Security storage and manufacturing flows subject to random fluctuations

Max-Olivier Hongler

EPFL/Laboratoire GALATEA

15 août 2016

Preliminary consideration

A wide set of multidisciplinary skills are required to conceive and control automatized **flexible manufacturing systems** able to satisfy **flows of highly customized demands**. As for each specific production-consumption network, the numerous and often very complex issues to be addressed differ, one could think, at least at first sight, that trying to construct a general set of lectures devote to such a **virtually infinite number a case-studies** is a vain goal. A similar dilemma is encountered when one has to construct and deliver lectures devoted to non-linear dynamical systems. It indeed exists a single way to be linear (and hence one can construct a set of appropriate lectures) whereas it exist an infinite number of different types of nonlinearities implying that another approach, which encompass sufficient generality, has to be adopted.
Common sense tells that highly skilled production engineers are those able to take appropriate decisions thanks to an optimal balance between **pure theoretical knowledge** (which can taught by a set of of lectures) and a long **practical know-how** (which can only be acquired on the manufacturing shop-floors). The goal of the present lecture notes is to address a few manufacturing **topics which offer sufficient generality and permanence** to allow a formal theoretical approach that can be exposed in lectures. Accordingly, the potential reader is warned that the following set of lecture notes cover pure theoretical knowledge which has therefore definitely to be completed by practical know-how.

1 Optimization issues and random environments

Whatever the production-consumption systems to be considered, random fluctuation will ubiquitously affect any idealized picture we always initially start with. On one hand, production systems are practically always failure prone and simultaneously, customers demands (like the types and quantities of finished products) are barely completely predictable. Adopting a highly stylized view, we can generically say that any production system delivers a randomly fluctuating flow of finished goods to cover a randomly fluctuating flow of customers demands. The unavoidable and ubiquitous presence of randomness raises a couple of basic engineering questions that we shall address, namely :

a) How to enhance the average finished goods flows and how to minimize the variability of complex automatized production networks composed of failure prone echelons ? (i.e. **determine the optimal capacity of buffers in production lines**).

b) How large should be the security stocks required to ensure on-hand delivery to incoming customers, (i.e.**determine the optimal capacity of hedging stocks**).

2 Characterizing the production output of a failure prone machines

Let us consider a production station $\mathcal{M}$ which delivers a single type of finished parts. In ideal operation condition, the time required to complete a single part is the **cycle time** τ_c ; it therefore has the physical dimension of a time. The inverse of the cycle time $U = \tau_c^{-1}$ is the production rate ; it hence has the physical dimension of a frequency. In the sequel, we shall focus on elementary machines for which both τ_c and U are constants. [1] Hence, ideally, the cumulate production $S(T)$ delivered during a time time interval $[0, T]$ exhibits a step structure, as represented in Figure 1, and reads :

[1]Observe that for certain types of production, (like for example fluids in chemistry and/or in food industries), it is possible to pilot the production rate but in the sequel we focus on constant rates.

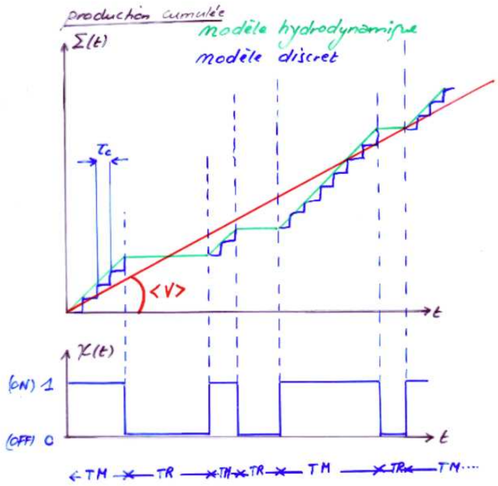

FIG. 1 – Cumulative production for a failure prone production center modeled via an "hydrodynamic" fluid queue picture.

