Proceedings of the 13th IFAC Symposium on
Information Control Problems in Manufacturing
Moscow, Russia, June 3-5, 2009

Optimal Lotsizing Within MRP Theory

Robert W. Grubbström FVR RI[1, 3], Marija Bogataj[2, 3] and Ludvik Bogataj[2, 3]

[1] *Linköping Institute of Technology, Sweden*
(e-mail: robert@grubbstrom.com)
[2] *University of Ljubljana, Slovenia*
(e-mail: marija.bogataj@guest.arnes.si, ludvik.bogataj@ef.uni-lj.si)
[3] *Mediterranean Institute for Advanced Studies, Šempeter pri Gorici, Slovenia*

Abstract: *MRP Theory* combines the use of Input-Output Analysis and Laplace transforms, enabling the development of a theoretical background for multi-level, multi-stage production-inventory systems together with their economic evaluation, in particular applying the Net Present Value principle (NPV). In this paper we concentrate our attention on the question of optimal lotsizing decisions within the MRP Theory framework. MRP Theory has mainly dealt with *assembly structures* by which items produced downstream (on a higher level in the product structure) contain one or more sub-items on lower levels, but at each stage, the assembly activity produces only one type of output. This enables the *input matrix*, after enumerating all items suitably, to be organised as a triangular matrix, with non-zero elements only appearing below its main diagonal. The introduction of a diagonal *lead time matrix* capturing the advanced timing when required inputs are needed, enables compact expressions to be obtained, explaining the development of key variables such as available inventory and backlogs in the frequency domain. Central in this theory is the *generalised input* matrix showing when and how much the internal (dependent) demand amounts to for any production plan. Previously has been demonstrated that in a one-product case, *inner-corner conditions* for an optimum production plan in continuous time reduce the number of possible replenishment times to a finite set of given points at which either a replenishment is made, or not. The dynamic lotsizing problem is thus turned into choosing from a set of zero-one decisions with 2^{m-1} alternatives, of which one (or possibly several equivalent) solution(s) must be optimal, where m is the number of requirement (demand) events. Given these points in time and the corresponding staircase function describing cumulative demand, the optimal plan may be obtained, for instance by employing the *Triple Algorithm* of dynamic lotsizing. This applies either an Average Cost approach, or the Net Present Value principle is applied. In this paper, we extend this analysis to a general multi-item system of the assembly type. Among other issues, it is shown how the number of internal demand events depends on product structure and number of external demand events. Also the inner-corner condition is proven still to be valid in this somewhat more complex situation, when demand for lower-level items no longer is given from the outset, but instead depends on decisions concerning production of items on higher levels in the product structure. A simple dynamic program procedure is provided offering a solution to maximising the NPV for any general assembly system. To solve for the optimal production plan, the Triple Algorithm may be applied to each stage.

1. INTRODUCTION

MRP Theory combines the use of Input-Output Analysis and Laplace transforms, enabling the development of a theoretical background for multi-level, multi-stage production-inventory systems together with their economic evaluation, in particular applying the Net Present Value principle. This theory has been developed during the last three decades and has mainly been applied to assembly systems (Grubbström and Bogataj, 1998, Grubbström and Tang, 2000), but there are also cases with systems in which the production processes are divergent (Grubbström, Bogataj and Bogataj, 2007, 2008, Grubbström, 2009), enabling more than one output from a given process, or with feedback (Grubbström, 1990, 1999a). Input-Output Analysis describes the requirements of materials and components (and also of capacities and other resources). This methodology was originally suggested for analysing macro-economic dependencies between different sectors of the economy by Wassily W. Leontief, 1906-1999 (Leontief, 1928, 1951), and has thereafter been developed also on the micro scale for describing multi-level product structures. In parallel to the later development of Input-Output Analysis, Andrew Vazsonyi, 1916-2004, (Vazsonyi, 1955, 1958), put forth an equivalent methodology adapted to the type of industrial problems that were subsequently treated by Material Requirements Planning (MRP), (Orlicky, 1975, Grubbström, 1996a). On the other hand, the Laplace transform, originally

978-3-902661-43-2/09/$20.00 © 2009 IFAC

10.3182/20090603-3-RU-2001.0562

developed by Pierre-Simon Laplace, 1749-1827, (and in a preliminary form by Leonard Euler, 1707-1783) is applied for capturing time developments and lead times of the processes concerned, and is also utilised for the economic evaluation when applying the Net Present Value principle (NPV), (Grubbström, 1996b, 1998, 1999c). Combining these two methodologies has given MRP theory a framework enabling an extremely compact analysis.

In this paper we focus our attention on the question of optimal lotsizing decisions within the MRP Theory framework. In production systems with *assembly structures*, items produced downstream (on a higher level in the product structure) contain one or more sub-items from lower levels, but at each stage, the activity produces only one type of output. This enables the *input matrix*, after enumerating all items suitably, to be organised as a triangular matrix, with non-zero elements only appearing below its main diagonal. The introduction of a diagonal *lead time matrix* capturing the advanced timing that required inputs need, enables compact expressions to be obtained, explaining the development of key variables such as available inventory and backlogs in the frequency domain. Central in this theory is the *generalised input* matrix showing when and how much the internal (dependent) demand amounts to for any given external demand.

Previously has been demonstrated that in a one-product case, *inner-corner conditions* for an optimum production plan in continuous time reduce the number of possible replenishment times to a finite set of given points at which either a replenishment is made, or not made. The dynamic lotsizing problem (allowing no backlogs) is then turned into choosing from a set of zero-one decisions with 2^{m-1} alternatives, of which one (or possibly several equivalent) solution(s) must be optimal, where m is the number of requirement events. Given these points in time and the corresponding staircase function describing cumulative requirements (demand), the optimal plan may be obtained, for instance by employing the *Triple Algorithm* for dynamic lotsizing. This applies either an average cost approach (AC) or the Net Present Value principle (NPV) is applied.

The two most well-known methods for solving the dynamic lotsizing problem are the *Wagner-Whitin* dynamic programming algorithm (W-W) developed by Harvey M. Wagner and Thomson M. Whitin (Wagner and Whitin, 1958), which leads to an optimal solution, and the Silver-Meal heuristic (Silver and Meal, 1973), which leads to a near-optimal solution. Also a few other discrete-time algorithms have been presented later, such as by Federgruen and Tzur (1991). It has been shown previously that W-W as well as the Silver-Meal heuristic may be stated both in discrete and continuous time, either the *Net Present Value* (NPV) or the *Average Cost* (AC) is applied as the objective function (Grubbström, 2005). Also our Triple Algorithm, which is of the forward type, see (Grubbström, 1999b, 2005), performs in continuous time as well as in discrete time. However, without any loss of generality, our analysis to follow will be confined to discrete events taking place in continuous time alone.

In our treatment below, we extend this analysis to a general multi-item system of the assembly type. After having given a brief overview of MRP Theory in Section 2, it is shown how the number of internal demand events depends on product structure and number of external demand events, among other issues (Section 3). Also the inner-corner condition is proven still to be valid in this somewhat more complex situation, when demand for lower-level items no longer is given from the outset, but instead depends on decisions concerning production of items on higher levels in the product structure (Section 4). In Section 5 an outline of the Triple Algorithm is provided. A simple dynamic programming procedure is developed in Section 6 for providing a solution to the problem of maximising the NPV for any general assembly system. To solve for the optimal production plan, the Triple Algorithm may be applied to each stage. A numerical example with five levels is analysed and solved in Section 7, followed by a final section summarising our conclusions.

2. OUTLINE OF BASIC MRP THEORY

The two basic ingredients in MRP Theory are Input-Output Analysis and the Laplace transform, see (Aseltine, 1958, Churchill, 1958, 1960, Zhixin and Grubbström, 2001). Starting with the latter, let us consider a given time function $x(t)$, where t is time and $t \geq 0$. The Laplace transform is then defined as the integral

$$£\{x(t)\} = \tilde{x}(s) = \int_0^\infty x(t)e^{-st}dt \,, \qquad (1)$$

where s is the so called Laplace frequency (complex frequency). In general, s is a complex variable $s = \sigma + \omega\sqrt{-1}$. Eq. (1) maps a time function into a new frequency function. The notation $£\{x(t)\}$ and $\tilde{x}(s)$ are two alternative notations for the transform, the latter explicitly showing it to be a function of s. In order for the transform to exist, the integral must converge, which implies that the function $x(t)$ may not grow faster than at an exponential rate. In all practical circumstances, the time functions allow the transform to exist, and there is a one-to-one relationship between the time function (in the *time domain*) and the transform (in the *frequency domain*). The time function corresponding to a given transform is called the *inverse transform*, often written as $x(t) = £^{-1}\{\tilde{x}(s)\}$. In this article, we make use of only a minute part of Laplace transform methodology.

A *step function* of level a starting at $t = 0$ has the transform:

$$£\{a\} = \int_0^\infty ae^{-st}dt = a/s \,, \qquad (2)$$

A special (generalised) function is the *Dirac impulse* (the *impulse function*) written $\delta(t-T)$. This time function only exists at $t = T$, where T is a given point in time, $T \geq 0$. It can be defined as an infinitely narrow and infinitely tall impulse with a unit area, such as for instance a rectangle with $\delta(t-T) = a$, for $t \geq T$ and $t \leq T + 1/a$, and $\delta(t-T) = 0$ elsewhere. The transform of $\delta(t-T) = 0$ is given by:

$$£\{\delta(t-T)\} = e^{-sT} . \tag{3}$$

Of the vast set of theorems existing for the Laplace transform, we mention the following:

Time differentiation and time integration

$$£\left\{\frac{dx(t)}{dt}\right\} = s\tilde{x}(s) ,\ £\left\{\int_0^t x(\tau)d\tau\right\} = \tilde{x}(s)/s . \tag{4}$$

Final value (and time average) and initial value, assuming these limits exist

$$\lim_{t\to\infty} x(t) = \lim_{t\to\infty} \frac{1}{t}\int_0^t x(\tau)d\tau = \lim_{s\to 0} s\tilde{x}(s) , \tag{5}$$

$$\lim_{t\to 0} x(t) = \lim_{s\to\infty} s\tilde{x}(s) . \tag{6}$$

Time translation

$$£\{x(t-T)\} = e^{-sT}\tilde{x}(s) . \tag{7}$$

When a function $x(t)$ is moved uniformly forwards in time, we have $T > 0$, and this formula holds defining $x(t - T)$ to be zero for $t < T$. But when we move backwards in time and $T < 0$, then the function might cross $t = 0$ and the function becomes truncated. So the formula assumes always that $x(t - T) = 0$ for $t < T$. This theorem is used for describing the effects of lead times.

A *staircase function* is made up of steps of various heights taking place at a set of given times, say the m points in time $t_1, t_2, ..., t_m$. Let the steps at these points have the heights $Q_1, Q_2, ..., Q_m$. Looking upon the staircase as a sum of steps, using (2) and (7) the jth step contributes with $Q_j e^{-st_j}/s$ giving the staircase the transform:

Staircase function

$$\sum_{j=1}^{m} Q_j e^{-st_j}/s . \tag{8}$$

But we can also look upon the staircase function as an accumulation (time integral) of contributions from Dirac impulses $Q_j e^{-st_j}$, in which case (3) and (4) provide the same expression $\left(\sum_{j=1}^{m} Q_j e^{-st_j}\right)/s$.

An additional application of the Laplace transform is its use in financial evaluations (Grubbström, 1967, Buser, 1986, Grubbström and Jiang, 1990). The *Net Present Value* (NPV) of a cash flow $x(t)$, existing in non-negative time, quite generally, may be written

*Net Present Value (*NPV*)*

$$\text{NPV} = \int_0^\infty x(t)e^{-\rho t}dt , \tag{9}$$

where ρ is the *continuous interest* rate. Comparing this integral with (1), we immediately see the identity:

$$\text{NPV} = \left[£\{x(t)\}\right]_{s=\rho} = \tilde{x}(\rho) . \tag{10}$$

So, if given the transform of the cash flow $\tilde{x}(s)$, its NPV is directly obtained by substituting the complex frequency by the continuous interest rate.

When there are discrete payments, these are interpreted as Dirac impulses.

