	0
	0	[0.0, 0.2)	2	0	1	[0.2, 0.4)	2	0
	1	[0.6, 0.8)	1	1				
	0	[0.6, 0.8)	1	1				
	1	[0.2, 0.4)	2	0				
$\mathbb{SP}2$	0	[0.2, 0.4)	2	0	0.5	[0.2, 0.4)	2	0
	1	[0.8, 1.0]	1	1	1	[0.8, 1.0]	1	1
	1	[0.2, 0.4)	2	0	0	[0.4, 0.6)	1	0
	0	[0.4, 0.6)	1	0	1	[0.4, 0.6)	2	1
	1	[0.4, 0.6)	2	1	0	[0.6, 0.8)	1	0
	0	[0.6, 0.8)	1	0				
$\mathbb{SP}3$	0	[0.2, 0.4)	2	0	0	[0.2, 0.4)	2	0
	1	[0.6, 0.8)	1	1	1	[0.6, 0.8)	1	1
	1	[0.8, 1.0]	1	1	1	[0.8, 1.0]	1	1
	0	[0.0, 0.2)	2	0	0	[0.0, 0.2)	2	0
	0	[0.0, 0.2)	1	1	0	[0.0, 0.2)	1	1
	1	[0.6, 0.8)	2	1	1	[0.6, 0.8)	2	1
$\mathbb{SP}4$	1	[0.6, 0.8)	1	1	1	[0.6, 0.8)	1	1
	0	[0.2, 0.4)	2	0	0	[0.2, 0.4)	2	0
	0	[0.0, 0.2)	3	0	1	[0.0, 0.2)	3	0
	1	[0.6, 0.8)	1	1	0	[0.0, 0.2)	2	0
	1	[0.0, 0.2)	3	0				
	0	[0.0, 0.2)	2	0				
$\mathbb{SP}5$	0	[0.0, 0.2)	2	0	0	[0.0, 0.2)	2	0
	1	[0.8, 1.0]	1	1	1	[0.8, 1.0]	1	1
	0	[0.0, 0.2)	3	0	0.5	[0.0, 0.2)	3	0
	0	[0.0, 0.2)	2	1	0	[0.0, 0.2)	2	1
	0	[0.2, 0.4)	1	0	0	[0.2, 0.4)	1	0
	1	[0.0, 0.2)	3	0				
$\mathbb{SP}6$	0	[0.0, 0.2)	2	0	0	[0.0, 0.2)	2	0
	0	[0.0, 0.2)	3	0	0	[0.0, 0.2)	3	0
	1	[0.6, 0.8)	1	1	1	[0.6, 0.8)	1	1
	0	[0.0, 0.2)	3	0	0	[0.4, 0.6)	1	1
	0	[0.4, 0.6)	1	1	1	[0.6, 0.8)	2	0
	1	[0.6, 0.8)	2	0				

- **Step IV: build generalized linear models by the GLM procedure in Statistical Analysis System (SAS).** The generalized linear models are built using the observations of $(\hat{Y}^{avg}, \bar{Y}, \xi, \dot{Y}^1_{hybrid})$ where $\hat{Y}^{avg}$ is the target variable and $\bar{Y}$, ξ, and $\dot{Y}^1_{hybrid}$ are the exploratory variables. We assume that the modeling output is as follows:

$$\varphi^1(Y) = 0.39 + \text{Match}[\bar{Y}] \begin{bmatrix} \text{“}[0.0, 0.2)\text{”} \Rightarrow -0.22 \\ \text{“}[0.2, 0.4)\text{”} \Rightarrow -0.08 \\ \text{“}[0.4, 0.6)\text{”} \Rightarrow -0.01 \\ \text{“}[0.6, 0.8)\text{”} \Rightarrow -0.01 \\ \text{“}[0.8, 1.0]\text{”} \Rightarrow 0.31 \end{bmatrix} + \text{Match}[\xi] \begin{bmatrix} \text{“1”} \Rightarrow 0.10 \\ \text{“2”} \Rightarrow -0.03 \\ \text{“3”} \Rightarrow -0.07 \end{bmatrix} + \text{Match}[\dot{Y}^1_{hybrid}] \begin{bmatrix} \text{“0”} \Rightarrow -0.06 \\ \text{“1”} \Rightarrow 0.06 \end{bmatrix}$$

where the values in quotation marks are the respective solution values associated with $\bar{Y}$, ξ, and $\dot{Y}^1_{hybrid}$, and the values next to “⇒” are the model coefficients.

- **Step V: obtain solution values $\bar{Y}$, ξ, and $\dot{Y}^1_{hybrid}$ for problems $\mathbb{P}$ and calculate $\varphi^1(Y)$ using the above equations for all setup variables.** We use a problem $\mathbb{P}$ that has a size J2-T2-M3 as an example. We assume its solution values $\bar{Y}$, ξ, and $\dot{Y}^1_{hybrid}$ are given at the left side of Table 6. $\varphi^1(Y)$ is calculated using the above equations and given at the right side of Table 6.