$$S(T) = \sum_{k=1}^{N(T)} \Theta(T - k\tau_c) \qquad \text{with } N(t) := \sup\{k : k\tau_c < T, \} \tag{1}$$

where $\Theta(T - k\tau_c)$ is the Heavyside step function :

$$\Theta(T - k\tau_c) = \begin{cases} 1 & \text{when} \quad T \geq k\tau_c, \\ \\ 0 & \text{when} \quad T > k\tau_c. \end{cases} \tag{2}$$

Generally, the station $\mathcal{M}$ can be interrupted either deterministically (night breaks for example) or randomly due to **unpredictable failures**. Accordingly, the cumulative production during a time horizon T Eq.(1) has to be modified as :

$$S(T) = \sum_{k=1}^{N(T)} \Theta(T - k\tau_c)\, \chi(k\tau_c) \qquad \text{with } N(t) := \sup\{k : k\tau_c < T, \} \tag{3}$$

where the state operating function $\chi(t) = 1$ when $\mathcal{M}$ is operational at time t and conversely $\chi(t) = 0$ when $\mathcal{M}$ is out-of-service at t. In the sequel, we shall focus on situations where $\chi(t)$ models failures. In this case, the operating function $\chi(t)$ will be a **random function which alternates between the**

values 0, (when $\mathcal{M}$ is OFF) **and** 1 (when $\mathcal{M}$ is ON) and the **sojourn time intervals** in the states 0 and 1 are random variables (r.v.) (i.e. time intervals having **random durations**). In the sequel, we shall always assume that successive alternating cycles[2] are statistically independent. The posotive definite r.v. characterizing the ON time duration $t_{\rm ON} \in \mathbb{R}^+$ is drawn from a probability density $f_{\rm ON}(\zeta)$ and similarly then r.v. characterizing the OFF time duration $t_{\rm OFF} \in \mathbb{R}^+$ is drawn from a probability density $f_{\rm OFF}(\zeta)$ and hence, we formally can write :

$$\begin{cases} \text{Prob}\,\{\zeta \ \leq t_{\rm ON} \ \leq (\zeta + d\zeta)\} := f_{\rm ON}(\zeta)d\zeta, \\ \\ \text{Prob}\,\{\zeta \leq t_{\rm OFF} \leq (\zeta + d\zeta)\} := f_{\rm OFF}(\zeta)d\zeta. \end{cases} \tag{4}$$

For future use, let us now introduce the notations for the first moments of these probability densities. We namely have the average[3]

$$\begin{cases} \mathbb{E}\,\{t_{\rm ON}\} := \int_0^\infty \zeta\, f_{\rm ON}\,(\zeta)\, d\zeta := \lambda^{-1}, \\ \\ \mathbb{E}\,\{t_{\rm OFF}\} := \int_0^\infty \zeta f_{\rm OFF}(\zeta) d\zeta := \mu^{-1}, \end{cases} \tag{5}$$

and similarly the associated variances are defined as :

$$\begin{cases} \sigma^2_{\rm ON} := \mathbb{E}\,\{\, t^2_{\rm ON}\,\} - [\mathbb{E}\,\{\, t_{\rm ON}\,\}]^2 = \int_0^\infty \zeta^2\, f_{\rm ON}\,(\zeta)\, d\zeta - \lambda^{-2}, \\ \\ \sigma^2_{\rm ON} := \mathbb{E}\,\{t^2_{\rm OFF}\} - [\mathbb{E}\,\{t_{\rm OFF}\}]^2 = \int_0^\infty \zeta^2 f_{\rm OFF}(\zeta) d\zeta - \mu^{-2}. \end{cases} \tag{6}$$

At this stage, we define the (dimensionless) unavailability factor $\mathcal{J}$ as :

$$\mathcal{J} = \frac{\text{average time to repair}}{\text{average time of operation}} = \frac{1/\mu}{1/\lambda} = \frac{\lambda}{\mu}. \tag{7}$$

We now also introduce the couple of dimensionless (positive definite) coefficients of variations :[4] :

$$\begin{cases} CV^2_{\rm ON} := \sigma^2_{\rm ON}\lambda^2, \\ \\ CV^2_{\rm OFF} := \sigma^2_{\rm OFF}\mu^2. \end{cases} \tag{8}$$

To proceed further, we shall now distinguish between two modeling frameworks defined by two time scales regimes of the $\mathcal{M}$ dynamics :

[2]A cycle is formed by a single ON and OFF alternation.