The *annuity stream* is the constant cash flow equivalence of a series of payments during a given interval T. For any NPV evaluated from a cash flow during an interval of length T, its corresponding annuity stream will be

Annuity Stream

$$\text{Annuity Stream} = \frac{\rho\text{NPV}}{1-e^{-\rho T}} . \tag{11}$$

Whereas several often-used mathematical operations in the time domain are quite cumbersome to perform, such as multiple integrations, their corresponding operations in the frequency domain become simple algebraic manipulations. The one-to-one property for a wide range of time functions and their transforms ensures that the resulting transform is a unique counter-part of the corresponding time function.

In our assembly system, we have altogether n items. The external demand $\mathbf{D}(t)$ and production $\mathbf{P}(t)$ are represented by n-dimensional column vectors each being a function of time. These vectors are defined as rates and converted into Laplace transforms $\tilde{\mathbf{D}}(s)$ and $\tilde{\mathbf{P}}(s)$. The vector $\tilde{\mathbf{P}}(s)$ also includes purchasing (which can be considered a special case of production) and it is defined for production at the time of its completion. We will be considering only batch production, which implies that $\tilde{\mathbf{P}}(s)$ is made up of a set of impulses, and similarly we assume external demand to occur in given amounts at given discrete points in time, also making components of $\tilde{\mathbf{D}}(s)$ sums of impulses.

For the production (assembly) of one unit of item j, there is a need in the amount of h_{kj} of other items k, $k = 1, 2, \dots, n$, and there is a *lead time* τ_j ahead of the completion of the production (assembly) of item j at which the components are needed. The h_{kj} are arranged into the square *input matrix* $\mathbf{H}$ describing the product structures of all relevant products. The lead times $\tau_1, \tau_2, \dots, \tau_n$, create internal demands in advance of production and are arranged into a diagonal matrix $\tilde{\boldsymbol{\tau}}(s)$, the *lead time matrix,* having $e^{s\tau_j}$ in its jth diagonal position:

$$\tilde{\boldsymbol{\tau}}(s) = \begin{bmatrix} e^{s\tau_1} & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & e^{s\tau_n} \end{bmatrix} . \tag{12}$$

The product $\mathbf{H}\tilde{\boldsymbol{\tau}}(s)$ is the *generalised input matrix* and it captures component requirements together with their required advance timing.

If a plan for production $\tilde{\mathbf{P}}(s)$ is given, this production will create an *internal* (*dependent*) *demand* amounting to $\mathbf{H}\tilde{\boldsymbol{\tau}}(s)\tilde{\mathbf{P}}(s)$, where the diagonal lead time matrix $\tilde{\boldsymbol{\tau}}(s)$ multiplies each column j of $\mathbf{H}$ with the factors $e^{s\tau_j}$ respectively, moving the requirements the respective lead time backwards in time, i. e. an application of the translation theorem (7). Internal demand represents ear-marked items designated for the given production and is an inflow of items into *work-in-progress* (*allocated component stock*, *allocations*). At the time of completion, items change identities, newly created products appear and corresponding ear-marked requirements disappear. The outflow from work-in-progress is thus $\mathbf{H}\tilde{\mathbf{P}}(s)$.

Available inventory $\tilde{\mathbf{R}}(s)$ is defined as total inventory less allocations and represents items that may satisfy *external* (*independent*) *demand* $\tilde{\mathbf{D}}(s)$. With an inflow of $\tilde{\mathbf{P}}(s)$ and an internal outflow of $\mathbf{H}\tilde{\boldsymbol{\tau}}(s)\tilde{\mathbf{P}}(s)$ and external outflow of $\tilde{\mathbf{D}}(s)$, disregarding backlog issues, we have

$$\tilde{\mathbf{R}}(s) = \frac{\tilde{\mathbf{R}}(0) + \left(\mathbf{I} - \mathbf{H}\tilde{\boldsymbol{\tau}}(s)\right)\tilde{\mathbf{P}}(s) - \tilde{\mathbf{D}}(s)}{s}, \tag{13}$$

where $\mathbf{R}(0)$ is initial available inventory, $\mathbf{I}$ the identity matrix, and where the time integration theorem (4) has been applied. Other similar fundamental equations of MRP Theory concern *allocations* and *backlogs*.

The non-negativity of available inventory over time, i. e. the vector $\mathbf{R}(t)$ satisfying $\mathbf{R}(t) \geq 0$

$$\mathbf{R}(t) = £^{-1}\left\{\tilde{\mathbf{R}}(s)\right\} = $$
$$= £^{-1}\left\{\frac{\tilde{\mathbf{R}}(0) + \left(\mathbf{I} - \mathbf{H}\tilde{\boldsymbol{\tau}}(s)\right)\tilde{\mathbf{P}}(s) - \tilde{\mathbf{D}}(s)}{s}\right\} \geq \mathbf{0}, \tag{14}$$

is one of the fundamental constraints of MRP systems. If $\mathbf{R}(t)$ would have some negative component at some time, this would mean a shortage and the plan represented by $\tilde{\mathbf{P}}(s)$ would not be feasible. Other similar constraints concern capacities (Grubbström and Wang, 2003, Huynh, 2006), but are not treated here, since we are avoiding to take capacity constraints into consideration in this current treatment.

In MRP terminology, $\tilde{\mathbf{D}}(s)$ may be called the *Master Production Schedule* (MPS). This is what is planned to be produced for external delivery. The *Lot-for-Lot* replenishment policy (L4L, or "As Required") determines $\tilde{\mathbf{P}}(s)$ exactly to satisfy the given MPS, i. e. to avoid any addition to available inventory. Therefore $\left(\mathbf{I} - \mathbf{H}\,\tilde{\boldsymbol{\tau}}(s)\right)\tilde{\mathbf{P}}(s)$ $-\tilde{\mathbf{D}}(s) = \mathbf{0}$, and

$$\tilde{\mathbf{P}}(s) = \left(\mathbf{I} - \mathbf{H}\,\tilde{\boldsymbol{\tau}}(s)\right)^{-1}\tilde{\mathbf{D}}(s), \tag{15}$$

explaining explicitly exactly how much to produce of all items at all times, in order to satisfy $\tilde{\mathbf{D}}(s)$. This simple compact formula covers every possible production plan for any MRP system with any MPS, when L4L is applied. For other standard ordering policies, such as *Fixed Order Quantity* (FOQ) or *Fixed Period Requirements* (FPR), the results are not equally explicit (Grubbström and Huynh, 2006).

The converse policy to L4L is "All at Once" ($\forall@1$). This policy minimises the number of setups and maximises available inventory) Here, setups only occur at the first necessary point in time of a production plan, and all items of a particular type are manufactured in one batch only. Although we do not have a simple algebraic formula available for the $\forall@1$ policy, the production plan is easily obtained from (15) by adding all amounts together for each item, and then timing these totals at the earliest of the times given by L4L (see example in Section 7).

The inverse matrix $\left(\mathbf{I} - \mathbf{H}\tilde{\boldsymbol{\tau}}(s)\right)^{-1}$ is a generalisation of the *Leontief inverse* $\left(\mathbf{I} - \mathbf{H}\right)^{-1}$ used in Input-Output Analysis, and $\lim_{s\to 0}\left(\mathbf{I} - \mathbf{H}\tilde{\boldsymbol{\tau}}(s)\right)^{-1} = \left(\mathbf{I} - \mathbf{H}\right)^{-1}$. If the characteristic roots (Eigen-values) of $\mathbf{H}$ all have absolute values less than unity, then $\left(\mathbf{I} - \mathbf{H}\right)^{-1}$ may be expanded into a *Neumann series*:

$$\left(\mathbf{I} - \mathbf{H}\right)^{-1} = \mathbf{I} + \mathbf{H} + \mathbf{H}^2 + \ldots\ . \tag{16}$$

For typical assembly systems, $\mathbf{H}$ is *triangular* with zeros on and above its main diagonal. In such a case, the series (16) converges in a finite number of steps, since $\mathbf{H}^k = \mathbf{0}$, for $k \geq n$. If $\mathbf{H}$ is triangular, so is $\left(\mathbf{I} - \mathbf{H}\right)^{-1}$ (but with unit elements along its main diagonal). The series expansion also shows that the Leontief inverse has only non-negative elements. Similar properties hold for the *generalised Leontief inverse* $\left(\mathbf{I} - \mathbf{H}\tilde{\boldsymbol{\tau}}(s)\right)^{-1}$.

In this paper, we will confine ourselves to deterministic situations with a given plan for deliveries from the system without considering opportunities to reschedule or to hold safety stocks. Also, we are refraining from considering capacity constraints. Issues related to these more general situations have been treated in other work (Bogataj and Ferbar, 1996, Segerstedt, 1996, Tang, 2000, Grubbström and Wang, 2003, Huynh, 2006). Other extensions and work related to MRP Theory can be found, for instance among (Martin, 1997, Bogataj, 1999, Ferbar and Bogataj, 1999, Bogataj and Bogataj, 2001, 2003, 2007, Grubbström, 2007).

For purposes of stating an objective function, MRP theory introduces economic parameters in the form of a price vector

p to describe unit revenues from sales, a unit production cost vector **c** containing a unit variable cost of production (avoiding any fixed costs), and a vector of setup costs **K** capturing the fixed cost associated with the production of a batch. Each of these vectors is an *n*-dimensional row vector, its *i*th component designating the parameter in question concerning item *i*. Furthermore the continuous interest rate ρ is used for discounting the relevant cash flows. The cash flow of revenue in-payments is thus $\mathbf{pD}(t)$, the cash flow of production cost out-payments is $\mathbf{cP}(t)$ and of fixed setup costs $\mathbf{KP}'(t)$, where the *setup vector* $\mathbf{P}'(t)$ describes the timing of production batches being completed as a sum of unit impulses.

In $\mathbf{KP}'(t)$ the out-payments for setups are thus associated with production completion times rather than the times at which the production of a batch commences. This is made only for the sake of notational simplicity. Without loss of generality, the possibility to discount a component to the lead time ahead of completion (or to any other intermediate time) may be included in the values of the components of **K**.

Applying the net present value theorem (9), we may write the economic consequences from a given production plan as:

$$\text{NPV} = \mathbf{p}\tilde{\mathbf{D}}(\rho) - \mathbf{c}\tilde{\mathbf{P}}(\rho) - \mathbf{K}\tilde{\mathbf{P}}'(\rho). \tag{17}$$

Since we are not allowing for any shortages to exist in the treatment below, we may disconsider any economic consequences concerning backlogs and lost sales. We may also note that only planned production is treated with external demand assumed to be known in advance, so all variables are deterministic.

3. DEVELOPMENT OF EXPRESSIONS FOR DEMAND EVENTS

In this section we develop expressions for the number of demand events (external and internal) that the system faces, when given sequences of external demand events during a finite horizon. Throughout our treatment in this and the following sections, the input matrix **H** is assumed to be written in triangular form, meaning that items on a higher level in the product structure have lower indices. Internal and external or different sources of internal demand may, by coincidence, happen to generate internal demand events at exactly the same time.

External demand $\tilde{\mathbf{D}}(s)$ is assumed to take place in discrete amounts at different times. For item no *i*, its external demand can be written:

$$\tilde{\mathbf{D}}_i(s) = \sum_{j=1}^{M_i} D_{ij} e^{-st_{ij}} = \begin{bmatrix} D_{i1} & \cdots & D_{iM_i} \end{bmatrix} \begin{bmatrix} e^{-st_{i1}} \\ \vdots \\ e^{-st_{iM_i}} \end{bmatrix}, \tag{18}$$

where D_{ij} is the size of demand at the *j*th demand event that takes place at time t_{ij}, and M_i is the number of *external* demand events of item *i*.

By replacing the amounts D_{ij} by unit elements, the sequence of events as a train of unit impulses written $\tilde{\mathbf{D}}_i'(s)$ will be

$$\tilde{\mathbf{D}}_i'(s) = \sum_{j=1}^{M_i} e^{-st_{ij}} = \begin{bmatrix} 1 & \cdots & 1 \end{bmatrix} \begin{bmatrix} e^{-st_{i1}} \\ \vdots \\ e^{-st_{iM_i}} \end{bmatrix}, \tag{19}$$

and therefore:

$$M_i = \lim_{s\to 0} \tilde{\mathbf{D}}_i'(s). \tag{20}$$

We may collect all $\tilde{\mathbf{D}}_i'(s)$ into a common vector $\tilde{\mathbf{D}}'(s)$,

$$\tilde{\mathbf{D}}'(s) = \begin{bmatrix} \tilde{\mathbf{D}}_1'(s) \\ \vdots \\ \tilde{\mathbf{D}}_n'(s) \end{bmatrix}, \tag{21}$$

and define a vector **M** according to

$$\mathbf{M} = \begin{bmatrix} M_1 \\ \vdots \\ M_n \end{bmatrix} = \lim_{s\to 0} \tilde{\mathbf{D}}'(s). \tag{22}$$

This vector contains the number of external demand events of all items (some possibly by coincidence double-counted).