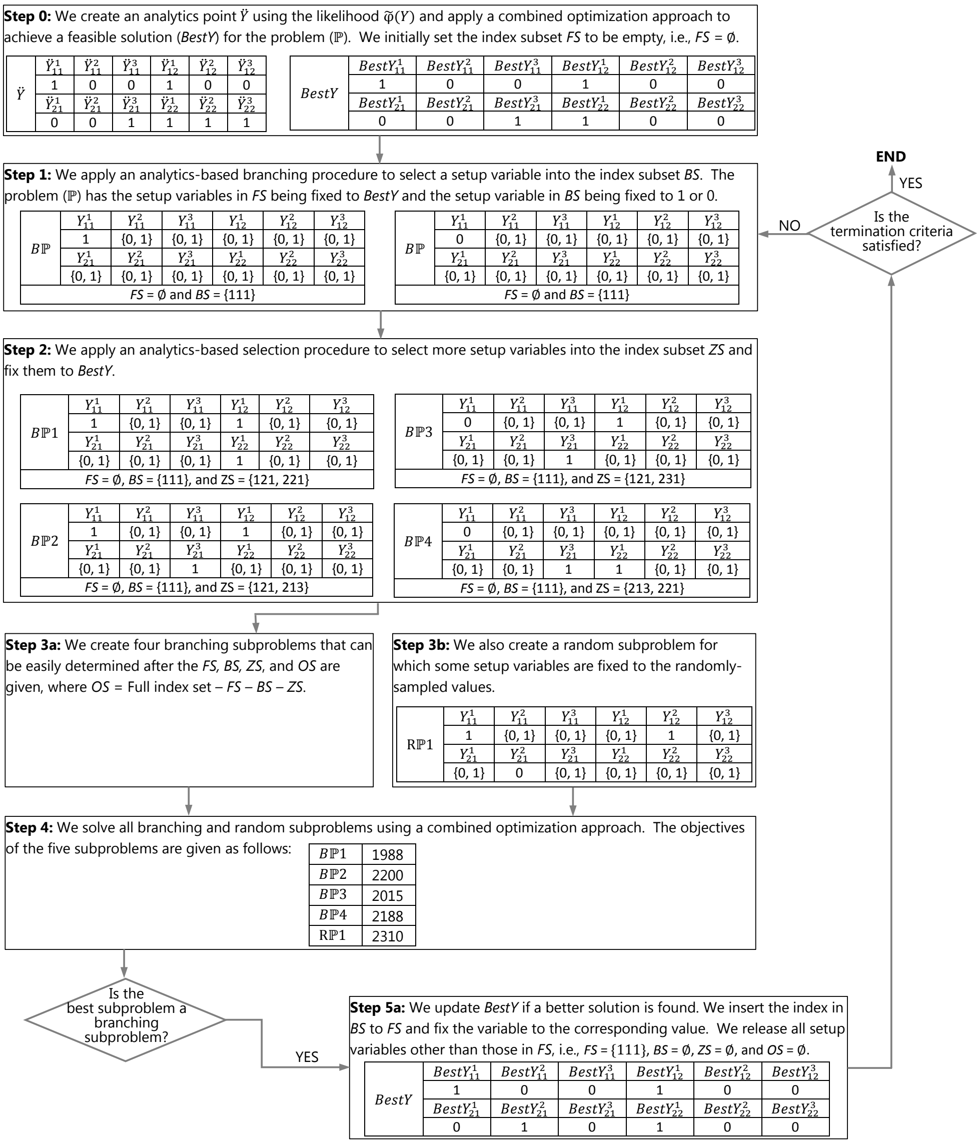
Table 6 **The solution values $\bar{Y}$, ξ, $\dot{Y}^1_{hybrid}$, and the insightful information $\varphi^1(Y)$**

Y^m_{jt}	$\bar{Y}^m_{jt}$	ξ^m_{jt}	$\dot{Y}^{m1}_{jt_{hybrid}}$	$\varphi^1(Y)$
Y^1_{11}	[0.8, 1.0]	1	1	$0.39 + 0.31 + 0.10 + 0.06 = 0.86$
Y^2_{11}	[0.2, 0.4)	2	0	$0.39 - 0.08 - 0.03 - 0.06 = 0.22$
Y^3_{11}	[0.4, 0.6)	3	0	$0.39 - 0.01 - 0.07 - 0.06 = 0.25$
Y^1_{12}	[0.6, 0.8)	1	1	$0.39 - 0.01 + 0.10 + 0.06 = 0.54$
Y^2_{12}	[0.0, 0.2)	3	0	$0.39 - 0.22 - 0.07 - 0.06 = 0.04$
Y^3_{12}	[0.2, 0.4)	2	0	$0.39 - 0.08 - 0.03 - 0.06 = 0.22$
Y^1_{21}	[0.0, 0.2)	1	0	$0.39 - 0.22 + 0.10 - 0.06 = 0.21$
Y^2_{21}	[0.0, 0.2)	2	1	$0.39 - 0.22 - 0.03 + 0.06 = 0.20$
Y^3_{21}	[0.4, 0.6)	3	1	$0.39 - 0.01 - 0.07 + 0.06 = 0.37$
Y^1_{22}	[0.8, 1.0]	2	1	$0.39 + 0.31 - 0.03 + 0.06 = 0.73$
Y^2_{22}	[0.6, 0.8)	3	1	$0.39 - 0.01 - 0.07 + 0.06 = 0.37$
Y^3_{22}	[0.2, 0.4)	1	0	$0.39 - 0.08 + 0.10 - 0.06 = 0.35$

From Table 6, we have that, at an optimal point, the setup variable Y^1_{11} is most likely (86%) 1 when the variable has solution values ([0.8, 1.0], 1, 1) for $(\bar{Y}, \xi, \dot{Y}_{hybrid})$. Meanwhile, the setup variable Y^2_{12} is least likely (4%) 1 when the variable has solution values ([0.0, 0.2], 3, 0) for $(\bar{Y}, \xi, \dot{Y}_{hybrid})$. We note that the above description is just for one iteration of the modeling procedure. Similarly, we perform such modeling procedure for 20 iterations for which $\dot{Y}_{hybrid}$ is refreshed during each iteration. We take the average of $\varphi^l(Y)$, $\forall\ l \in \{1, 2, ..., 20\}$, i.e., $\tilde{\varphi}(Y)$ $(=\frac{\sum_{l\in\{1,...,20\}} \varphi^l(Y)}{20})$ and use $\tilde{\varphi}(Y)$ as the likelihood information for the analytics branching and selection method.

1.2 An illustrative example for the analytics branching and selection method

We will next use the above illustrative example to describe the procedure of the analytics branching and selection method in Figure 1.



Step 0: We create an analytics point $\ddot{Y}$ using the likelihood $\tilde{\varphi}(Y)$ and apply a combined optimization approach to achieve a feasible solution (*BestY*) for the problem ($\mathbb{P}$). We initially set the index subset *FS* to be empty, i.e., *FS* = ∅.

$\ddot{Y}$	$\ddot{Y}_{11}^1$	$\ddot{Y}_{11}^2$	$\ddot{Y}_{11}^3$	$\ddot{Y}_{12}^1$	$\ddot{Y}_{12}^2$	$\ddot{Y}_{12}^3$
	1	0	0	1	0	0
	$\ddot{Y}_{21}^1$	$\ddot{Y}_{21}^2$	$\ddot{Y}_{21}^3$	$\ddot{Y}_{22}^1$	$\ddot{Y}_{22}^2$	$\ddot{Y}_{22}^3$
	0	0	1	1	1	1

$BestY$	$BestY_{11}^1$	$BestY_{11}^2$	$BestY_{11}^3$	$BestY_{12}^1$	$BestY_{12}^2$	$BestY_{12}^3$
	1	0	0	1	0	0
	$BestY_{21}^1$	$BestY_{21}^2$	$BestY_{21}^3$	$BestY_{22}^1$	$BestY_{22}^2$	$BestY_{22}^3$
	0	0	1	1	0	0

Step 1: We apply an analytics-based branching procedure to select a setup variable into the index subset *BS*. The problem ($\mathbb{P}$) has the setup variables in *FS* being fixed to *BestY* and the setup variable in *BS* being fixed to 1 or 0.