[3]The operator $\mathbb{E}(Z)$ stands for the **expectation** of the r.v. Z.

[4]The coefficient of variation which is a dimensionless factor, is only defined for strictly positive r.v. The CV^2 factor directly measures the strength of the fluctuations. Observe in particular that a r.v. with $CV^2 = 0$ is effectively an ordinary deterministic quantity.

a) *Discrete queuing dynamics*. This regime occurs when the time intervals τ_c, λ^{-1} and μ^{-1} have identical orders of magnitude. In this case, the discrete nature of the production output cannot be ignored and the appropriate theory to model the production output is the **classical queuing dynamics**. Typical example will be production of cars, engines and most of highly complex devices for which the cycle time is pretty long.

b) *Fluid queueing dynamics*. This regime is characterized by the fact that $\tau_c >> \lambda^{-1}$ and $\tau_c >> \mu^{-1}$. This characterizes high production flows like those encountered in productions of chocolate, cigarettes, screws and many other simple and usually tiny parts typical in micro-engineering shop-floors. For these regimes, the discrete nature of the production flows almost disappears and we may consider that we effectively deal with a fluid of parts for which the appropriate dynamics is known as the **fluid queue dynamics**. In the sequel, we shall focus on this type of production.

3 The fluid queue approach

For high production rate, the production output is assimilable to a fluid flow (like one would view the flow of sand grains in an hour glass). It is interrupted during a random time interval τ_{ON} (i.e when $\mathcal{M}$ is OFF) and then restituted to its nominal value after that $\mathcal{M}$ has been repaired. The random reparation time is τ_{OFF}. In the fluid approximation, without failure, (i.e. $\chi(t) \equiv 1 \; \forall t$ the cumulative production output $S(t)$ is simply an increasing a straight line with slope U. Alternatively, in presence of failure (i.e. $\chi(t)$ randomly alternates between 0 and 1), the cumulative output $S(t)$ delivered by $\mathcal{M}$ is a random piecewise increasing function (i.e. a time-dependent **stochastic process**) as sketched in Figure 2. Let us now characterize the stochastic process $S(t)$. The following proposition can be proved :

Proposition 1.[5]For large production time horizon $T >> \max\{\lambda^{-1}, \mu^{-1}\}$, the random cumulative production $S(T)$ is approximately characterized by

[5]The proof of this proposition can be found in the contribution by Ph. Ciprut, M.-O. Hongler and Y. Salama, *"On the Variance of the Production Output of Transfer Lines"* published in IEEE Trans. on Robotics and Automation **15**, (1999), 33-43. Basically the proof relies on the Central Limit Theorem.

a Normal law $\mathcal{N}(\text{mean}, \text{variance}; \zeta)$ with characteristics :

$$\begin{cases} S(T) = U\int_0^T \chi(s)ds, \\ \text{Prob}\left\{\zeta \le S(T) \le \zeta + d\zeta\right\} \cong \mathcal{N}\left(\frac{UT}{1+\mathcal{J}}, \sigma_S^2(T); \zeta\right) := \frac{1}{\sqrt{2\pi\sigma_S^2 T}} e^{-\frac{[\zeta - \mathbb{E}\{S(T)\}]^2}{2\sigma_S^2(T)}} d\zeta, \\ \mathbb{E}\left\{S(T)\right\} = \frac{U}{1+\mathcal{J}} T, \\ \sigma_S^2(T) = \left[CV_{\text{ON}}^2 + CV_{\text{OFF}}^2\right] \frac{U^2\mathcal{J}}{\mu(1+\mathcal{J})^3} T. \end{cases} \tag{9}$$

An alternative question to be addressed is the characterization of the random time τ_B needed to complete a (relatively large) production batch B. For this quantity we can prove the :