The immediate internal demand of components and raw materials generated from external demand $\tilde{\mathbf{D}}(s)$ is $\mathbf{H}\tilde{\boldsymbol{\tau}}(\mathbf{s})\tilde{\mathbf{D}}(s)$, which in its turn generates the internal demand $(\mathbf{H}\tilde{\tau}(s))^2\,\tilde{\mathbf{D}}(s)$, etc., so when production exactly matches demand, total internal demand will be

$$\sum_{k=1}^{n-1} \left(\mathbf{H}\tilde{\boldsymbol{\tau}}(s)\right)^k \tilde{\mathbf{D}}(s) = \left(\left(\mathbf{I} - \mathbf{H}\tilde{\boldsymbol{\tau}}(s)\right)^{-1} - \mathbf{I}\right)\tilde{\mathbf{D}}(s), \tag{23}$$

since $\left(\mathbf{H}\tilde{\boldsymbol{\tau}}(s)\right)^k = \mathbf{0}$, for $k \geq n$, $\mathbf{H}\tilde{\tau}(s)$ being a triangular matrix with zeros in its main diagonal, and where the sum is interpreted as a Neumann expansion of the inverse matrix.

By replacing positive elements in the input matrix by unit elements, writing this new *modified input matrix* $\mathbf{H}'$, the internal demand events as a whole will be given by:

$$\left(\left(\mathbf{I} - \mathbf{H}'\tilde{\boldsymbol{\tau}}(s)\right)^{-1} - \mathbf{I}\right)\tilde{\mathbf{D}}'(s). \tag{24}$$

Here, the lead time matrix $\tilde{\boldsymbol{\tau}}(s)$ does not affect the number of events, only their advanced timing. If we add external demand events $\tilde{\mathbf{D}}'(s)$, the total sequences of external *and* internal demand events will be

$$\tilde{\mathbf{m}}(s)=\begin{bmatrix}\tilde{m}_1(s)\\ \vdots\\ \tilde{m}_n(s)\end{bmatrix}=$$

$$=\left(\left(\mathbf{I}-\mathbf{H}'\tilde{\boldsymbol{\tau}}(s)\right)^{-1}-\mathbf{I}\right)\tilde{\mathbf{D}}'(s)+\tilde{\mathbf{D}}'(s)= \tag{25}$$

$$=\left(\mathbf{I}-\mathbf{H}'\tilde{\boldsymbol{\tau}}(s)\right)^{-1}\tilde{\mathbf{D}}'(s),$$

defining the vector $\tilde{\mathbf{m}}(s)$, and where $\tilde{m}_i(s)$ is the sequence of all events for item *i*.

We may compare this equation with the production plan following the Lot-for-Lot policy (L4L) according to (15), in which case available inventory is kept at zero levels throughout. Hence the maximum number of demand events equals the number of production events when following the L4L policy. The total number of events can thus be conveniently written:

$$\mathbf{m}=\begin{bmatrix}m_1\\ \vdots\\ m_n\end{bmatrix}=\lim_{s\to 0}\begin{bmatrix}\tilde{m}_1(s)\\ \vdots\\ \tilde{m}_n(s)\end{bmatrix}=$$

$$=\lim_{s\to 0}\left(\mathbf{I}-\mathbf{H}'\tilde{\boldsymbol{\tau}}(s)\right)^{-1}\tilde{\mathbf{D}}'(s)= \tag{26}$$

$$=\left(\mathbf{I}-\mathbf{H}'\right)^{-1}\begin{bmatrix}M_1\\ \vdots\\ M_n\end{bmatrix}=\left(\mathbf{I}-\mathbf{H}'\right)^{-1}\mathbf{M},$$

since $\lim_{s\to 0}\tilde{\boldsymbol{\tau}}(s)=\mathbf{I}$, and where $\left(\mathbf{I}-\mathbf{H}'\right)^{-1}$ is the *modified Leontief inverse* obtained when replacing all positive elements of the input matrix by unit elements.

The obvious case, when the number of demand events is the highest possible, is obtained for a matrix $\mathbf{H}'=\hat{\mathbf{H}}'$ that contains unit elements in all positions below its diagonal. Then all items in the product structure will require all types of items on lower levels (i. e. manufactured upstream). We derive the following theorem for this case:

Theorem 1: Let $\hat{\mathbf{H}}'$ be a modified input matrix with unit-valued elements in all positions below its main diagonal and zeros elsewhere. Given a lead time matrix $\tilde{\boldsymbol{\tau}}(s)$, its generalised Leontief inverse $\left(\mathbf{I}-\hat{\mathbf{H}}'\tilde{\boldsymbol{\tau}}(s)\right)^{-1}$ will be:

$$\left[\left(\mathbf{I}-\hat{\mathbf{H}}'\tilde{\boldsymbol{\tau}}(s)\right)^{-1}\right]_{ij}=\begin{cases}0, \text{ for } i<j,\\ 1, \text{ for } i=j,\\ e^{\tau_j s}, \text{ for } i=j+1,\\ e^{\tau_j s}\prod_{k=j+1}^{i-1}\left(1+e^{\tau_k s}\right), \text{ for } i>j+1.\end{cases} \tag{27}$$

Proof: Since $\mathbf{I}-\hat{\mathbf{H}}'\tilde{\boldsymbol{\tau}}(s)$ is triangular with unit elements along its main diagonal, it has a determinant equal to unity and is regular. Hence, its inverse is unique. Therefore, it suffices to show that the product $\left(\mathbf{I}-\hat{\mathbf{H}}'\tilde{\boldsymbol{\tau}}(s)\right)\left(\mathbf{I}-\hat{\mathbf{H}}'\tilde{\boldsymbol{\tau}}(s)\right)^{-1}$ is the identity matrix **I**. Taking the scalar product of row *i* and column *j* in the two matrices, it is trivially seen that

$$\sum_{l=1}^{n}\left[\left(\mathbf{I}-\hat{\mathbf{H}}'\tilde{\boldsymbol{\tau}}(s)\right)\right]_{il}\left[\left(\mathbf{I}-\hat{\mathbf{H}}'\tilde{\boldsymbol{\tau}}(s)\right)^{-1}\right]_{lj}=\begin{cases}0, \text{ for } i<j,\\ 1, \text{ for } i=j.\end{cases}$$

For $i>j$, we develop the non-zero terms of the product according to

$$\sum_{l=1}^{n}\left[\left(\mathbf{I}-\hat{\mathbf{H}}'\tilde{\boldsymbol{\tau}}(s)\right)\right]_{il}\left[\left(\mathbf{I}-\hat{\mathbf{H}}'\tilde{\boldsymbol{\tau}}(s)\right)^{-1}\right]_{lj}=$$

$$e^{s\tau_j}\left(-1-\sum_{l=j+1}^{i-1}e^{s\tau_l}\prod_{k=j+1}^{l-1}\left(1+e^{\tau_k s}\right)+\prod_{k=j+1}^{i-1}\left(1+e^{\tau_k s}\right)\right)=$$

$$e^{s\tau_j}\left(-1-\sum_{l=j+2}^{i}\prod_{k=j+1}^{l-1}\left(1+e^{\tau_k s}\right)+\sum_{l=j+1}^{i-1}\prod_{k=j+1}^{l-1}\left(1+e^{\tau_k s}\right)+\prod_{l=k+1}^{i-1}\left(1+e^{\tau_k s}\right)\right)=0,$$

where the summation has been moved one step in the second term, and the convention $\prod_{k=j+1}^{j}\left(1+e^{\tau_k s}\right)=1$ applied. ■

With this structure of the input matrix, it will generate the largest number of internal demand events, and their advanced timing is given by (27). This leads us immediately to:

Theorem 2: The maximum number of internal demand events occurs for a system having the following modified Leontief inverse

$$\left[\left(\mathbf{I}-\hat{\mathbf{H}}'\right)^{-1}\right]_{ij}=\begin{cases}0, \text{ for } i>j,\\ 1, \text{ for } i=j,\\ 2^{i-j-1}, \text{ for } i<j.\end{cases} \tag{28}$$

Proof: In (27), we have $\lim_{s\to 0}\tilde{\boldsymbol{\tau}}(s)=\mathbf{I}$, and therefore $\lim_{s\to 0}\left(\mathbf{I}-\hat{\mathbf{H}}'\tilde{\boldsymbol{\tau}}(s)\right)^{-1}=\left(\mathbf{I}-\hat{\mathbf{H}}'\right)^{-1}$. Taking the similar limit of the right-hand member of (27), we obtain zero, for $i<j$, unity for $i=j$ and $i=j+1$, and finally $\lim_{s\to 0}e^{\tau_j s}\prod_{k=j+1}^{i-1}\left(1+e^{\tau_k s}\right)=2^{i-j-1}$, for $i>j$. ■

As an example we choose a 5×5 system, in which case:

$$\mathbf{I}_{5\times 5}-\hat{\mathbf{H}}'_{5\times 5}=\begin{bmatrix}1&0&0&0&0\\-1&1&0&0&0\\-1&-1&1&0&0\\-1&-1&-1&1&0\\-1&-1&-1&-1&1\end{bmatrix}$$

and

$$\left(\mathbf{I}_{5\times5} - \hat{\mathbf{H}}'_{5\times5}\right)^{-1} = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 \\ 2 & 1 & 1 & 0 & 0 \\ 4 & 2 & 1 & 1 & 0 \\ 8 & 4 & 2 & 1 & 1 \end{bmatrix}.$$

For the maximum-event case, for the total number of demand events we have $\mathbf{m} = \left(\mathbf{I} - \hat{\mathbf{H}}'\right)^{-1}\mathbf{M}$, so:

$$m_i = \sum_{j=1}^{i-1} 2^{i-j-1} M_j + M_i\,. \quad (29)$$

4. THE INNER-CORNER CONDITION FOR OPTIMAL LOTSIZING

We now turn our attention to the necessary *inner-corner condition* for a production plan to be optimal. This condition provides us with the opportunity to formulate the lotsizing problem using *binary* decision variables.

For the one-product system it has been shown previously that the optimal production plan must satisfy the inner-corner condition (Grubbström and Molinder, 1994, Grubbström, 1999b, Grubbström, 2005). Dropping the first subscript of relevant variables temporarily (the item index), this condition can be illustrated as a relationship between a right staircase function describing cumulative requirements (demand) $\bar{D}$ and a left staircase describing cumulative production (replenishments) $\bar{P}$ as shown in Fig. 1. Either the AC or the NPV is applied, at the optimum the two staircases must fit into each other at "inner corners" (indicated by circles in the figure). Thus the optimal number of batches produced can never be higher than the number of events generating requirements *m*. Also the final level of cumulative production and cumulative demand always must coincide, $\bar{P}_m = \bar{D}_m$. There is no use of any surplus after the end time t_m.

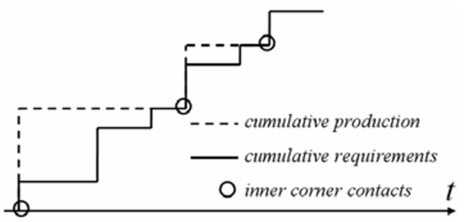

Fig. 1. The staircases of cumulative production (dashed) and cumulative demand (solid). Inner corner contacts are marked by circles.

There must always be a setup at the initial time t_1, so in the one-product case, the inner-corner condition limits the different alternative candidates for the optimum production staircase to the number 2^{m-1}, which increases fast in *m*. For *m* = 30, this number is about 537 million, for *m* = 40 about 550 billion, and for *m* = 100 about $0.63\cdot10^{30}$.

We therefore state the following theorem:

Theorem 3: For an assembly system with no backlogs allowed, it is always optimal to complete production at points belonging to the set of inner corners, either the Net Present Value Principle or the Average Cost measure is applied.

Proof: The net present value given any feasible production plan $\tilde{\mathbf{P}}(s)$ (assuming no backlogging) is given by (17) $\text{NPV} = \mathbf{p}\tilde{\mathbf{D}}(\rho) - \mathbf{c}\tilde{\mathbf{P}}(\rho) - \mathbf{K}\tilde{\mathbf{P}}'(\rho)$, where $\tilde{\mathbf{P}}'(s)$ is the setup vector, i. e. the production vector $\tilde{\mathbf{P}}(s)$ with batch volumes replaced by unit values. Feasibility of $\tilde{\mathbf{P}}(s)$ is given by the available inventory constraint (14).