$B\mathbb{P}$	Y_{11}^1	Y_{11}^2	Y_{11}^3	Y_{12}^1	Y_{12}^2	Y_{12}^3
	1	{0, 1}	{0, 1}	{0, 1}	{0, 1}	{0, 1}
	Y_{21}^1	Y_{21}^2	Y_{21}^3	Y_{22}^1	Y_{22}^2	Y_{22}^3
	{0, 1}	{0, 1}	{0, 1}	{0, 1}	{0, 1}	{0, 1}
FS = ∅ and *BS* = {111}						

$B\mathbb{P}$	Y_{11}^1	Y_{11}^2	Y_{11}^3	Y_{12}^1	Y_{12}^2	Y_{12}^3
	0	{0, 1}	{0, 1}	{0, 1}	{0, 1}	{0, 1}
	Y_{21}^1	Y_{21}^2	Y_{21}^3	Y_{22}^1	Y_{22}^2	Y_{22}^3
	{0, 1}	{0, 1}	{0, 1}	{0, 1}	{0, 1}	{0, 1}
FS = ∅ and *BS* = {111}						

Step 2: We apply an analytics-based selection procedure to select more setup variables into the index subset *ZS* and fix them to *BestY*.

$B\mathbb{P}1$	Y_{11}^1	Y_{11}^2	Y_{11}^3	Y_{12}^1	Y_{12}^2	Y_{12}^3
	1	{0, 1}	{0, 1}	1	{0, 1}	{0, 1}
	Y_{21}^1	Y_{21}^2	Y_{21}^3	Y_{22}^1	Y_{22}^2	Y_{22}^3
	{0, 1}	{0, 1}	{0, 1}	1	{0, 1}	{0, 1}
FS = ∅, *BS* = {111}, and ZS = {121, 221}						

$B\mathbb{P}3$	Y_{11}^1	Y_{11}^2	Y_{11}^3	Y_{12}^1	Y_{12}^2	Y_{12}^3
	0	{0, 1}	{0, 1}	1	{0, 1}	{0, 1}
	Y_{21}^1	Y_{21}^2	Y_{21}^3	Y_{22}^1	Y_{22}^2	Y_{22}^3
	{0, 1}	{0, 1}	1	{0, 1}	{0, 1}	{0, 1}
FS = ∅, *BS* = {111}, and ZS = {121, 231}						

$B\mathbb{P}2$	Y_{11}^1	Y_{11}^2	Y_{11}^3	Y_{12}^1	Y_{12}^2	Y_{12}^3
	1	{0, 1}	{0, 1}	1	{0, 1}	{0, 1}
	Y_{21}^1	Y_{21}^2	Y_{21}^3	Y_{22}^1	Y_{22}^2	Y_{22}^3
	{0, 1}	{0, 1}	1	{0, 1}	{0, 1}	{0, 1}
FS = ∅, *BS* = {111}, and ZS = {121, 213}						

$B\mathbb{P}4$	Y_{11}^1	Y_{11}^2	Y_{11}^3	Y_{12}^1	Y_{12}^2	Y_{12}^3
	0	{0, 1}	{0, 1}	{0, 1}	{0, 1}	{0, 1}
	Y_{21}^1	Y_{21}^2	Y_{21}^3	Y_{22}^1	Y_{22}^2	Y_{22}^3
	{0, 1}	{0, 1}	1	1	{0, 1}	{0, 1}
FS = ∅, *BS* = {111}, and ZS = {213, 221}						

Step 3a: We create four branching subproblems that can be easily determined after the *FS, BS, ZS*, and *OS* are given, where *OS* = Full index set – *FS* – *BS* – *ZS*.

Step 3b: We also create a random subproblem for which some setup variables are fixed to the randomly-sampled values.

$R\mathbb{P}1$	Y_{11}^1	Y_{11}^2	Y_{11}^3	Y_{12}^1	Y_{12}^2	Y_{12}^3
	1	{0, 1}	{0, 1}	{0, 1}	1	{0, 1}
	Y_{21}^1	Y_{21}^2	Y_{21}^3	Y_{22}^1	Y_{22}^2	Y_{22}^3
	{0, 1}	0	{0, 1}	{0, 1}	{0, 1}	{0, 1}

Step 4: We solve all branching and random subproblems using a combined optimization approach. The objectives of the five subproblems are given as follows:

$B\mathbb{P}1$	1988
$B\mathbb{P}2$	2200
$B\mathbb{P}3$	2015
$B\mathbb{P}4$	2188
$R\mathbb{P}1$	2310

Step 5a: We update *BestY* if a better solution is found. We insert the index in *BS* to *FS* and fix the variable to the corresponding value. We release all setup variables other than those in *FS*, i.e., *FS* = {111}, *BS* = ∅, *ZS* = ∅, and *OS* = ∅.

$BestY$	$BestY_{11}^1$	$BestY_{11}^2$	$BestY_{11}^3$	$BestY_{12}^1$	$BestY_{12}^2$	$BestY_{12}^3$
	1	0	0	1	0	0
	$BestY_{21}^1$	$BestY_{21}^2$	$BestY_{21}^3$	$BestY_{22}^1$	$BestY_{22}^2$	$BestY_{22}^3$
	0	1	0	1	0	0

Figure 1 **An illustrative example for the analytics branching and selection method**

- **Step 0:** We create an analytics point $\ddot{Y}$ using the likelihood $\tilde{\varphi}(Y)$. In this example, we assume $\tilde{\varphi}(Y)$ is the same as $\varphi^1(Y)$ given above and set κ to 0.3 to calculate the analytics point $\ddot{Y}$ that is given in Step 0 of Figure 1. We apply a combined optimization approach to achieve a feasible solution ($BestY$) for the problem ($\mathbb{P}$) that is given in Step 0 of Figure 1. We initially set the index subset FS to be empty, i.e., $FS = \emptyset$. Go to Step 1.

- **Step 1:** We apply an analytics-based branching procedure to select a setup variable into the index subset BS. The problem ($\mathbb{P}$) has the setup variables in FS being fixed to $BestY$ and the setup variable in BS being fixed to 1 or 0. In this example, we branch the setup variable Y^1_{11} at first because it has the largest value at $\tilde{\varphi}(Y)$. Go to Step 2.
- **Step 2:** We apply an analytics-based selection procedure to select more setup variables into the index subset ZS and fix them to $BestY$. In this example, we select two variables into the index subset ZS and have 4 different selections of ZS. Go to Step 3.
- **Step 3:** We therefore have four branching subproblems that can be easily determined after the FS, BS, ZS, and OS are given, where OS = the full index set of the setup variables $- FS - BS - ZS$. The definition of the branching subproblem is given in the main manuscript of this paper. Meanwhile, we also create a random subproblem for which some setup variables are fixed to the randomly-sampled values. Go to Step 4.
- **Step 4:** We solve all branching and random subproblems using a combined optimization approach. In this example, we assume the objectives of the five subproblems are given in the table of Step 4 of Figure 1. Go to Step 5.
- **Step 5:** Because the best subproblem is a branching subproblem, we insert the index in BS to FS and fix the variable to the corresponding value. We release all setup variables other than those in FS, i.e., $FS = \{111\}, BS = \emptyset, ZS = \emptyset$, and $OS = \emptyset$. We also assume a better solution is found such that we update $BestY$ that is given in Step 5 of Figure 1. Go to Step 6.
- **Step 6:** We terminate the method either when the computation time is larger than the preset time-limit (T_{lim}), or all setup variables Y have been fixed, otherwise go to Step 1. In this example, we have $FS = \{111\}$ for the next iteration.