Proposition 2. For production batches[6] $B >> B_{\text{inf}} = U\lambda^{-1}$, the first hitting time τ_B for which $S(\tau_B) = B$ is a r.v. which is approximately characterized by the Inverse Gaussian law $\mathcal{G}(\text{mean}, \text{variance}; \zeta)$:

$$\begin{cases} \text{Prob}\left\{\zeta \le \tau_B \le (\zeta + d\zeta)\right\} \simeq \mathcal{G}(\mathbb{E}\left\{\tau_b\right\}, \sigma_{\tau_B}^2) := \frac{B}{\sqrt{2\pi\sigma_{\tau_B}^2 \zeta^3}} e^{-\frac{\left[B - \zeta\left(\frac{U}{1+\mathcal{J}}\right)\right]^2}{2\sigma_{\tau_B}^2 \zeta}} d\zeta, \\ \mathbb{E}\left\{\tau_b\right\} := \frac{B}{U}(1+\mathcal{J}), \\ \sigma_{\tau_B}^2 := \frac{\left[CV_{\text{ON}}^2 + CV_{\text{OFF}}^2\right] B\mathcal{J}}{\mu U}, \\ CV_{\tau_B}^2 = \frac{\sigma_{\tau_B}^2}{[\mathbb{E}\{\tau_b\}]^2} = \frac{U}{B}\left\{\frac{\left[CV_{\text{ON}}^2 + CV_{\text{OFF}}^2\right]\mathcal{J}}{\mu(1+\mathcal{J})^2}\right\}. \end{cases} \tag{10}$$

Sketch of the proof of Proposition 2. For relatively large B, the output of $\mathcal{M}$ is approximately a drifted Brownian motion (BM) $Y(t)$ characterized by the stochastic differential equation :

[6] For the proposition to be valid, the size of the batch B has to be large enough to approximately allow the stationary production regime to be reached. This can be attained only if numerous ON-OFF cycles have been completed before the production target B to be reached for the first time, (see Figure 2 fro a sketch of the situation). Summarizing, the size of B should be large enough to allow the diffusive approximation of $S(t)$ to be valid.

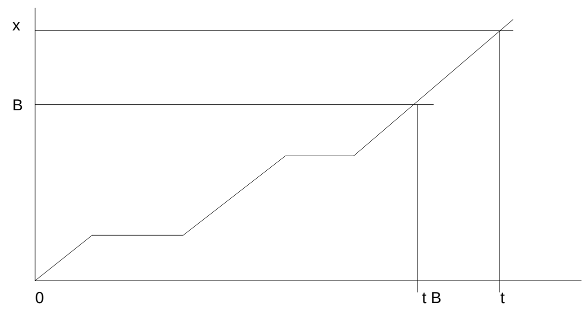

FIG. 2 – Sketch of the first passage time of the process x to the level τ_B.

$$\begin{cases} dY(t) = m\,dt + \sigma_S\,dW_t, \\ m = \frac{U}{1+\mathcal{J}}, \\ \sigma_S = [CV^2_{\mathrm{ON}} + CV^2_{\mathrm{OFF}}]\,\frac{U^2\mathcal{J}}{\mu(1+\mathcal{J})^3}, \end{cases} \tag{11}$$

where dW_t stands for the White Gaussian Noise (WGN) and the other notations are those used in Proposition 1. Now the first hitting time τ_B of the process $Y(t)$ at the level B for the drifted BM Eq.(11) is given by the inverse Gaussian probability law[7] where the mean and the variance are given in Eq.(10).

CQFD

From Eq.(10), we may draw the following remarks :

a) As $CV^2_{\tau_B}$ decreases with the batch size B meaning that the larger B, the less fluctuations (i.e the less uncertainties) on the required time to compete a batch. This matches intuition as for large B, a smoothing effect reduces the relative amplitudes of local randomness.

b) Using the inverse Gaussian law $\mathcal{G}(\mathbb{E}\{\tau_b\}, \sigma^2_{\tau_B}; \zeta)$ law, we may also introduce a risk coefficient $\mathcal{R}(t_{\mathrm{op}}) \in [0, 1]$ defined as the probability that, having produced during the time t_{op}, the batch B is not yet completed. Invoking Eq.(10), we shall have :

[7] V. Seshadri, "*The Inverse Gaussian Law*", Prentice Hall (1994).

$$\text{Prob}\{\text{production batch of size } B \text{ is not completed during time } t_{\text{op}}\} =$$

$$\text{Prob}\{t_{\text{op}} \leq \tau_b \leq \infty\} := \mathcal{R}(t_{\text{op}}) = \int_{t_{\text{op}}}^{\infty} \mathcal{G}(\text{mean}, \text{variance}; \zeta) d\zeta. \tag{12}$$

The quadrature in Eq.(12) required to be computed numerically.