First, we pay attention to an end item (facing external demand only). If a plan $\tilde{\mathbf{P}}(s)$ contains an end item production batch $\tilde{\mathbf{P}}_{ij}(s) = Q_{ij}e^{-st_{ij}}$ *not* being at an inner corner of external demand for this item, then either this production may be delayed a finite interval Δt, i. e. $Q_{ij}e^{-s(t_{ij}+\Delta t)}$, or the level Q_{ij} may be reduced by a finite amount ΔQ_{ij}, or both types of change may be made, without making the new plan infeasible.

In the first case, the change in available inventory is $\pounds^{-1}\left\{\Delta\tilde{\mathbf{R}}(s)\right\} = \pounds^{-1}\left\{\left(\mathbf{I} - \mathbf{H}\tilde{\boldsymbol{\tau}}(s)\right)_{\text{column}\,i} Q_{ij}e^{-st_{ij}}\left(e^{-s\Delta t} - 1\right)/s\right\}$. The column vector $\left(\mathbf{I} - \mathbf{H}\tilde{\boldsymbol{\tau}}(s)\right)_{\text{column}\,i}$ has a unit element in position *i*, zeros above the main diagonal, and elements of the type $-H_{ki}e^{s\tau_i}$, $k > i$, below the main diagonal. The change in completion time of end product *i* creates a reduction in its own available inventory by $\pounds^{-1}\left\{\Delta\tilde{R}_i(s)\right\} = \pounds^{-1}\left\{Q_{ij}e^{-st_{ij}}\left(e^{-s\Delta t} - 1\right)/s\right\}$ which is feasible as long as the new completion time $t_{ij} + \Delta t$ does not exceed the time of the adjacent inner corner of cumulative external demand $\tilde{D}_i(s)/s$. For items *k* below item *i*, the change in available inventory becomes $\pounds^{-1}\left\{\Delta\tilde{R}_k(s)\right\} = \pounds^{-1}\left\{H_{ki}e^{s\tau_i}Q_{ij}e^{-st_{ij}}\left(1 - e^{-s\Delta t}\right)/s\right\}$, which is an increase in available inventory and therefore makes the new plan feasible.

Instead, for a reduction in batch size by ΔQ_{ij} (negative), as long as the new level does not violate cumulative external demand $\tilde{D}_i(s)/s$, the reduction in available inventory of item *i* amounts to $\pounds^{-1}\left\{\Delta\tilde{R}_i(s)\right\} = \pounds^{-1}\left\{\Delta Q_{ij}e^{-st_{ij}}/s\right\}$ and for lower levels *k* an increase obtains by $\pounds^{-1}\left\{\Delta\tilde{R}_k(s)\right\} = \pounds^{-1}\left\{-H_{ki}e^{s\tau_i}\Delta Q_{ij}e^{-st_{ij}}/s\right\}$ since ΔQ_{ij} is negative. Hence, the available inventory constraint is not violated by small finite changes in either completion time or production level (when these changes comply with the end product cumulative demand).

For a postponement by Δt, the net present value changes by $\Delta\text{NPV} = -\mathbf{c}\Delta\tilde{\mathbf{P}}(\rho) - \mathbf{K}\Delta\tilde{\mathbf{P}}'(\rho)$, where $\Delta\tilde{P}_i(\rho) =$

$Q_{ij}e^{-\rho t_{ij}}\left(e^{-\rho\Delta t}-1\right)<0$ and $\Delta P_i'(\rho)=e^{-\rho t_{ij}}(e^{-\rho\Delta t}-1)<0$, and $\Delta\tilde{P}_k(\rho)=0$ and $\Delta\tilde{P}_k'(\rho)=0$, for $k\neq i$. Therefore, $\Delta\text{NPV}=-\mathbf{c}\Delta\tilde{\mathbf{P}}(\rho)-\mathbf{K}\Delta\tilde{\mathbf{v}}(\rho)=\left(c_iQ_{ij}+K_i\right)e^{-\rho t_{ij}}\left(1-e^{-\rho\Delta t}\right)>0.$ Instead, for a reduction in production level, we have simply $\Delta\text{NPV}=-\mathbf{c}\Delta\tilde{\mathbf{P}}(\rho)=-c_i\Delta Q_{ij}e^{-\rho t_{ij}}$, which also is positive. Hence, a feasible production plan for an end item not having production at inner corners can never be optimal.

For an item on a lower production level (upstream), similar conditions apply. Postponing production, if this is possible cannot reduce available inventory of downstream items, but reduces its own available inventory. Neither can such a postponement create any shortage concerning upstream items as long as cumulative production does not exceed cumulative demand. Instead, if the production level of this item is possible to reduce, without violating the requirements (created by downstream items alone in an assembly system), such a reduction cannot create any shortage for either upstream or downstream items. So with positive vectors **c** and **K**, and with a positive interest rate ρ, if a move to an adjacent inner corner is possible, then this increases the value of NPV.

Applying the average cost measure, the consequences are similar. The number of setups will not be changed, and therefore not total setup costs, and available inventory will be reduced. ■

Therefore, an optimal production plan involves at most production events at the points generated by $\mathbf{m}=\left(\mathbf{I}-\mathbf{H}'\right)^{-1}\mathbf{M}$ in (26). Reductions in utilising all available points, i. e. refraining from the L4L policy on a higher level, reduces the number of opportunities upstream. Also, the number of requirement events may be lower due to more than one event being generated at the same time. This may happen for combinations of external demand events and lead times by coincidence.

We therefore conclude that the optimal production plan for the system as a whole is to be found from a zero/one problem concerning at most all points generated by $\left(\mathbf{I}-\mathbf{H}'\tilde{\boldsymbol{\tau}}(s)\right)^{-1}\tilde{\mathbf{D}}'(s)$.

The relation between the zero/one decisions α_i, a production plan $\tilde{\mathbf{P}}(s)$, and requirements, is now studied. Assume the production plan $\tilde{\mathbf{P}}(s)$ is given as concerns items $k=1,2,\ldots,i-1$. Then there is a given total demand (requirements) for item i, written as $\tilde{\hat{D}}_i(s)$, amounting to

$$\tilde{\hat{D}}_i(s)=\left[\mathbf{H}\tilde{\boldsymbol{\tau}}(s)\right]_{\text{row } i}\tilde{\mathbf{P}}(s)+\tilde{D}_i(s)\,, \tag{30}$$

which is made up of internal demand (first term) and external demand (second term). Since **H** is triangular for our assembly system, only production plans $\tilde{P}_k(s)$ for items $k=1,2,\ldots,i-1$, will contribute to the first term of internal demand. Let us write cumulative requirements as $\tilde{\bar{\hat{D}}}_i(s)=\tilde{\hat{D}}_i(s)/s$, where the division by s represents time integration. The timing of the requirements are given by $\left[\mathbf{H}'\tilde{\boldsymbol{\tau}}(s)\right]_{\text{row } i}\tilde{\mathbf{P}}'(s)+\tilde{D}_i'(s)$.

From our binary decision variable approach, we choose zero/one variables α_{ij}, for $j=1,2,3,\ldots,m_i$, as decision variables with $\alpha_{ij}=1$ meaning that production occurs at t_{ij}, and $\alpha_{ij}=0$ not. It always applies that $\alpha_{i1}=1$ (the earliest inner corner). Hence, when this plan is implemented, we have $\tilde{P}_i'(s)=\sum_{j=1}^{m_i}\alpha_{ij}e^{-st_{ij}}$ as regards timing. The production plan for item i will be $\tilde{P}_i(s)=\sum_{j=1}^{m_i}Q_{ij}e^{-st_{ij}}$ and cumulative production $\tilde{\bar{P}}_i(s)=\tilde{P}_i(s)/s=\sum_{j=1}^{m_i}Q_{ij}e^{-st_{ij}}/s$, cf. (8). The level of cumulative production immediately after the completion at t_{ik} is written $\bar{P}_{ik}=\sum_{j=1}^{k}Q_{ij}$. The possible times t_{ij} depend on the production plans for items downstream $\tilde{P}_k'(s)$, but these times, for an optimal plan, must belong to the set generated by $\left[\mathbf{H}'\tilde{\boldsymbol{\tau}}(s)\right]_{\text{row } i}\tilde{\mathbf{D}}'(s)$.

For simplicity, we temporarily drop the first index i (the item index), by writing a decision α_j, cumulative production $\bar{P}_j$, and number of events m, rather than α_{ij}, $\bar{P}_{ij}$, m_i, etc. We also let $\bar{\hat{D}}_j$ represents cumulative total requirements at time t_j (including both internal and external demand) and $\hat{D}_j=\bar{\hat{D}}_j-\bar{\hat{D}}_{j-1}$ denote requirements at t_j.

Consider Fig. 1. If $\alpha_j=0$, the cumulative production remains at its earlier level, $\bar{P}_j=\bar{P}_{j-1}$, i. e. $Q_j=0$, and if $\alpha_j=1$, then $\bar{P}_{j-1}=\bar{\hat{D}}_{j-1}$, which implies $\alpha_j\left(\bar{P}_{j-1}-\bar{\hat{D}}_{j-1}\right)=0$ and $\left(1-\alpha_j\right)\left(\bar{P}_j-\bar{P}_{j-1}\right)=0$. Subtraction between these two zero-valued expressions yields:

$$\left(1-\alpha_j\right)\bar{P}_j-\bar{P}_{j-1}+\alpha_j\bar{\hat{D}}_{j-1}=0\,. \tag{31}$$

This is a first-order difference equation in the cumulative $\bar{P}_j$, when given cumulative requirements $\bar{\hat{D}}_j$ and decisions α_j. The following theorem provides its solution.

Theorem 4: The difference equation (31) has the unique solution:

$$\overline{P}_j = \left(\hat{\overline{D}}_j + \sum_{k=j+1}^{m} \hat{D}_k \prod_{l=j+1}^{k} (1-\alpha_l) \right). \quad (32)$$

and, as a consequence, the lotsize at time t_j may be written:

$$Q_j = \alpha_j \left(\hat{D}_j + \sum_{k=j+1}^{m} \hat{D}_k \prod_{l=j+1}^{k} (1-\alpha_l) \right) = \alpha_j \sum_{k=j}^{m} \hat{D}_k \prod_{l=j+1}^{k} (1-\alpha_l). \quad (33)$$

Proof: Inserting the expression (32) for $\overline{P}_j$ and $\overline{P}_{j-1}$ into (31) immediately yields

$$\begin{aligned}
&\left(1-\alpha_j\right)\overline{P}_j - \overline{P}_{j-1} + \alpha_j \hat{\overline{D}}_{j-1} = \\
&= \left(1-\alpha_j\right)\left(\hat{\overline{D}}_j + \sum_{k=j+1}^{m} \hat{D}_k \prod_{l=j+1}^{k} (1-\alpha_l) \right) - \\
&-\left(\hat{\overline{D}}_{j-1} + \sum_{k=j}^{m} \hat{D}_k \prod_{l=j}^{k} (1-\alpha_l) \right) + \alpha_j \hat{\overline{D}}_{j-1} = \\
&= \left(1-\alpha_j\right)\hat{D}_j + \left(1-\alpha_j\right) \sum_{k=j+1}^{m} \hat{D}_k \prod_{l=j+1}^{k} (1-\alpha_l) - \\
&- \sum_{k=j}^{m} \hat{D}_k \prod_{l=j}^{k} (1-\alpha_l) = 0.
\end{aligned}$$

By taking the difference between (32) for subscripts j and $(j$ - $1)$, we obtain:

$$\begin{aligned}
Q_j &= \overline{P}_j - \overline{P}_{j-1} = \\
&= \hat{\overline{D}}_j + \sum_{k=j+1}^{m} \hat{D}_k \prod_{l=j+1}^{k} (1-\alpha_l) - \hat{\overline{D}}_{j-1} - \sum_{k=j}^{m} \hat{D}_k \prod_{l=j}^{k} (1-\alpha_l) = \\
&= \alpha_j \left(\hat{D}_j + \sum_{k=j+1}^{m} \hat{D}_k \prod_{l=j+1}^{k} (1-\alpha_l) \right).
\end{aligned}$$

Hence (32) satisfies (31) and so does (33), and there can be no other solution, since all expressions each generate a unique sequence. ■