2. The per-item Dantzig–Wolfe decomposition

Here we describe the LP-relaxation of the Dantzig–Wolfe decomposition for the three formulations. For LSNM, the definitions of all parameters and variables are given in the main manuscript. We denote the LP-relaxation of the master problem by MP-D$^{LR}_{LSNM\_ID}$ that is given as follows:

$$\min \sum_{j\in J}\sum_{t\in T}\sum_{m\in\mathcal{M}_j}\sum_{o\in O^j} sc^m_j \cdot Y^{mo}_{jt} \cdot \lambda_{jo} + \sum_{j\in J}\sum_{t\in T}\sum_{m\in\mathcal{M}_j}\sum_{o\in O^j} pc^m_j \cdot X^{mo}_{jt} \cdot \lambda_{jo} + \sum_{j\in J}\sum_{t\in T}\sum_{o\in O^j} hc_j \cdot I^o_{jt} \cdot \lambda_{jo}$$

$$
\begin{aligned}
\text{Subject to:}\quad & \sum_{j\in\mathcal{J}_m}\sum_{o\in O^j}(pt_j^m\cdot X_{jt}^{mo}+st_j^m\cdot Y_{jt}^{mo})\cdot\lambda_{jo}\le C_t^m,\ \forall\, t\in T,\ m\in M\\
& \sum_{o\in O^j}\lambda_{jo}=1,\ \forall\, j\in J \qquad\qquad (\text{MP-D}_{LSNM\_ID}^{LR})\\
& \lambda_{jo}\ge 0,\ \forall\, j\in J,\ o\in O^j
\end{aligned}
$$

The pricing subproblem is same as the subproblem SP-D$_{LSNM_ID}$ given in the main manuscript.

For FLNM, besides O^j, Y_{jt}^{mo}, λ_{jo}, β_t^m and ρ_j defined as previously, U_{jts}^{mo} is defined as the amount of production of item j on machine m in period t used to satisfy demand in period s for a given possible setup schedule $o\in O^j$. FLNM is decomposed into several pricing subproblems (SP-D$_{FLNM\_ID}$) and a master problem (MP-D$_{FLNM\_ID}^{LR}$) as follows:

$$
\begin{aligned}
\min\quad & \sum_{t\in T}\sum_{m\in\mathcal{M}_j} sc_j^m\cdot Y_{jt}^m+\sum_{t\in T}\sum_{s\in\{t,\dots,|T|\}}\sum_{m\in\mathcal{M}_j} vc_j^m\cdot U_{jts}^m\\
& -\sum_{t\in T}\sum_{s\in\{t,\dots,|T|\}}\sum_{m\in\mathcal{M}_j}\beta_t^m\cdot pt_j^m\cdot U_{jts}^m-\sum_{t\in T}\sum_{m\in\mathcal{M}_j}\beta_t^m\cdot st_j^m\cdot Y_{jt}^m-\rho_j,\quad \forall\, j\in J\\
\text{Subject to:}\quad & \sum_{t=1}^{s}\sum_{m\in\mathcal{M}_j}U_{jts}^m=d_{js},\ \forall\, s\in T\\
& U_{jts}^m\le d_{js}\cdot Y_{jt}^m,\ \forall\, t\in T,\ s\in\{t,\dots,|T|\},\ m\in\mathcal{M}_j \qquad\qquad (\text{SP-D}_{FLNM\_ID})\\
& \sum_{m\in\mathcal{M}_j}Y_{jt}^m\le mm_j,\ \forall\, t\in T\\
& U_{jts}^m\ge 0,\quad Y_{jt}^m\in\{0,1\},\ \forall\, t\in T,\ s\in\{t,\dots,|T|\},\ m\in\mathcal{M}_j
\end{aligned}
$$

$$
\begin{aligned}
\min\quad & \sum_{j\in J}\sum_{t\in T}\sum_{m\in\mathcal{M}_j}\sum_{o\in O^j} sc_j^m\cdot Y_{jt}^{mo}\cdot\lambda_{jo}+\sum_{j\in J}\sum_{t\in T}\sum_{s\in\{t,\dots,|T|\}}\sum_{m\in\mathcal{M}_j}\sum_{o\in O^j} vc_j^m\cdot U_{jts}^{mo}\cdot\lambda_{jo}\\
\text{Subject to:}\quad & \sum_{j\in\mathcal{J}_m}\sum_{s\in\{t,\dots,|T|\}}\sum_{o\in O^j} pt_j^m\cdot U_{jts}^{mo}\cdot\lambda_{jo}+\sum_{j\in\mathcal{J}_m}\sum_{o\in O^j} st_j^m\cdot Y_{jt}^{mo}\cdot\lambda_{jo}\le C_t^m,\ \forall\, t\in T,\ m\in M\\
& \sum_{o\in O^j}\lambda_{jo}=1,\ \forall\, j\in J \qquad\qquad (\text{MP-D}_{FLNM\_ID}^{LR})\\
& \lambda_{jo}\ge 0,\ \forall\, j\in J,\ o\in O^j
\end{aligned}
$$

For SPNM, W_{jtq}^{mo} is defined as percentage of production of item j on machine m in period t used to satisfy the accumulated demand for item j from period t to period q for a given possible setup schedule $o \in O^j$. SPNM is decomposed into several pricing subproblems (SP-D$_{SPNM\_ID}$) and a master problem (MP-D$^{LR}_{SPNM\_ID}$) as follows:

$$
\begin{aligned}
\min\quad & \sum_{t\in T}\sum_{m\in\mathcal{M}_j} sc_j^m \cdot Y_{jt}^m + \sum_{t\in T}\sum_{s\in\{t,\dots,|T|\}}\sum_{m\in\mathcal{M}_j} tc_{jts}^m \cdot W_{jts}^m \\
& -\sum_{t\in T}\sum_{q\in\{t,\dots,|T|\}}\sum_{s\in\{q,\dots,|T|\}}\sum_{m\in\mathcal{M}_j} \beta_t^m \cdot pt_j^m \cdot d_{jq}\cdot w_{jts} \\
& -\sum_{t\in T}\sum_{m\in\mathcal{M}_j} \beta_t^m \cdot st_j^m \cdot Y_{jt}^m - \rho_j, \quad \forall\, j\in J \\
\text{Subject to:}\quad & \sum_{s\in T}\sum_{m\in\mathcal{M}_j} W_{j1s}^m = 1, \\
& \sum_{s=1}^{t-1}\sum_{m\in\mathcal{M}_j} W_{js(t-1)}^m = \sum_{s\in\{t,\dots,|T|\}}\sum_{m\in\mathcal{M}_j} W_{jts}^m,\ \forall\, t\in\{2,\dots,|T|\} \qquad (\text{SP-D}_{SPNM\_ID}) \\
& \sum_{s\in\{t,\dots,|T|\}} W_{jts}^m \le Y_{jt}^m,\ \forall\, t\in T,\ m\in M \\
& \sum_{m\in\mathcal{M}_j} Y_{jt}^m \le mm_j,\ \forall\, t\in T \\
& W_{jts}^m \ge 0,\quad Y_{jt}^m \in\{0,1\},\ \forall\, t\in T,\ s\in\{t,\dots,|T|\},\ m\in\mathcal{M}_j
\end{aligned}
$$

$$
\begin{aligned}
\min\quad & \sum_{j\in J}\sum_{t\in T}\sum_{m\in\mathcal{M}_j}\sum_{o\in O^j} sc_j^m\cdot Y_{jt}^{mo}\cdot\lambda_{jo} + \sum_{j\in J}\sum_{t\in T}\sum_{s\in\{t,\dots,|T|\}}\sum_{m\in\mathcal{M}_j}\sum_{o\in O^j} tc_{jts}^m\cdot W_{jts}^{mo}\cdot\lambda_{jo} \\
\text{Subject to:}\quad & \sum_{j\in\mathcal{J}_m}\sum_{q\in\{t,\dots,|T|\}}\sum_{s\in\{q,\dots,|T|\}}\sum_{o\in O^j} pt_j^m\cdot d_{jq}\cdot W_{jts}^{mo}\cdot\lambda_{jo} + \sum_{j\in\mathcal{J}_m}\sum_{o\in O^j} st_j^m\cdot Y_{jt}^{mo}\cdot\lambda_{jo} \le C_t^m, \\
& \qquad\qquad \forall\, t\in T,\ m\in M \\
& \sum_{o\in O^j}\lambda_{jo} = 1,\ \forall\, j\in J \qquad (\text{MP-D}^{LR}_{SPNM\_ID}) \\
& \lambda_{jo}\ge 0,\ \forall\, j\in J,\ o\in O^j
\end{aligned}
$$

For the per-item Dantzig–Wolfe decomposition of FLNM and SPNM, the integrality must be imposed on the original setup variables Y_{jt}^m for the master problems. We skip the formulations for brevity.

3. The per-period Dantzig–Wolfe decomposition

Here we describe the LP-relaxation of the per-period Dantzig–Wolfe decomposition for LSNM, FLNM, and SPNM. For LSNM, we define K^t as the set of all possible setup schedules for period t where $K^t = \{(Y_{1t}^1,\dots,Y_{jt}^m,\dots)|Y_{jt}^m\in\{0,1\}\}$, $\forall\ j\in J,\ m\in\mathcal{M}_j$. For a given possible setup schedule $k \in K^t$, we define Y_{jt}^{mk} as the setup value for item j on machine m in period t, X_{jt}^{mk} as the amount of production of item j on machine m in period t, I_{jt}^{k} as the amount of inventories of item j at the end of period t, and π_{tk} as the fraction of production of period t. In addition, γ_{jt} ($\forall\ j \in J,\ t \in T$) represents dual variables related with the demand-satisfaction constraints; and δ_t ($\forall\ t \in T$) represents the dual variables for the convexity constraints. With these definitions, LSNM is decomposed into T pricing subproblems (SP-D$_{LSNM\_PD}$) and a master problem (MP-D$^{LR}_{LSNM\_PD}$) as follows:

$$\begin{aligned}
\min \quad & \sum_{j\in J}\sum_{m\in\mathcal{M}_j} sc_j^m \cdot Y_{jt}^m + \sum_{j\in J}\sum_{m\in\mathcal{M}_j} pc_j^m \cdot X_{jt}^m + \sum_{j\in J} hc_j \cdot I_{jt} \\
& \qquad - \sum_{j\in J}\sum_{m\in\mathcal{M}_j} \gamma_{jt} \cdot X_{jt}^m + \sum_{j\in J} (\gamma_{jt} - \gamma_{j(t+1)}) \cdot I_{jt} - \delta_t, \ \ \forall\, t \in T \\
\text{Subject to:} \quad & \sum_{j\in\mathcal{J}_m} pt_j^m \cdot X_{jt}^m + \sum_{j\in\mathcal{J}_m} st_j^m \cdot Y_{jt}^m \le C_t^m, \ \forall\, m \in M \\
& X_{jt}^m \le BM_{jt}^m \cdot Y_{jt}^m, \ \forall\, j \in J,\ m \in \mathcal{M}_j \qquad \text{(SP-D}_{LSNM\_PD}\text{)} \\
& \sum_{m\in\mathcal{M}_j} Y_{jt}^m \le mm_j, \ \forall\, j \in J \\
& I_{jt} \le \sum_{q=t}^{T} d_{jq}, \ \forall\, j \in J \\
& X_{jt}^m,\ I_{jt} \ge 0,\ Y_{jt}^m \in \{0,1\}, \ \forall\, j \in J,\ m \in \mathcal{M}_j
\end{aligned}$$

$$\begin{aligned}
\min \quad & \sum_{j\in J}\sum_{t\in T}\sum_{m\in\mathcal{M}_j}\sum_{k\in K^t} sc_j^m \cdot Y_{jt}^{mk} \cdot \pi_{tk} + \sum_{j\in J}\sum_{t\in T}\sum_{m\in\mathcal{M}_j}\sum_{k\in K^t} pc_j^m \cdot X_{jt}^{mk} \cdot \pi_{tk} \\
& \qquad + \sum_{j\in J}\sum_{t\in T}\sum_{k\in K^t} hc_j \cdot I_{jt}^k \cdot \pi_{tk} \\
\text{Subject to:} \quad & \sum_{k\in K^t}\sum_{m\in\mathcal{M}_j} X_{jt}^{mk} \cdot \pi_{tk} + \sum_{k\in K^t} I_{j(t-1)}^k \cdot \pi_{(t-1)k} - \sum_{k\in K^t} I_{jt}^k \cdot \pi_{tk} = d_{jt}, \ \forall\, j \in J,\ t \in T \\
& \sum_{k\in K^t} \pi_{tk} = 1, \ \forall\, t \in T \qquad \text{(MP-D}^{LR}_{LSNM\_PD}\text{)} \\
& \pi_{tk} \ge 0, \ \forall\, t \in T,\ k \in K^t
\end{aligned}$$

For FLNM, U_{jts}^{mk} is defined as production of item j on machine m in period t used to satisfy the demand for item j in period s for a given possible setup schedule $k \in K^t$.