4 The production dipole $\mathcal{D}$

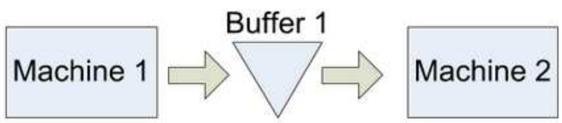

FIG. 3 – Production dipole configuration $\mathcal{D}$.

Let us now focus on the **production dipole** a very common configuration in production lines. In this case, we basically have a couple of failure prone machines say $\mathcal{M}_1$ and $\mathcal{M}_2$ separated by a buffer stock $\mathcal{B}_{12}$ of capacity h items. The machines $\mathcal{M}_1$ and $\mathcal{M}_2$ are failure prone and the introduction of $\mathcal{B}_{12}$ aims at enhancing the **average throughput** $Z(h)$ of the production of the dipole $\mathcal{D}$ when it is a quipped with a buffer capacity h. Here again, we shall focus on high production flows which can be assimilated to fluid flows. Now two complementary scenarii may happen : i) the intermediate buffer $\mathcal{B}_{12}$ is filled up with items and the downstream machine $\mathcal{M}_2$ fails, then **the upstream machine** $\mathcal{M}_1$ **is blocked**, alternatively ii) the intermediate buffer $\mathcal{B}_{12}$ is empty and the upstream machine $\mathcal{M}_1$ fails, then **the downstream machine** $\mathcal{M}_2$ **is starved**. In both situations (i.e. either blocked or starved), the $\mathcal{D}$ throughput $Z(h) = 0$. In the sequel, we may assume that the transfer time of items between $\mathcal{M}_1$ and $\mathcal{M}_2$ is short enough to be negligible in our calculations [8].

[8] If this is not the case, one can consult the contribution by C. Commault and A. Semery, *"Taking into account delays in buffers for analytical performance of transfer lines", IEE Trans.* **22**, *(1990), 133.*

In full generality, only approximative expressions for $Z(h)$ can be derived. However, introducing additional hypotheses on the stochastic behavior of $\mathcal{M}_1$ and $\mathcal{M}_2$ confers a Markovian character to the dynamics which ultimately enables to derive exact expressions for $Z(h)$. Specifically, exact expressions for $Z(h)$ can be obtained provided the r.v. $t_{\rm ON}$ and $t_{\rm OFF}$ are exponentially distributed with respective parameters λ and μ, i.e. formally :

$$\begin{cases} \text{Prob}\{\zeta \leq t_{\rm ON} \leq (\zeta + d\zeta)\} = \lambda e^{-\lambda\zeta} d\zeta & \Rightarrow \quad \mathbb{E}\{t_{\rm ON}\} = \lambda^{-1}, \\ \text{Prob}\{\zeta \leq t_{\rm OFF} \leq (\zeta + d\zeta)\} = \mu e^{-\mu\zeta} d\zeta & \Rightarrow \quad \mathbb{E}\{t_{\rm OFF}\} = \mu^{-1}. \end{cases} \tag{13}$$

Remark. For exponential probability laws as those appearing in Eq.(13), it is easy to verify that the associated coefficient of variations are $CV_{\rm ON}^2 \equiv 1$ and $CV_{\rm OFF}^2 \equiv 1$, (i.e. for exponential laws, the CV^2's are independent of the single parameter characterizing the law).