By defining a triangular matrix $\mathbf{A}_m$ as

$$\mathbf{A}_m = \begin{bmatrix} \alpha_1 & \alpha_1(1-\alpha_2) & \cdots & \alpha_1(1-\alpha_2)\cdot\ldots\cdot(1-\alpha_m) \\ 0 & \alpha_2 & \cdots & \alpha_2(1-\alpha_3)\cdot\ldots\cdot(1-\alpha_m) \\ 0 & 0 & \ddots & \vdots \\ 0 & 0 & \cdots & \alpha_m \end{bmatrix}, \quad (34)$$

where m, as before, is the number of requirement events, (33) may be written in the convenient compact form:

$$\mathbf{Q} = \mathbf{A}_m \hat{\mathbf{D}}. \quad (35)$$

By further using a vector $\tilde{\mathbf{t}} = \left[e^{-t_1 s}, \ldots, e^{-t_m s} \right]$ identified from $\tilde{\hat{D}} = \tilde{\mathbf{t}}\hat{\mathbf{D}}$ to describe the timing of requirement events, production in binary form will be given by

$$\tilde{P}(s) = \tilde{\mathbf{t}}\mathbf{Q} = \tilde{\mathbf{t}}\mathbf{A}_m \hat{\mathbf{D}}. \quad (36)$$

Reintroducing the first subscript i (the item index) and the notation $\hat{D}_{ij}$ for requirements (internal plus external demand), it is now clear that the decisions α_{ij} constitute the coefficients of the production timing (setup) vector $\tilde{\mathbf{P}}'(s)$

$$\tilde{\mathbf{P}}'(s) = \begin{bmatrix} \tilde{P}_1'(s) \\ \vdots \\ \tilde{P}_n'(s) \end{bmatrix} = \begin{bmatrix} \sum_{k=1}^{m_1} \alpha_{1k} e^{-st_{1k}} \\ \vdots \\ \sum_{k=1}^{m_n} \alpha_{nk} e^{-st_{nk}} \end{bmatrix} = \begin{bmatrix} \tilde{\mathbf{t}}_1 \boldsymbol{\alpha}_{1m_1} \\ \vdots \\ \tilde{\mathbf{t}}_n \boldsymbol{\alpha}_{nm_n} \end{bmatrix}, \quad (37)$$

where we have introduced the vectors

$$\boldsymbol{\alpha}_{im_i} = \begin{bmatrix} \alpha_{i1} \\ \vdots \\ \alpha_{im_i} \end{bmatrix}, i = 1, 2, \ldots, n, \quad (38)$$

and so the production plan $\tilde{\mathbf{P}}(s)$, using (33) and (36), may be written in terms of the decisions $\boldsymbol{\alpha}_{im_i}$ and requirements $\hat{D}_{ij}$ only:

$$\tilde{\mathbf{P}}(s) = \begin{bmatrix} \tilde{P}_1(s) \\ \vdots \\ \tilde{P}_n(s) \end{bmatrix} = \begin{bmatrix} \sum_{j=1}^{m_1} Q_{1j} e^{-st_{1j}} \\ \vdots \\ \sum_{j=1}^{m_n} Q_{nj} e^{-st_{nj}} \end{bmatrix} =$$

$$= \begin{bmatrix} \sum_{j=1}^{m_1} \alpha_{1j} \sum_{k=j}^{m_1} \hat{D}_{1k} \prod_{l=j+1}^{k} (1-\alpha_{1l}) e^{-st_{1j}} \\ \vdots \\ \sum_{j=1}^{m_n} \alpha_{nj} \sum_{k=j}^{m_n} \hat{D}_{nk} \prod_{l=j+1}^{k} (1-\alpha_{nl}) e^{-st_{nj}} \end{bmatrix} = \begin{bmatrix} \tilde{\mathbf{t}}_1 \mathbf{A}_{1m} \hat{\mathbf{D}}_1 \\ \vdots \\ \tilde{\mathbf{t}}_n \mathbf{A}_{nm} \hat{\mathbf{D}}_n \end{bmatrix}. \quad (39)$$

Whereas external demand events never have zero demand. it may well happen that internal demand for an item is zero at some of the time points generated by the L4L solution, i.e. that some $\hat{\mathbf{D}}_i$ contain one or more zero-valued components.

5. THE TRIPLE ALGORITHM FOR DYNAMIC LOTSIZING

We devote this section to explaining the Triple Algorithm for the one-item case. In production-economic literature, the *dynamic lot sizing problem* is a problem frequently referred to.

The Triple Algorithm and the standard W-W algorithm apply to a deterministic one-level demand process allowing no backlogging, and similarly the Federgruen-Tzur algorithm, cf. Section 1, the latter developed for discrete time. Based our statement on the inner-corner condition in Section 4, the dynamic lotsizing problem may be stated in terms of a set of zero/one decision variables, because, at the optimum, production takes place only at inner corners. In the following, we apply this binary formulation to the Triple algorithm.

In the same way as at the end of the previous section, we drop the first index *i* (the index designating a specific item), so for this single item we have the requirements

$$\tilde{\tilde{\mathbf{D}}}(s) = \begin{bmatrix} \hat{D}_1 & \cdots & \hat{D}_m \end{bmatrix} \begin{bmatrix} e^{-st_1} \\ \vdots \\ e^{-st_m} \end{bmatrix} = \sum_{j=1}^{m} \hat{D}_j e^{-st_j} . \tag{40}$$

The dynamic lotsizing problem is now formulated in the following way. There is a finite horizon during which requirements for a single product are given by the amounts $\hat{D}_1$, $\hat{D}_2$, … $\hat{D}_m$ at times t_1, t_2, ... , t_m. Cumulative requirements are written $\bar{\hat{D}}_i = \sum_{j=1}^{i} \hat{D}_j$ and $\bar{\hat{D}}(t) = \bar{\hat{D}}_i$, $t_i \le t < t_{i+1}$. The issue is to determine the optimal amount to produce at each point in time when no shortages are allowed.

As an objective function we may apply either the average cost, or the net present value. In general, any cumulatively produced volume at time *t*, written $\bar{P}(t)$, satisfying $\bar{P}(t) \ge \bar{\hat{D}}(t)$, would be feasible.

According to (17), the economic consequences are limited to a setup cost *K* at each time a production takes place (a batch is completed) and a variable production cost *c* for each unit produced, since revenue in-payments are given and so their NPV, $\mathbf{p}\tilde{\mathbf{D}}(\rho)$. The continuous interest rate ρ is used when applying the NPV principle, and the inventory holding cost per unit and time unit is *h* when applying AC as the objective. In standard situations, but not always, see (Grubbström, 1980, Klein Haneveld and Teunter, 1998, Teunter and van der Laan, 2002), we normally have $h = \rho c$ when comparing the AC and NPV approaches with first-order approximations of the exponential functions concerned.

It is evident that if the setup cost for each batch is sufficiently high compared to the holding cost (or the interest rate), there must be exactly one setup at the very beginning (“All at Once”, $\forall @1$), and if it is sufficiently low, then a setup is afforded each time a requirement occurs (the L4L policy “As Required”). Therefore, the dynamic lotsizing problem is non-trivial, only when the ratio between these costs neither is very high, nor very low. One may thus conjecture that when this ratio falls into an intermediate range, more complexity is created, than outside of this range. The L4L policy harmonises with the popular idea of “just in time” as the best guidance. This topic, however, will be discussed elsewhere.

With *m* points in time with requirements, we are dealing with a staircase function of cumulative requirements having *m* steps. Our problem is to find the best out of the 2^{m-1} possible cumulative production staircases, in this one-product formulation. We choose the NPV (17) as the objective to be maximised and disregard revenues. Then we have

$$\mathrm{NPV}_{\mathrm{production}} = -K\tilde{P}'(\rho) - c\tilde{P}(\rho) =$$

$$-K\sum_{j=1}^{m} \alpha_j e^{-\rho t_j} - c\sum_{j=1}^{m} \alpha_j e^{-\rho t_j} \sum_{k=j}^{m} \hat{D}_k \prod_{l=j+1}^{k} (1-\alpha_l)\, e^{-\rho t_j} =$$

$$-K\tilde{\mathbf{t}}(\rho)\boldsymbol{\alpha}_m - c\tilde{\mathbf{t}}(\rho)\mathbf{A}_m\hat{\mathbf{D}}\, , \tag{41}$$

where (33), (36) and (38) are used. The first sum in the last member is the discounted value of the fixed setup costs, and the second term the discounted value of the production costs, each item requiring an out-payment of *c* when it is produced (completed).

If, instead applying the traditional *average cost* (AC) approach, the AC can be interpreted as a zeroth/first-order approximation of the *annuity stream* corresponding to the NPV, i. e. the equivalent constant cash flow generating the same NPV, cf. (11), where the holding cost *h* may be interpreted as ρc. But the net present value principle, generally speaking, is considered superior to the traditional average cost approach (Hadley, 1964, Trippi and Levin, 1974, Grubbström, 1980, Grubbström and Thorstenson, 1986, Kim, Philippatos and Chung, 1986, Thorstenson, 1988, Johansen and Thorstenson, 1996, Grubbström and Kingsman, 2004).

The Triple Algorithm (Grubbström, 1999b, 2005) works as follows. We consider three points in time at which demand events occur t_i, t_j, and t_k, where $i < j < k$, cf. Fig. 2. If there is a setup at t_j (the in-between time), this setup has to be justified by the inventory reduction represented by the shaded rectangle exceeding consequences of the setup cost *K*.

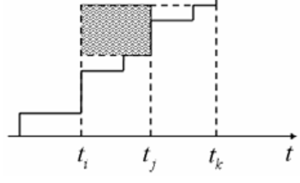

Fig. 2: The shaded area represents g_{ijk} (Eqs (42)-(43)), its size determining whether or not a setup at t_j is justified.

The setup cost K is temporally located at t_j and has an NPV discounted to t_i amounting to $Ke^{-\rho(t_j-t_i)}$. The inventory reduction in physical units amounts to $\bar{\hat{D}}_{k-1}-\bar{\hat{D}}_{j-1}$ and has a corresponding NPV amounting to $-c\left(1-e^{-\rho(t_j-t_i)}\right)\left(\bar{\hat{D}}_{k-1}-\bar{\hat{D}}_{j-1}\right)$, from moving the production cost $c\left(\bar{\hat{D}}_{k-1}-\bar{\hat{D}}_{j-1}\right)$ from t_i to t_j. Comparing these consequences by the ratio g_{ijk} defined as

$$g_{ijk}=\frac{c\left(1-e^{-\rho(t_j-t_i)}\right)\left(\bar{\hat{D}}_{k-1}-\bar{\hat{D}}_{j-1}\right)}{Ke^{-\rho(t_j-t_i)}}=$$

$$=\frac{c\left(e^{\rho(t_j-t_i)}-1\right)\left(\bar{\hat{D}}_{k-1}-\bar{\hat{D}}_{j-1}\right)}{K}, \tag{42}$$

the intermediate setup at t_j is justified when $g_{ijk}>1$, and not justified in the converse case. Assuming ρ to be small, a Maclaurin expansion of the exponential factor $e^{\rho(t_j-t_i)}\approx 1+\rho(t_j-t_i)$, we obtain the corresponding ratio for the average cost case

$$g_{ijk}=\frac{\rho c\left(t_j-t_i\right)\left(\bar{\hat{D}}_{k-1}-\bar{\hat{D}}_{j-1}\right)}{K}. \tag{43}$$

where ρc is interpreted as the inventory holding cost. With this approximation, the inventory savings are directly proportional to the shaded area in Fig. 2, i. e. to $\left(t_j-t_i\right)\left(\bar{\hat{D}}_{k-1}-\bar{\hat{D}}_{j-1}\right)$. The ratio g_{ijk} forms the basic measure on which of the Triple algorithm is founded.

Assume that there is a setup at t_i. The next two setups t_j, and t_k cannot be considered as candidates for the optimal solution unless $g_{ijk}>1$. Furthermore, for t_k temporarily assumed known, the in between t_j must be chosen so as to maximise the ratio g_{ijk}, otherwise further savings would be possible by adjusting t_j. When k is increased from $i+1$ and upwards, the next setup after i has to occur at the *latest* for a *highest k* such that $g_{ijk}\le 1$ for all in between values of j, $i<j<k$, otherwise a setup should have been taken at an intermediate t_j (and not wait until t_k). With an increasing k, the intermediate ratios are increasing for any intermediate j (as seen from Fig. 2).