FLNM is decomposed into T pricing subproblems (SP-D$_{FLNM\_PD}$) and a master problem (MP-D$^{LR}_{FLNM\_PD}$) as follows:

$$\begin{aligned}
\min\quad & \sum_{j\in J}\sum_{m\in\mathcal{M}_j} sc^m_j\cdot Y^m_{jt} + \sum_{j\in J}\sum_{s\in\{t,\dots,|T|\}}\sum_{m\in\mathcal{M}_j} vc^m_j\cdot U^m_{jts} \\
& \qquad\qquad - \sum_{j\in J}\sum_{q\in\{t,\dots,|T|\}}\sum_{m\in\mathcal{M}_j}\gamma_{jq}\cdot U^m_{jtq} - \delta_t,\quad \forall\, t\in T \\
\text{Subject to:}\quad & \sum_{j\in\mathcal{J}_m}\sum_{s\in\{t,\dots,|T|\}} pt^m_j\cdot U^m_{jts} + \sum_{j\in\mathcal{J}_m} st^m_j\cdot Y^m_{jt} \le C^m_t,\ \forall\, m\in M \\
& U^m_{jts}\le d_{js}\cdot Y^m_{jt},\ \forall\, j\in J,\ s\in\{t,\dots,|T|\} \\
& \sum_{m\in\mathcal{M}_j} Y^m_{jt}\le mm_j,\ \forall\, j\in J,\ m\in\mathcal{M}_j \\
& U^m_{jts}\ge 0,\ Y^m_{jt}\in\{0,1\},\ \forall\, j\in J,\ t\in T,\ s\in\{t,\dots,|T|\},\ m\in\mathcal{M}_j
\end{aligned}\tag{SP-D$_{FLNM\_PD}$}$$

$$\begin{aligned}
\min\quad & \sum_{j\in J}\sum_{t\in T}\sum_{m\in\mathcal{M}_j}\sum_{k\in K^t} sc^m_j\cdot Y^{mk}_{jt}\cdot\pi_{tk} + \sum_{j\in J}\sum_{t\in T}\sum_{s\in\{t,\dots,|T|\}}\sum_{m\in\mathcal{M}_j}\sum_{k\in K^t} vc^m_j\cdot U^{mk}_{jts}\cdot\pi_{tk} \\
\text{Subject to:}\quad & \sum_{q=1}^{t}\sum_{m\in\mathcal{M}_j}\sum_{k\in K^q} U^{mk}_{jqt}\cdot\pi_{qk} = d_{jt},\ \forall\, j\in J,\ t\in T \\
& \sum_{k\in K^t}\pi_{tk}=1,\ \forall\, t\in T \\
& \pi_{tk}\ge 0,\ \forall\, t\in T,\ k\in K^t
\end{aligned}\tag{MP-D$^{LR}_{FLNM\_PD}$}$$

For SPNM, W^{mk}_{jtq} is defined as percentage of production of item j on machine m in period t used to satisfy the accumulated demand for item j from period t to period q for a given possible setup schedule $k\in K^t$. SPNM is decomposed into T pricing subproblems (SP-D$_{SPNM\_PD}$) and a master problem (MP-D$^{LR}_{SPNM\_PD}$) as follows:

$$\begin{aligned}
\min\quad & \sum_{j\in J}\sum_{m\in\mathcal{M}_j} sc^m_j\cdot Y^m_{jt} + \sum_{j\in J}\sum_{s\in\{t,\dots,|T|\}}\sum_{m\in\mathcal{M}_j} tc^m_{jts}\cdot W^m_{jts} \\
& \quad - \sum_{j\in J}\sum_{q\in\{t,\dots,|T|\}}\sum_{m\in\mathcal{M}_j}\gamma_{jt}\cdot w^m_{jtq} + \sum_{j\in J}\sum_{q\in\{t,\dots,|T|-1\}}\sum_{m\in\mathcal{M}_j}\gamma_{j(q+1)}\cdot w^m_{jtq} \\
& \quad - \delta_t,\quad \forall\, t\in T \\
\text{Subject to:}\quad & \sum_{j\in\mathcal{J}_m}\sum_{s\in\{t,\dots,|T|\}}\sum_{q\in\{s,\dots,|T|\}} pt^m_j\cdot d_{js}\cdot W^m_{jtq} + \sum_{j\in\mathcal{J}_m} st^m_j\cdot Y^m_{jt}\le C^m_t,\ \forall\, m\in M \\
& \sum_{s\in\{t,\dots,|T|\}} W^m_{jts}\le Y^m_{jt},\ \forall\, j\in J,\ m\in\mathcal{M}_j
\end{aligned}\tag{SP-D$_{SPNM\_PD}$}$$

$$\sum_{m\in\mathcal{M}_j} Y_{jt}^m \le mm_j,\ \forall\, j\in J$$

$$W_{jts}^m \ge 0,\ Y_{jt}^m\in\{0,1\},\ \forall\, j\in J,\ t\in T,\ s\in\{t,...,|T|\},\ m\in\mathcal{M}_j$$

$$\begin{aligned}
\min\quad & \sum_{j\in J}\sum_{t\in T}\sum_{m\in\mathcal{M}_j}\sum_{k\in K^t} sc_j^m\cdot Y_{jt}^{mk}\cdot\pi_{tk} + \sum_{j\in J}\sum_{t\in T}\sum_{s\in\{t,...,|T|\}}\sum_{m\in\mathcal{M}_j}\sum_{k\in K^t} tc_{jts}^m\cdot W_{jts}^{mk}\cdot\pi_{tk}\\
\text{Subject to:}\quad & \sum_{q\in T}\sum_{m\in\mathcal{M}_j}\sum_{k\in K^1} W_{j1q}^{mk}\cdot\pi_{1k} = 1,\ \forall\, j\in J\\
& \sum_{q=1}^{t-1}\sum_{m\in\mathcal{M}_j}\sum_{k\in K^q} W_{jq(t-1)}^{mk}\cdot\pi_{qk} = \sum_{q\in\{t,...,|T|\}}\sum_{m\in\mathcal{M}_j}\sum_{k\in K^t} W_{jtq}^{mk}\cdot\pi_{tk},\ \forall\, j\in J,\ t\in\{2,...,|T|\}\\
& \sum_{k\in K^t}\pi_{tk} = 1,\ \forall\, t\in T \qquad\qquad (\text{MP-D}_{SPNM\_PD}^{LR})\\
& \pi_{tk}\ge 0,\ \forall\, t\in T,\ k\in K^t
\end{aligned}$$

For the per-period Dantzig–Wolfe decomposition of LSNM, FLNM and SPNM, the integrality must be imposed on the original setup variables Y_{jt}^m for the master problems. We skip the formulations for brevity.