Assume now that Eq.(13) holds and assume further that $U_1 = U_2 = U$ where U_1 and U_2 are the production rates of $\mathcal{M}_1$ respectively $\mathcal{M}_2$. We shall write λ_1^{-1}, μ_1^{-1} and $\mathcal{J}_1$ respectively λ_2^{-1}, μ_2^{-1} and $\mathcal{J}_2$ for the average time between failure, average reparation time and unavailability of machine $\mathcal{M}_1$ respectively $\mathcal{M}_2$. In our modeling, we shall assume that failures can occur only when machines are in operation. In other words a **starved or a blocked machine cannot fail**. Finally, for future use, we also introduce a couple of dimensionless ratio : $\lambda_2/\lambda_1 := \alpha$ and $\mu_2/\mu_1 := \beta$. Accordingly, the **stationary average throughput** [9] $Z(h)$ of the dipole $\mathcal{D}$ can be expressed as :

$$Z(h) = \frac{U}{1 + \mathcal{J}_{\rm dip}(h)}, \tag{14}$$

where the h-dependent dipole unavailability $\mathcal{J}_{\rm dip}(h)$ reads[10] :

[9] The explicit expressions given by Eqs.(14) and 15) are exclusively valid in the stationary regimes reached when all transient have elapsed. In other words, we first let $\mathcal{D}$ operate during a long (ideally infinite) time and then we monitor the statistical behavior of the buffer content and the average stationary throughput $Z(h)$ of $\mathcal{D}$ delivered at the output of $\mathcal{M}_2$.

[10] The derivation of the following expressions is beyond the scope of the present notes. The interested reader can find the details in two contributions : a) C. Terracol and R. David, *"Performance d'une ligne composée de machine et de stock intermédiaires"*, RAIRO A.P.I.I. **21**(3), (1987), 239-241 and b) D. Dubois and J.-P. Forestier, *"Productivité et encours moyens d'un ensemble de deux machines séparées par une zone de stockage"*, RAIRO Autom. Syst. Analysis and Control **16**, (1981), 105-122.

$$\mathcal{J}_{\rm dip}(h) = \begin{cases} \mathcal{J}_1 \left[1 + \frac{1}{1+\frac{\alpha}{\alpha+1}\mathcal{F}(h)(1+\mathcal{J}_1)}\right] & \text{with} \quad \mathcal{F}(h) := \frac{\mu_1 h}{U} \text{ and when } \frac{U}{1+\mathcal{J}_1} = \frac{U}{1+\mathcal{J}_2}, \\ \\ \mathcal{J}_1 \left[\frac{\frac{\alpha^2}{\beta^2}\exp(\Gamma h)-1}{\frac{\alpha}{\beta}\exp(\Gamma h)-1}\right] & \text{with } \Gamma := \frac{\alpha-\beta}{U}\left[\frac{\mu_1}{1+\alpha} + \frac{\lambda_1}{(1+\beta)}\right] \text{ and when } \mathcal{J}_1 \neq \mathcal{J}_2. \end{cases} \tag{15}$$

Note that the condition $\frac{U}{1+\mathcal{J}_1} = \frac{U}{1+\mathcal{J}_2}$, implies that $\mathcal{J}_1 = \mathcal{J}_2 =$ and therefore $\lambda_2 = \alpha\lambda_1$ and $\mu_2 = \alpha\mu_1$. Now we shall explore the rather rich content delivered by Eq.(15). Hence, using Eq.(15), we can finally simply express the average throughput $Z(h)$ of $\mathcal{D}$ as :

$$Z(h) = \frac{U}{1+\mathcal{J}_{\rm dip}(h)}. \tag{16}$$

4.1 Perfectly balanced production dipole

.