Hence, for a given i, we determine a latest point $L(i)$ by:

$$L(i)=\operatorname*{Max}_{j,\,g_{ijk}\le 1} k \,. \tag{44}$$

For k above $L(i)$, there must be a setup somewhere in the interval between i and k. For both k and the intermediate j to be optimal, the j must be chosen to maximise g_{ijk}. As k increases, the maximising argument j cannot decrease. Viewing j as a function of k, this function is non-decreasing. Therefore, if the setup at t_i is known, the *earliest* time index at which a next possible setup is optimal, written $E(i)$, is for the j maximising $g_{ij,L(i)+1}$:

$$E(i)=\operatorname*{Argmax}_{j} g_{ij,L(i)} \,. \tag{45}$$

For any i, we can therefore define an interval [$E(i)$, $L(i)$], within which the next optimal setup must occur.

For a given i, we investigate which next k (second next setup) that can be a possible candidate. This k is at least $L(i)+1$ and not higher than that the intermediate maximising j falls within [$E(i)$, $L(i)$]. Any given i thus generates a set of possible triples (i, j, k) which are consistent with this condition. Starting at $t_1=0$, at which there must be a setup, a first set of triples (1, j, k) is generated. For each such triple, starting with its intermediate element, a new set of triples is generated, and so on.

For an optimal sequence of setups, consecutive triples must be consistent with each other in the sense that the later interval of one triple (t_k-t_j) equals the former interval (t_j-t_i) of the triple following. While generating the sets of triples, one may therefore discard any triple that does not meet this requirement. The remaining triples must be compatible forwards and backwards. Striking out a triple at a late stage, might provide reductions also at very early stages, and vice versa.

This procedure generates a limited set of feasible paths (sequences), of which (at least) one is optimal. Having reached the end stage, the remaining sequences may be evaluated by complete enumeration determining the optimum.

6. SOLUTION TO THE LOTSIZING PROBLEM FOR A GENERAL ASSEMBLY SYSTEM WITHIN MRP THEORY

Having explained basic MRP Theory, developed expressions for how demand events propagate upstream in Sections 3 - 4, and outlined the Triple Algorithm for dynamic lotsizing in Section 5, we are now in a position to state and solve the problem of lotsizing for a general multi-level, multi-product system having an assembly structure.

The very maximum number of different time points at which requirements can occur, has been determined in Section 3. Essentially, this timing is given by the L4L lotsizing policy. However, on a lower level (upstream) requirements might not occur at all points, since production at the higher level might restrict itself to only using a smaller subset of the points available, and this in its turn reduces the number of available points on lower levels. A fewer number of terms in $\tilde{P}_i(s)$ (and therefore in $\tilde{P}_i'(s)$) will reduce the number of possible optimal time points of production on lower levels, i. e. possible terms in $\tilde{P}_j(s)$, $j>i$.

Let us consider the case that there is only one end product, i. e. there is never any demand for spare parts etc. In such a case, only $M_1 > 0$, and all other $M_j = 0$, for $j > 0$. Then,

$$m_i = \sum_{j=1}^{i-1} 2^{i-j-1} M_j + M_i = M_1 \times \begin{cases} 1, \; i = 1, \\ 1, \; i = 2, \\ 2^{i-2}, \; i > 2. \end{cases} \tag{46}$$

So, at the bottom level n, the maximum number of points in time to consider will be $2^{n-2} M_1$. If instead all levels have a similar external demand, say M, we have

$$m_i = M \sum_{j=1}^{i-1} 2^{i-j-1} + M = M\left(\frac{2^{i-1}-1}{2-1} + 1\right) = 2^{i-1} M , \tag{47}$$

which is (only) twice the number of the former case, if $M = M_1$.

We know that total requirements for item i, are generated only by production events $\tilde{P}_k'(s)$ on higher levels, $k < i$, and its external demand $\tilde{D}_i(s)$. Given these downstream events (and external demand), provides the timing belonging to $\tilde{P}_i'(s)$ as well as the size of requirements at each time, i. e. the staircase $\hat{\tilde{D}}_i(s)/s$. We can therefore apply the Triple Algorithm to the NPV of each $\tilde{P}_i'(s)$ given all downstream $\tilde{P}_k'(s)$, provided that the NPV of the optimal plan for upstream production is added to the objective function. This forms a dynamic programming procedure enabling an optimisation of the system as a whole.

As a description of the state at stage i, we choose the requirements $\hat{\tilde{D}}_i = \tilde{\mathbf{t}}_i \hat{\mathbf{D}}_i$, where the timing $\tilde{\mathbf{t}}_i$ can be computed in advance from the L4L solution. The coefficients in $\hat{\mathbf{D}}_i$ depend on decisions at earlier stages (downstream) according to

$$\left[\tilde{\mathbf{t}}_i \hat{\mathbf{D}}_i\right] = \sum_{k=1}^{i-1} \left[\mathbf{H}\tilde{\boldsymbol{\tau}}(s)\right]_{ik} \tilde{\mathbf{t}}_k \mathbf{A}_{km_k} \hat{\mathbf{D}}_k + \tilde{\mathbf{D}}_k(s) , \tag{48}$$

where the matrices $\mathbf{A}_{km_k}$ contain elements depending on decisions at stage k only, i. e. on $\boldsymbol{\alpha}_{km_k}$ (defined in (38)), $k = 1 , \ldots , i - 1$. So (48) defines the state transition, when decisions are taken at stage i - 1.

Thus, if we use this vector to collect the binary decisions for the production events $\tilde{P}_i'(s)$ and let $W_{i+1}\left(\boldsymbol{\alpha}_{im_i}, \tilde{\mathbf{t}}_i \hat{\mathbf{D}}_i\right)$ denote the maximum NPV of the production of all upstream items $k > i$, when given $\tilde{\mathbf{t}}_i \hat{\mathbf{D}}_i$, then the problem may be formulated in the following recursive manner:

$$W_i\left(\boldsymbol{\alpha}_{i-1m_{i-1}}, \tilde{\mathbf{t}}_{i-1}\hat{\mathbf{D}}_{i-1}\right) = \\ = \underset{\boldsymbol{\alpha}_{im_i}}{\mathrm{Max}} \left\{ \mathrm{NPV}_{\mathrm{production}}\left(\boldsymbol{\alpha}_{im_i} \middle| \tilde{\mathbf{t}}_i \hat{\mathbf{D}}_i\right) + W_{i+1}\left(\boldsymbol{\alpha}_{im_i}, \tilde{\mathbf{t}}_i \hat{\mathbf{D}}_i\right)\right\} , \tag{49}$$

where $\mathrm{NPV}_{\mathrm{production}}\left(\boldsymbol{\alpha}_{im_i} \middle| \tilde{\mathbf{t}}_i \hat{\mathbf{D}}_i\right) = -K_i \tilde{\mathbf{t}}_i \boldsymbol{\alpha}_{im_i} - c_i \tilde{\mathbf{t}}_i \mathbf{A}_{im_i} \hat{\mathbf{D}}_i$ is the contribution to the NPV from the ith level. For $i = n$, we have $W_{n+1}\left(\boldsymbol{\alpha}_{nm_n}\right) = 0$, so $W_n\left(\boldsymbol{\alpha}_{n-1m_{n-1}}, \tilde{\mathbf{t}}_{n-1}\hat{\mathbf{D}}_{n-1}\right) = \underset{\boldsymbol{\alpha}_{nm_n}}{\mathrm{Max}} \left\{ \mathrm{NPV}_{\mathrm{production}}\left(\boldsymbol{\alpha}_{nm_n} \middle| \tilde{\mathbf{t}}_n \hat{\mathbf{D}}_n\right)\right\}$, etc. When carrying out each such maximisation, the Triple Algorithm may be applied. If a component of $\hat{\mathbf{D}}_i$ happens to be zero-valued, the corresponding component of $\boldsymbol{\alpha}_{im_i}$ must be set to zero. This may create a substantial reduction in the size of the problem.

7. NUMERICAL EXAMPLE

As a numerical example, we choose the five-item product structure in Fig. 3.

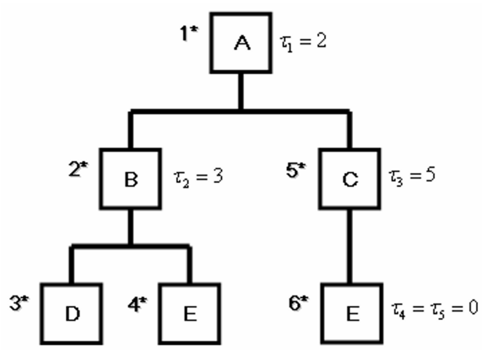

Fig. 3. Product structure of example, showing input matrix elements and lead times.

The input matrix, lead time matrix and generalised Leontief inverse are given by:

$$\mathbf{H} = \begin{matrix} \mathrm{A} \\ \mathrm{B} \\ \mathrm{C} \\ \mathrm{D} \\ \mathrm{E} \end{matrix} \begin{bmatrix} 0 & 0 & 0 & 0 & 0 \\ 2 & 0 & 0 & 0 & 0 \\ 5 & 0 & 0 & 0 & 0 \\ 0 & 3 & 0 & 0 & 0 \\ 0 & 4 & 6 & 0 & 0 \end{bmatrix}, \quad \tilde{\boldsymbol{\tau}}(s) = \begin{bmatrix} e^{2s} & 0 & 0 & 0 & 0 \\ 0 & e^{3s} & 0 & 0 & 0 \\ 0 & 0 & e^{5s} & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix},$$

$$\left(\mathbf{I} - \mathbf{H}\tilde{\boldsymbol{\tau}}(s)\right)^{-1} = \begin{matrix} \mathrm{A} \\ \mathrm{B} \\ \mathrm{C} \\ \mathrm{D} \\ \mathrm{E} \end{matrix} \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 2e^{2s} & 1 & 0 & 0 & 0 \\ 5e^{2s} & 0 & 1 & 0 & 0 \\ 6e^{5s} & 3e^{3s} & 0 & 1 & 0 \\ \left(8e^{5s} + 30e^{7s}\right) & 4e^{3s} & 6e^{5s} & 0 & 1 \end{bmatrix}.$$

Let external demand be

$$\tilde{\mathbf{D}}(s)=\begin{bmatrix}\tilde{D}_1(s)\\ \tilde{D}_2(s)\\ 0\\ 0\\ 0\end{bmatrix}=\begin{bmatrix}2e^{-12s}+3e^{-14s}\\ 4e^{-13s}\\ 0\\ 0\\ 0\end{bmatrix},$$

meaning that two units of item A are demanded at time $t = 12$, three at time $t = 14$, and that four units of item B are demanded at $t = 13$. The L4L solution will then become

$$\tilde{\mathbf{P}}_{\mathrm{L4L}}(s)=\left(\mathbf{I}-\mathbf{H}\tilde{\boldsymbol{\tau}}(s)\right)^{-1}\tilde{\mathbf{D}}(s)=$$

$$\begin{bmatrix}1&0&0&0&0\\ 2e^{2s}&1&0&0&0\\ 5e^{2s}&0&1&0&0\\ 6e^{5s}&3e^{3s}&0&1&0\\ \left(8e^{5s}+30e^{7s}\right)&4e^{3s}&6e^{5s}&0&1\end{bmatrix}\begin{bmatrix}2e^{-12s}+3e^{-14s}\\ 4e^{-13s}\\ 0\\ 0\\ 0\end{bmatrix}=$$

$$=\begin{bmatrix}2e^{-12s}+3e^{-14s}\\ 4e^{-10s}+6e^{-12s}+4e^{-13s}\\ 10e^{-10s}+15e^{-12s}\\ 12e^{-7s}+18e^{-9s}+12e^{-10s}\\ 60e^{-5s}+106e^{-7s}+24e^{-9s}+16e^{-10s}\end{bmatrix},$$

which shows the volumes completed as well as their timing for all five items. The setup vector for this L4L case will be

$$\tilde{\mathbf{P}}'(s)=\left(\mathbf{I}-\mathbf{H}'\tilde{\boldsymbol{\tau}}(s)\right)^{-1}\tilde{\mathbf{D}}'(s)=$$

$$\begin{bmatrix}1&0&0&0&0\\ e^{2s}&1&0&0&0\\ e^{2s}&0&1&0&0\\ e^{5s}&e^{3s}&0&1&0\\ \left(e^{5s}+e^{7s}\right)&e^{3s}&e^{5s}&0&1\end{bmatrix}\begin{bmatrix}e^{-12s}+e^{-14s}\\ e^{-13s}\\ 0\\ 0\\ 0\end{bmatrix}=$$

$$\begin{bmatrix}e^{-12s}+e^{-14s}\\ e^{-10s}+e^{-12s}+e^{-13s}\\ e^{-10s}+e^{-12s}\\ e^{-7s}+e^{-9s}+e^{-10s}\\ e^{-5s}+e^{-7s}+e^{-9s}+e^{-10s}\end{bmatrix}.$$