4. The proof procedures of the propositions

4.1 The proof procedure of Proposition 1: $\mathcal{O}^j = \{(X_{jt}^m, I_{jt}, Y_{jt}^m)\in(2),(4)-(7);\ I_{j,t-1}\cdot X_{jt}^m = 0,\ \forall\, j\in J,\ t\in T,\ m\in\mathcal{M}_j\}$

Proof. See Degraeve and Jans (2007), Proposition 1, for the multi-item capacitated lot sizing problem with setup times. The procedure is extendable to the CMLS-NM problem studied in this paper and we skip it for brevity. ☐

4.2 The proof procedure of Proposition 2: $|\mathcal{O}^j| = (\sum_{n=0}^{mm_j} C_{\mathcal{M}_j}^n\cdot 2^n)^{|T|}$.

Proof. For an extreme point, there are two possibilities ($X_{jt}^m = 0$ and $X_{jt}^m > 0$) for the production variable when $Y_{jt}^m = 1$; and there is one possibility ($X_{jt}^m = 0$) for the production variable when $Y_{jt}^m = 0$. Herein, for each of the $|T|$ periods, there are exactly $C_{\mathcal{M}_j}^n\cdot 2^n\cdot 1^{mm_j-n}$ possibilities for the setup and production variable under the scenario that $\sum_{m\in\mathcal{M}_j} Y_{jt}^m = n$, $\forall\, n\in\{0,...,mm_j\}$. Combining the possibilities over the $|T|$ periods gives $(C_{\mathcal{M}_j}^0\cdot 2^0 + C_{\mathcal{M}_j}^1\cdot 2^1 + C_{\mathcal{M}_j}^2\cdot 2^2 + ... + C_{\mathcal{M}_j}^{mm_j}\cdot 2^{mm_j})^{|T|}$ extreme points in total. ☐

We note that Degraeve and Jans (2007) discussed a similar proposition for the single-machine multi-item lot sizing problem.

4.3 The proof procedure of Proposition 3: $|O^j| = (\sum_{n=0}^{mm_j} C^n_{\mathcal{M}_j})^{|T|}$.

Proof. Because only dominant production plans are considered in O^j, there is one possibility ($X^m_{jt} > 0$) for the production variable when $Y^m_{jt} = 1$; and there is one possibility ($X^m_{jt} = 0$) for the production variable when $Y^m_{jt} = 0$ for an extreme point. Herein, for each of the $|T|$ periods, there are exactly $C^n_{\mathcal{M}_j} \cdot 1^n \cdot 1^{mm_j - n}$ possibilities for the setup and production variable under the scenario that $\sum_{m\in\mathcal{M}_j} Y^m_{jt} = n, \ \forall\ n \in \{0, ..., mm_j\}$. Combining the possibilities over the $|T|$ periods gives $(C^0_{\mathcal{M}_j} + C^1_{\mathcal{M}_j} + C^2_{\mathcal{M}_j} + ... + C^{mm_j}_{\mathcal{M}_j})^{|T|}$ extreme points in total. □

4.4 The proof procedure of Proposition 4: $\mathbf{LB}_{LSNM-PD} \le \mathbf{LB}_{FLNM-PD}$

We prove the proposition by showing that a feasible solution $(\breve{U}^m_{jtq}, \breve{Y}^m_{jt})$ of the per-period decomposition of FLNM corresponds to a feasible solution $(\breve{X}^m_{jt}, \breve{I}_{jt}, \breve{Y}^m_{jt})$ of the per-period decomposition of LSNM with the same objective value but the reverse is not necessarily true.

If $(\breve{U}^m_{jtq}, \breve{Y}^m_{jt})$ is a feasible solution, it means that a) the solution is contained in the convex hull defined by constraints (10)-(14) and b) it satisfies constraints (9). Herein, $(\breve{U}^m_{jtq}, \breve{Y}^m_{jt})$ can be expressed as a convex combination of all extreme points $(U^{mk}_{jtq}, Y^{mk}_{jt})$ of the convex hull of constraints (10)-(14):

$U^m_{jtq} = \sum\limits_{k\in K^t} \pi_{tk} \cdot U^{mk}_{jtq}$ and $Y^m_{jt} = \sum\limits_{k\in K^t} \pi_{tk} \cdot Y^{mk}_{jt}, \ \forall\ j \in J,\ t \in T,\ q \in \{t, ..., |T|\},\ m \in \mathcal{M}_j$.

Projecting the feasible point $(\breve{U}^m_{jtq}, \breve{Y}^m_{jt})$ onto the $(X,\ I,\ Y)$-space, we have the corresponding solution:

$$\breve{X}^m_{jt} = \sum_{q\in\{t,...,|T|\}} \breve{U}^m_{jtq} = \sum_{q\in\{t,...,|T|\}} \sum_{k\in K^t} \pi_{tk} \cdot U^{mk}_{jtq},\ \forall\ j \in J,\ t \in T,\ m \in \mathcal{M}_j$$

$$\breve{I}_{jt} = \sum_{q=1}^{t} \sum_{s=t+1}^{|T|} \sum_{m\in\mathcal{M}_j} \breve{U}^m_{jqs} = \sum_{q=1}^{t} \sum_{s=t+1}^{|T|} \sum_{m\in\mathcal{M}_j} \sum_{k\in K^t} \pi_{tk} \cdot U^{mk}_{jqs},\ \forall\ j \in J,\ t \in T$$

$$\breve{Y}^m_{jt} = \sum_{k\in K^t} \pi_{tk} \cdot \breve{Y}^{mk}_{jt},\ \forall\ j \in J,\ t \in T,\ m \in \mathcal{M}_j$$