In contrast, for the "All at Once" solution we have:

$$\tilde{\mathbf{P}}_{\forall @1}(s)=\begin{bmatrix}5e^{-12s}\\ 14e^{-10s}\\ 25e^{-10s}\\ 42e^{-7s}\\ 206e^{-5s}\end{bmatrix}.$$

Requirements making up internal and external demand are obtained as:

$$\tilde{\tilde{\mathbf{D}}}(s)=\mathbf{H}\tilde{\boldsymbol{\tau}}(s)\tilde{\mathbf{P}}(s)+\tilde{\mathbf{D}}(s)=$$

$$=\begin{bmatrix}0&0&0&0&0\\ 2e^{2s}&0&0&0&0\\ 5e^{2s}&0&0&0&0\\ 0&3e^{3s}&0&0&0\\ 0&4e^{3s}&6e^{5s}&0&0\end{bmatrix}\tilde{\mathbf{P}}(s)+\begin{bmatrix}2e^{-12s}+3e^{-14s}\\ 4e^{-13s}\\ 0\\ 0\\ 0\end{bmatrix}.$$

We may now write the production opportunities (48) in binary form. With $\hat{\mathbf{D}}_1=\begin{bmatrix}\hat{D}_{11}\\ \hat{D}_{12}\end{bmatrix}=\begin{bmatrix}2\\ 3\end{bmatrix}$ and $\tilde{\mathbf{t}}_1=\left[e^{-12s},e^{-14s}\right]$, using (36) we have $\tilde{P}_1(s)=\tilde{\mathbf{t}}_1\mathbf{A}_{12}\hat{\mathbf{D}}_1=\ \alpha_{11}\left(2+3\left(1-\alpha_{12}\right)\right)e^{-12s}$ $+3\alpha_{12}e^{-14s}$, where the first index of $\mathbf{A}_{12}$ is item index and the second the number of events. Since $\tilde{P}_1(s)=\ \tilde{\mathbf{t}}_1\mathbf{A}_{12}\hat{\mathbf{D}}_1=$ $\left(\alpha_{11}\left(2+3\left(1-\alpha_{12}\right)\right)e^{-12s}+\ 3\alpha_{12}e^{-14s}\right)$, we identify

$$\hat{\mathbf{D}}_2=\begin{bmatrix}\hat{D}_{21}\\ \hat{D}_{22}\\ \hat{D}_{23}\end{bmatrix}=\begin{bmatrix}\alpha_{11}\left(4+6\left(1-\alpha_{12}\right)\right)\\ 6\alpha_{12}\\ 4\end{bmatrix}$$

and

$$\tilde{\mathbf{t}}_2=\left[e^{-10s},e^{-12s},e^{-13s}\right].$$

Hence,

$$\tilde{P}_2(s)=\left[e^{-10s},e^{-12s},e^{-13s}\right]\mathbf{A}_{23}\begin{bmatrix}2\alpha_{11}\left(2+3\left(1-\alpha_{12}\right)\right)\\ 6\alpha_{12}\\ 4\end{bmatrix}=$$

$$=\left(\alpha_{21}\begin{pmatrix}2\alpha_{11}\left(2+3\left(1-\alpha_{12}\right)\right)\\ +6\alpha_{12}(1-\alpha_{22})+4(1-\alpha_{22})(1-\alpha_{23})\end{pmatrix}\right)e^{-10s}$$

$$+\left(\alpha_{22}\left(6\alpha_{12}+4(1-\alpha_{23})\right)\right)e^{-12s}+4\alpha_{23}e^{-13s}.$$

Further computations yield

$$\hat{\mathbf{D}}_3=\begin{bmatrix}\alpha_{11}\left(10+15\left(1-\alpha_{12}\right)\right)\\ 15\alpha_{12}\end{bmatrix},\ \tilde{\mathbf{t}}_3=\left[e^{-12s},e^{-14s}\right],$$

and

$$\tilde{P}_3(s) = \tilde{\mathbf{t}}_3\mathbf{A}_{23}\hat{\mathbf{D}}_3 = \left(\alpha_{11}\left(10+15\left(1-\alpha_{12}\right)\right)\alpha_{31} + 15\alpha_{12}\alpha_{31}\left(1-\alpha_{32}\right)\right)e^{-10s} + 15\alpha_{12}\alpha_{32}e^{-12s},$$

$$\hat{\mathbf{D}}_4 = \begin{bmatrix} \left[6\alpha_{11}\left(2+3(1-\alpha_{12})\right)\alpha_{21} + 18\alpha_{12}\alpha_{21}\left(1-\alpha_{22}\right) \right. \\ \left. +12\alpha_{21}\left(1-\alpha_{22}\right)\left(1-\alpha_{23}\right)\right] \\ 18\alpha_{12}\alpha_{22} + 12\alpha_{22}\left(1-\alpha_{23}\right) \\ 12\alpha_{23} \end{bmatrix}$$

and

$$\tilde{\mathbf{t}}_4 = \left[e^{-7s}, e^{-9s}, e^{-10s}\right],\ \tilde{P}_4(s) = \tilde{\mathbf{t}}_4\mathbf{A}_{43}\hat{\mathbf{D}}_4,$$

and finally,

$$\hat{\mathbf{D}}_5 = \begin{bmatrix} \alpha_{11}\left(60+90\left(1-\alpha_{12}\right)\right)\alpha_{31} + 90\alpha_{12}\alpha_{31}\left(1-\alpha_{32}\right) \\ \left[\alpha_{11}\left(16+24\left(1-\alpha_{12}\right)\right)\alpha_{21} + 24\alpha_{12}\alpha_{21}\left(1-\alpha_{22}\right) \right. \\ \left. +16\alpha_{21}\left(1-\alpha_{22}\right)\left(1-\alpha_{23}\right) + 90\alpha_{12}\alpha_{32}\right] \\ 24\alpha_{12}\alpha_{22} + 16\alpha_{22}\left(1-\alpha_{23}\right) \\ 16\alpha_{23} \end{bmatrix},$$

and

$$\tilde{\mathbf{t}}_5 = \left[e^{-5s}, e^{-7s}, e^{-9s}, e^{-10s}\right],\ \tilde{P}_5(s) = \tilde{\mathbf{t}}_5\mathbf{A}_{54}\hat{\mathbf{D}}_5.$$

It is easily verified that the L4L solution is obtained when all α_{ij} are set to unity, and the $\forall@1$ solution when only the α_{i1} are unit-valued.

The production timing may be written $\tilde{P}_i'(s) = \tilde{\mathbf{t}}_i\boldsymbol{\alpha}_{im_i}$. All $\tilde{\mathbf{t}}_i$ and m_i are found from the explicit L4L solution $\left(\mathbf{I}-\mathbf{H}\tilde{\boldsymbol{\tau}}(s)\right)^{-1}\tilde{\mathbf{D}}(s)$.

As economic parameters, let us further assume the production and setup cost vectors

$\mathbf{c} = [200,\ 180,\ 160,\ 140,\ 120]$, $\mathbf{K} = [110,\ 100,\ 130,\ 80,\ 70]$,

and an interest rate $\rho = 0.01$. The NPV is then

$$\text{NPV}_{\text{production}} = -\mathbf{c}\tilde{\mathbf{P}}(\rho) - \mathbf{K}\tilde{\mathbf{P}}'(\rho) =$$

$$-[200,\ 180,\ 160,\ 140,\ 120]\begin{bmatrix} \tilde{\mathbf{t}}_1\mathbf{A}_{12}\hat{\mathbf{D}}_1 \\ \tilde{\mathbf{t}}_2\mathbf{A}_{23}\hat{\mathbf{D}}_2 \\ \tilde{\mathbf{t}}_3\mathbf{A}_{33}\hat{\mathbf{D}}_3 \\ \tilde{\mathbf{t}}_4\mathbf{A}_{43}\hat{\mathbf{D}}_4 \\ \tilde{\mathbf{t}}_4\mathbf{A}_{54}\hat{\mathbf{D}}_5 \end{bmatrix}_{s=\rho=0.01}$$

$$-[110,\ 100,\ 130,\ 80,\ 70]\begin{bmatrix} \tilde{\mathbf{t}}_1\boldsymbol{\alpha}_{12} \\ \tilde{\mathbf{t}}_2\boldsymbol{\alpha}_{23} \\ \tilde{\mathbf{t}}_3\boldsymbol{\alpha}_{33} \\ \tilde{\mathbf{t}}_4\boldsymbol{\alpha}_{43} \\ \tilde{\mathbf{t}}_4\boldsymbol{\alpha}_{54} \end{bmatrix}_{s=\rho=0.01}.$$

With the given parameter values, the optimum is found for the solution:

$$\boldsymbol{\alpha}_{12} = \begin{bmatrix}1\\1\end{bmatrix},\ \boldsymbol{\alpha}_{23} = \begin{bmatrix}1\\0\\0\end{bmatrix},\ \boldsymbol{\alpha}_{32} = \begin{bmatrix}1\\1\end{bmatrix},\ \boldsymbol{\alpha}_{43} = \begin{bmatrix}1\\0\\0\end{bmatrix},\ \boldsymbol{\alpha}_{54} = \begin{bmatrix}1\\1\\0\\0\end{bmatrix},$$

which gives an $\text{NPV}_{\text{production}}$ = -36,141, of which the fixed setup charges account for -741, and the production variable out-payments for -35,400. This solution gives an optimal production plan $\tilde{\mathbf{P}}(s)$ and an optimal production timing vector $\tilde{\mathbf{P}}'(s)$ according to:

$$\tilde{\mathbf{P}}(s) = \begin{bmatrix} 2e^{-12s} + 3e^{-14s} \\ 14e^{-10s} \\ 10e^{-10s} + 15e^{-12s} \\ 42e^{-7s} \\ 60e^{-5s} + 146e^{-7s} \end{bmatrix},\ \tilde{\mathbf{P}}'(s) = \begin{bmatrix} e^{-12s} + e^{-14s} \\ e^{-10s} \\ e^{-10s} + e^{-12s} \\ e^{-7s} \\ e^{-5s} + e^{-7s} \end{bmatrix}.$$

We may compare this solution with the two extreme replenishment policies L4L ("As Required") and the converse policy $\forall@1$ ("All-at-Once"). With L4L replenishments, all $\boldsymbol{\alpha}_{im_i}$ have unit values in all positions. For the $\forall@1$ policy, there will be zero values for all components of $\boldsymbol{\alpha}_{im_i}$ except the first components having unit values. The $\forall@1$ policy gives $\text{NPV}_{\text{production}}$ = -36,239, made up of a setup contribution amounting to -456, and a variable cost contribution to -35,783. Instead, the L4L policy yields $\text{NPV}_{\text{production}}$ = -36,353, with a setup contribution of -1,190 and a variable production cost contribution of -35,162.

If we instead choose the average cost principle (AC), the optimum will be found for

$$\boldsymbol{\alpha}_{12} = \begin{bmatrix}1\\0\end{bmatrix},\ \boldsymbol{\alpha}_{23} = \begin{bmatrix}1\\0\\0\end{bmatrix},\ \boldsymbol{\alpha}_{32} = \begin{bmatrix}1\\0\end{bmatrix},\ \boldsymbol{\alpha}_{43} = \begin{bmatrix}1\\0\\0\end{bmatrix},\ \boldsymbol{\alpha}_{54} = \begin{bmatrix}1\\1\\0\\0\end{bmatrix}.$$

which differs from the NPV optimum by two less setups. Here, total costs amount to 604, of which setup costs are 570 and holding costs 34. The L4L and $\forall@1$ solutions give total costs of 1,320 and 668, respectively, of which setup costs are 1,320 and 500, respectively, and holding costs zero and 168, respectively. The holding costs are based on available inventory and therefore do not include costs for work-in-progress, which amount to 296 in all cases.

If the NPV is computed for the AC solution, we have $\mathrm{NPV}_{\mathrm{production}} = -36{,}178$, of which the setups contribute with -521 and production costs with -35,657. This creates a cost increase of 37, compared to the NPV optimum, and the other solutions do not differ much in their consequences either, a circumstance which is common when balancing setup and holding charges. The most costly solution generates a mere $\mathrm{NPV}_{\mathrm{production}} = -36596$.