For the extreme points $(U^{mk}_{jtq}, Y^{mk}_{jt})$ of the convex hull of constraints (10)-(14), we, by projection, have the corresponding solution in the $(X,\ I,\ Y)$-space:

$$X^{mk}_{jt} = \sum_{q\in\{t,...,|T|\}} U^{mk}_{jtq},\ \forall\ j \in J,\ t \in T,\ m \in \mathcal{M}_j$$

$$I_{jt}^{k} = \sum_{q=1}^{t}\sum_{s=t+1}^{|T|}\sum_{m\in\mathcal{M}_j} U_{jqs}^{mk} = \sum_{q=1}^{t}\sum_{s=t+1}^{|T|}\sum_{m\in\mathcal{M}_j}\sum_{k\in K^t} U_{jqs}^{mk},\ \forall\, j\in J,\ t\in T$$

$$Y_{jt}^{mk} = Y_{jt}^{mk},\ \forall\, j\in J,\ t\in T,\ m\in\mathcal{M}_j$$

From the above equalities, we have that $\breve{X}_{jt}^{m} = \sum_{q=t}^{|T|}\sum_{k\in K^t}\pi_{tk}\cdot U_{jtq}^{mk} = \sum_{k\in K^t}\pi_{tk}\cdot X_{jt}^{mk}$, $\breve{I}_{jt} = \sum_{q=1}^{t}\sum_{s=t+1}^{|T|}\sum_{m\in\mathcal{M}_j}\sum_{k\in K^t}\pi_{tk}\cdot U_{jqs}^{mk} = \sum_{k\in K^t}\pi_{tk}\cdot I_{jt}^{k}$, and $\breve{Y}_{jt}^{m} = \sum_{k\in K^t}\pi_{tk}\cdot Y_{jt}^{mk}$. Also, the solution ($X_{jt}^{mk}$, I_{jt}^{k}, Y_{jt}^{mk}) satisfies constraints (3)-(7), which implies that the solution ($\breve{X}_{jt}^{m}$, $\breve{I}_{jt}$, $\breve{Y}_{jt}^{m}$) is a convex combination of solutions (X_{jt}^{mk}, I_{jt}^{k}, Y_{jt}^{mk}) and is in the convex hull of constraints (3)-(7). Besides, the solution ($\breve{X}_{jt}^{m}$, $\breve{I}_{jt}$, $\breve{Y}_{jt}^{m}$) satisfies constraints (2) because the original solution ($\breve{U}_{jtq}^{m}$, $\breve{Y}_{jt}^{m}$) satisfies constraints (9). Herein, ($\breve{X}_{jt}^{m}$, $\breve{I}_{jt}$, $\breve{Y}_{jt}^{m}$) is a feasible solution of the per-period decomposition of LSNM. In addition, the definition of cost coefficients in FLNM guarantees that both solutions have the same objective value.

According to the computational tests, most test instances in J3-T8-M2 and J4-T6-M3 can be examples that show the inequality is possibly strict. □

4.5 The proof procedure of Proposition 5: $\mathbf{LB}_{FLNM-PD} = \mathbf{LB}_{SPNM-PD}$

We prove the proposition by showing that a feasible solution ($\breve{W}_{jtq}^{m}$, $\breve{Y}_{jt}^{m}$) of the per-period decomposition of SPNM corresponds to a feasible solution ($\breve{U}_{jtq}^{m}$, $\breve{Y}_{jt}^{m}$) of the per-period decomposition of FLNM with the same objective value and vice versa.

If ($\breve{W}_{jtq}^{m}$, $\breve{Y}_{jt}^{m}$) is a feasible solution, it means that a) the solution is contained in the convex hull defined by constraints (18)-(21) and b) it satisfies constraints (16)-(17). Herein, ($\breve{W}_{jtq}^{m}$, $\breve{Y}_{jt}^{m}$) can be expressed as a convex combination of all extreme points (W_{jtq}^{mk}, Y_{jt}^{mk}) of the convex hull of constraints (18)-(21):

$W_{jtq}^{m} = \sum_{k\in K^t}\pi_{tk}\cdot W_{jtq}^{mk}$ and $Y_{jt}^{m} = \sum_{k\in K^t}\pi_{tk}\cdot Y_{jt}^{mk}$, $\forall\, j\in J,\ t\in T,\ q\in\{t,\ldots,|T|\},\ m\in\mathcal{M}_j$

Projecting the feasible point ($\breve{W}_{jtq}^{m}$, $\breve{Y}_{jt}^{m}$) onto the (U, Y)-space, we have the corresponding solution:

$$\breve{U}_{jtq}^{m} = \sum_{s=q}^{|T|} d_{jq}\cdot W_{jts}^{m} = \sum_{s=q}^{|T|}\sum_{k\in K^t}\pi_{tk}\cdot d_{jq}\cdot W_{jts}^{mk},\ \forall\, j\in J,\ t\in T,\ q\in\{t,\ldots,|T|\},\ m\in\mathcal{M}_j$$

$$\breve{Y}_{jt}^{m} = \sum_{k\in K^t}\pi_{tk}\cdot\breve{Y}_{jt}^{mk},\ \forall\, j\in J,\ t\in T,\ m\in\mathcal{M}_j$$

For the extreme points $(W_{jtq}^{mk},\, Y_{jt}^{mk})$ of the convex hull of constraints (18)-(21), we can find the corresponding solution in the $(U,\, Y)$-space by projection as well:

$$U_{jtq}^{mk} = \sum_{s=q}^{|T|} d_{jq} \cdot W_{jts}^{mk},\ \forall\, j \in J,\ t \in T,\ q \in \{t, ..., |T|\},\ m \in \mathcal{M}_j$$

$$Y_{jt}^{mk} = Y_{jt}^{mk},\ \forall\, j \in J,\ t \in T,\ m \in \mathcal{M}_j$$

From the above equalities, we have $\breve{U}_{jtq}^{m} = \sum_{s=q}^{|T|} \sum_{k \in K^t} \pi_{tk} \cdot d_{jq} \cdot W_{jts}^{mk} = \sum_{k \in K^t} \pi_{tk} \cdot U_{jtq}^{mk}$ and $\breve{Y}_{jt}^{m} = \sum_{k \in K^t} \pi_{tk} \cdot Y_{jt}^{mk}$. Also, the solution $(U_{jtq}^{mk},\, Y_{jt}^{mk})$ satisfies constraints (10)-(14), which implies that the solution $(\breve{U}_{jtq}^{m},\, \breve{Y}_{jt}^{m})$ is a convex combination of solutions $(U_{jtq}^{mk},\, Y_{jt}^{mk})$ and is in the convex hull of constraints (10)-(14). Besides, the solution $(\breve{U}_{jtq}^{m},\, \breve{Y}_{jt}^{m})$ satisfies constraints (9) because the original solution $(W_{jtq}^{m},\, Y_{jt}^{m})$ satisfies constraints (16)-(17). Herein, $(U_{jtq}^{m},\, Y_{jt}^{m})$ is a feasible solution for the per-period decomposition of FLNM. The definition of cost coefficients in SPNM guarantees that both solutions have the same objective value.

The above proof process is reversible. If we let $(\breve{U}_{jtq}^{m},\, \breve{Y}_{jt}^{m})$ be a feasible solution of the per-period decomposition of FLNM and project it onto the space of the per-period decomposition of SPNM, we can show that $(W_{jtq}^{m},\, Y_{jt}^{m})$ is a feasible solution of the space of the per-period decomposition of SPNM. □

References

Degraeve, Z., R. Jans. 2007. A new Dantzig–Wolfe reformulation and branch-and-price algorithm for the capacitated lot-sizing problem with setup times. *Operations Research* **55** 909–920.