The structure of this example is as follows. We have five stages with five vectors $\boldsymbol{\alpha}_{im_i}$, $i = 1, \ldots , 5$, holding a total of $\sum_{i=1}^{5} m_i = 14$ binary variables, of which each must have a unit first component. Therefore, there are $2^9 = 512$ solutions to choose from, which is not a large number. One potential requirement event disappeared concerning item E. This is due to two terms in the generalised Leontief inverse in position (5, 1) by coincidence having the same exponential factor e^{-7s}. Otherwise there would have been 10 decision variables and 1024 solutions. But the number is also reduced due to zero internal demand at certain points in time. In this example there are 402 such cases reducing the number from 512 to 110.

However, as shown earlier, this number increases rapidly when the number of product structure levels and number of external demand events increase, as well as when the input matrix is more filled with positive elements. If the matrix would have been filled completely with positive elements beneath its diagonal, assuming the same external demand events, from (29) we would have had 35 binary variables (rather than 9 or 10) and a total of $2^{35} = 549{,}755{,}813{,}888$ solutions to choose from (rather than 512 or 1024).

Fig. 4 contains a scatter diagram of all 110 solutions of this example, indicating the relationship between the four extreme solutions.

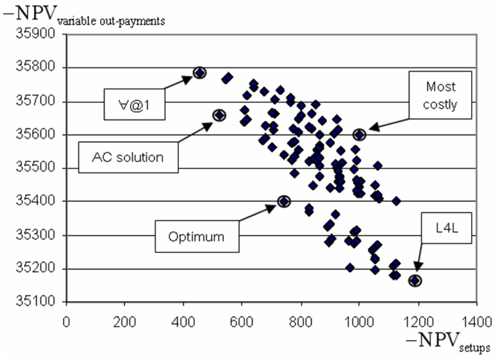

Fig. 4. Scatter diagram of solutions. Encircled points show the four extreme and the AC solutions.

8. CONCLUSIONS

In the foregoing, we have analysed the problem of deciding an optimal production plan for a general assembly system, when maximising the Net Present Value. As a background, we developed expressions for how external demand events (the "Master Production Schedule" in MRP terminology) generate requirement events on lower levels upstream in the production system. Adopting the same representation for decisions as when applying the Triple Algorithm, namely introducing binary decision variables for having a setup or not at time points determined by each requirement event, enabled us to formulate the general problem in a compact way utilising the framework of MRP Theory. The final findings were in the form of a recursive equation for NPV maximisation, in parallel with an equation explaining the development of requirements for each stage in the decision process, the latter acting as the state transition equation.

The simplicity in the structure of the solution, due to its compact formulation, makes it very easy to design a computer program, intended to find the optimal production plan for any general assembly system.

In further studies, the natural line of development, on the one hand, will be to introduce stochastics into external demand, allowing for such demand to be backlogged or lost, and introducing safety precautions in the form of safety stocks and/or safety lead times. Secondly, the introduction of capacity constraints into this type of lotsizing decisions will be a further important extension.

This will be a challenging continuation.

REFERENCES

Aseltine, J.A. (1958). *Transform Method in Linear System Analysis,* McGraw-Hill, New York, N.Y.

Bogataj, M. (1999). Inventory allocation and customer travelling problem in spatial duopoly. *International Journal of Production Economics*, 59, pp. 271-279.

Bogataj, D., Bogataj, M. (2007). Measuring the supply chain risk and vulnerability in frequency space, *International Journal of Production Economics*, 108, pp. 291-301.

Bogataj, M., Bogataj, L. (2001). Supply chain coordination in spatial games, *International Journal of Production Economics*, 71, pp. 277-286.

Bogataj, M., Bogataj, L. (2003). On the compact presentation of the lead time perturbations in distribution networks, *International Journal of Production Economics*, 88, pp. 145-155.

Bogataj, L., Ferbar, L. (1996). Stochastic considerations of Grubbström-Molinder model of MRP, input-output and multi-echelon inventory systems, *International Journal of Production Economics*, 45, pp. 329-336.

Buser, S. A. (1986). LaPlace Transforms as Present Value Rules: A Note , *Journal of Finance*, 41, pp. 243-247.

Churchill, R. V. (1958). *Operational Mathematics,* McGraw-Hill, New York, N.Y.

Churchill, R. V. (1960). *Complex Variables and Applications,* McGraw-Hill, New York, N.Y.

Federgruen, A., Tzur, M. (1991). A Simple Forward Algorithm to solve General Dynamic Lot Sizing Models with *n* Periods in 0(*n* log *n*) or 0(*n*) Time, *Management Science*, 37, pp. 909-925.

Ferbar, L., Bogataj, L. (1999). A market game with the characteristic function according to the MRP and input-output analysis model. *International Journal of Production Economics*, 59, pp. 281-288.

Grubbström, R. W. (1967). On the application of the Laplace transform to certain economic problems, *Management Science*, 13, pp. 558-567.

Grubbström, R. W. (1980). A principle for determining the correct capital costs of work-in-progress and inventory, *International Journal of Production Research,* 18, pp. 259-271.

Grubbström, R. W. (1990). The Distribution of an Additive in a Chemical Process - An Application of Input-Output Theory, *Engineering Costs and Production Economics,* 19, 333-340.

Grubbström, R. W. (1996a). Material Requirements Planning and Manufacturing Resources Planning, in Warner, M., (Ed.), *International Encyclopedia of Business and Management,* Routledge, London.

Grubbström, R. W. (1996b). Stochastic Properties of a Production-Inventory Process With Planned Production Using Transform Methodology, *International Journal of Production Economics*, 45, pp. 407-419.

Grubbström, R. W. (1998). A Net Present Value Approach to Safety Stocks in Planned Production , *International Journal of Production Economics*, 56-57, pp. 213-229.

Grubbström, R. W. (1999a). The Fundamental Equations of MRP Theory of a Simple Feed-Back System with Lead Times, *Working Paper WP-264*, Department of Production Economics, Linköping Institute of Technology, June.

Grubbström, R. W. (1999b). An Optimal Forward Dynamic Lot-Sizing Policy for the NPV and Average Cost Cases in Continuous Time, in Hillery, M.T., Lewis, H.J., (Eds), *The 15th International Conference on Production Research. ICPR 15 Manufacturing for a Global Market,* Limerick, Ireland, August 9-13, pp. 381-384.

Grubbström, R. W. (1999c). A net present value approach to safety stocks in a multi-level MRP system, *International Journal of Production Economics*, 59, pp. 361-365.

Grubbström, R. W. (2005). Algorithms for Optimal Dynamic Lotsizing in Discrete and Continuous Time, *International Journal of Agile Manufacturing*, 8(2).

Grubbström, R. W. (2007). Transform Methodology Applied to Some Inventory Problems, *Zeitschrift für Betriebswirtschaft*, 77, pp. 297-324.

Grubbström, R. W. (2009). Extensions of MRP Theory with Application to Distribution and Remanufacturing, *Proceedings, Second International Workshop on Successful Strategies in Supply Chain Management*, January 8-9, Hong Kong, pp. 3-19.

Grubbström, R. W., Bogataj, L., (Eds), (1998). *Input-Output Analysis and Laplace Transforms in Material Requirements Planning*. Storlien 1997. FPP, Portorož.

Grubbström, R. W., Bogataj, L., Bogataj, M. (2007). A Compact Representation and Optimisation of Distribution and Reverse Logistics in the Value Chain, WP, (*Mathematical economics, operational reseach and logistics*, 5), Ljubljana: Faculty of Economics, KMOR, Slovenia.

Grubbström, R. W., Bogataj, L., Bogataj, M. (2008). Extensions of MRP Theory from Assembly Systems to Arborescent Systems, *Working Paper*, Mediterranean Institute for Advanced Studies, Šempeter pri Gorici, Slovenia.

Grubbström, R. W., Huynh, T.T.T. (2006). Analysis of standard ordering policies within the framework of MRP theory, *International Journal of Production Research*, 44, pp. 3759-3773.

Grubbström, R. W., Jiang, Y. (1990). A survey and analysis of the application of the Laplace transform to present value problems, *Revista di matematica per le scienze economiche e sociale*, 12, 43-62.

Grubbström, R. W., Kingsman, B. G. (2004). Ordering and Inventory Policies for Step Changes in the Unit Item Cost: A Discounted Cash Flow Approach, *Management Science*, 50, pp. 253-267.

Grubbström, R. W., Molinder, A. (1994). Further theoretical consideration on the relationship between MRP, Input-Output analysis and multi echelon inventory systems, *International Journal of Production Economics*, 35, pp. 299-311.

Grubbström, R. W., Tang, O. (2000). An Overview of Input-Output Analysis Applied to Production-Inventory Systems, *Economic Systems Review,* 12, pp. 3-25.

Grubbström, R. W., Thorstenson, A. (1986). Evaluation of capital costs in a multi-level inventory system by means of the annuity stream principle, *European Journal of Operational Research*, 24, pp. 136-145.

Grubbström, R. W., Wang, Z. (2003). A Stochastic Model of Multi-Level/Multi-Stage Capacity-Constrained Production-Inventory Systems, *International Journal of Production Economics,* 81-82, pp. 483-494.

Hadley, G. (1964). A comparison of order quantities using the average annual costs and the discounted costs, *Management Science*, 10, pp. 472-476.

Huynh, T. T. T. (2006), *Capacity Constraints in Multi-Stage Production-Inventory Systems - Applying Material Requirements Planning Theory*, Production-Economic Research in Linköping, PROFIL 23, Linköping.

Johansen, S. Glud, Thorstenson, A. (1996). Optimal (r, Q) inventory policies with Poisson demands and lost sales: discounted and undiscounted cases, *International Journal of Production Economics,* 46-47, pp. 359-371.

Kim, Y. H., Philippatos, G. C., Chung, K. H. (1986). Evaluating Investment in Inventory Policy: A Net Present Value Framework, *The Engineering Economist*, 31, pp. 119-136.

Klein Haneveld, W. K., Teunter, R. H. (1998). Effects of discounting and demand rate variability on the EOQ, *International Journal of Production Economics*, 54, pp. 173-192.

Leontief, W. W. (1928). Die Wirtschaft als Kreislauf, *Archiv für Sozialwissenschaft und Sozialpolitik,* 60, pp. 577-623.

Leontief, W. W. (1951). *The Structure of the American Economy*, Oxford University Press, London.

Martin, R. L. (1997). Modelling flow of demand through a bill of manufacture, *International Journal of Production Research*, 35, pp. 1911-1924.

Orlicky, J. A. (1975). *Material Requirements Planning,* McGraw-Hill, New York, N.Y.

Segerstedt, A. (1996). A Capacity- Constrained Multi-level Inventory and Production Control Problem, *International Journal of Production Economics*, 45, pp. 449-461.

Silver, E. A., Meal, H. C. (1973). A Heuristic for Selecting Lot Size Quantities for the Case of a Deterministic Time Varying Demand Rate and Discrete Opportunities for Replenishment, *Production and Inventory Management*, 14, pp. 64-74.

Tang, O. (2000). *Planning and Replanning within the Material Requirements Planning Environment – A Transform Approach*, Production-Economic Research in Linköping, PROFIL 16, Linköping.

Teunter, R., van der Laan, E. (2002). On the non-optimality of the average cost approach for inventory models with remanufacturing, *International Journal of Production Economics*, 79, pp. 67-73.

Thorstenson, A. (1988). *Capital Costs in Inventory Models - A Discounted Cash Flow Approach*, Production-Economic Research in Linköping, PROFIL 8, Linköping.

Trippi, R. R., Levin, D. E. (1974). A present value formulation of the classical EOQ problem, *Decision Sciences*, 5, pp. 30-35.

Vazsonyi, A. (1955). The Use of Mathematics in Production and Inventory Control, *Management Science,* 1, pp. 70-85.

Vazsonyi, A. (1958). *Scientific Programming in Business and Industry,* Wiley, New York, N.Y.

Wagner, H. M., Whitin, T. M. (1958). A Dynamic Version of the Economic Lot Size Model, *Management Science*, 5, pp. 89-96.

Zhixin, Y., Grubbström, R. W. (2001). The Origin of the Laplace Transform and Its Applications, *Working Paper WP–283*, Department of Production Economics, Linköping Institute of Technology, June.