Let us consider the first line in Eq.(15) which, by imposing $\mathcal{J}_1 = \mathcal{J}_2$, describes a perfectly balanced production dipole. Indeed, since we assumed identical nominal production rates $U_1 = U_2 = U$, the unavailability relation $\mathcal{J}_1 = \mathcal{J}_2$ implies that, on average, the production output of $\mathcal{M}_1$ coincides with the production output of $\mathcal{M}_2$. In other words, while the buffer content of $\mathcal{B}_{12}$ fluctuates, there will be not net tendency for $\mathcal{B}_{12}$ to reach either fully filled or fully empty buffer states. Two limiting situations are of direct interest :

i) *Absence of buffer*. In this limiting case, we simply select $h = 0$ in Eq.(15) and obtain :

$$h = 0 \quad \Rightarrow \quad \mathcal{F}(0) = 0 \quad \text{and hence} \quad \mathcal{J}(h=0) = 2\mathcal{J}_1. \tag{17}$$

Therefore, according to Eq.(14), the $\mathcal{D}$-throughput will be given as :

$$Z(0) = \frac{U}{1+2\mathcal{J}_1}. \tag{18}$$

Remark 1. In absence of $\mathcal{B}_{12}$, (i.e. when $h = 0$), both machines $\mathcal{M}_1$ and $\mathcal{M}_2$ are directly coupled. So as soon as $\mathcal{M}_1$ fails, $\mathcal{M}_2$ is immediately starved and conversely, when $\mathcal{M}_2$ fails, $\mathcal{M}_1$ is immediately blocked. So the statistical behaviors of $\mathcal{M}_1$ and $\mathcal{M}_2$ are not totally independent as, by hypothesis, failures appear only when machines are in operation.

And consistently, we observe that if they where independent, we could write :

$$Z(0) = \frac{U}{(1+\mathcal{J}_1)(1+\mathcal{J}_2)} = \frac{U}{(1+\mathcal{J}_1)^2} = \frac{U}{1+2\mathcal{J}_2+\mathcal{J}_2^2} < \frac{U}{1+2\mathcal{J}_1}. \tag{19}$$

Remark 2. The dimensionless parameter $\mathcal{F}(h)$ has an interesting content. To fix the idea, let us choose a buffer capacity leading to $\mathcal{F}(h) = 2$. So we have :

$$\mathcal{F}(h) = 2 \quad \Rightarrow \quad \frac{h}{U} = \frac{2}{\mu} = 2\mathbb{E}\left\{t_{\mathrm{OFF}}\right\}. \tag{20}$$

But now we note that h/U is the time required to empty (or fill) a completely the buffer content when producing at rate U without failure. So we see that $\mathcal{F}(h) = 2$ implies that we select the buffer capacity to ensure that when starting with a half-filled buffer, we can absorb the incoming upstream production of $\mathcal{M}_1$ (conversely feed the downstream production of $\mathcal{M}_2$) during one average reparation time $\mathbb{E}\left\{t_{\mathrm{OFF}}\right\}$.

ii) *Very large buffer content.* In this limiting case, we naturally select $h = \infty$ in Eq.(15) and obtain :

$$\mathcal{J}(h=\infty) \ \Rightarrow \ \mathcal{F}(\infty) = \infty \quad \text{and hence} \quad \mathcal{J}(h=\infty) = \mathcal{J}_1. \tag{21}$$

As intuitively expected, when $h = \infty$, the (stationary) statistical dynamics of $\mathcal{M}_1$ and $\mathcal{M}_2$ is fully decoupled. Accordingly, in the stationary regime, the output of $\mathcal{M}_2$ alone coincides with the throughput of $\mathcal{D}$ itself and we can write :

$$Z(h=\infty) = \frac{U}{1+\mathcal{J}_2} = \frac{U}{1+\mathcal{J}_1}, \tag{22}$$

where the last equality follows by the hypothesis $\mathcal{J}_1 = \mathcal{J}_2$.

From the first line of Eq.(15), one concludes that $\mathcal{J}_{\mathrm{dip}}(h)$ is a monotonously decreasing function of the buffer capacity h (hence of the dimensionless parameter $\mathcal{F}(h)$) and we more precisely we can write :

$$\mathcal{J}_{\mathrm{dip}}(0) = 2\mathcal{J}_1 \geq \mathcal{J}_{\mathrm{dip}}(h) \geq \mathcal{J}_{\mathrm{dip}}(\infty) = \mathcal{J}_1. \tag{23}$$

Accordingly in view of Eqs. (14) and (23), we conclude that the throughput $Z(h)$ **itself is a monotonously increasing function of the buffer**

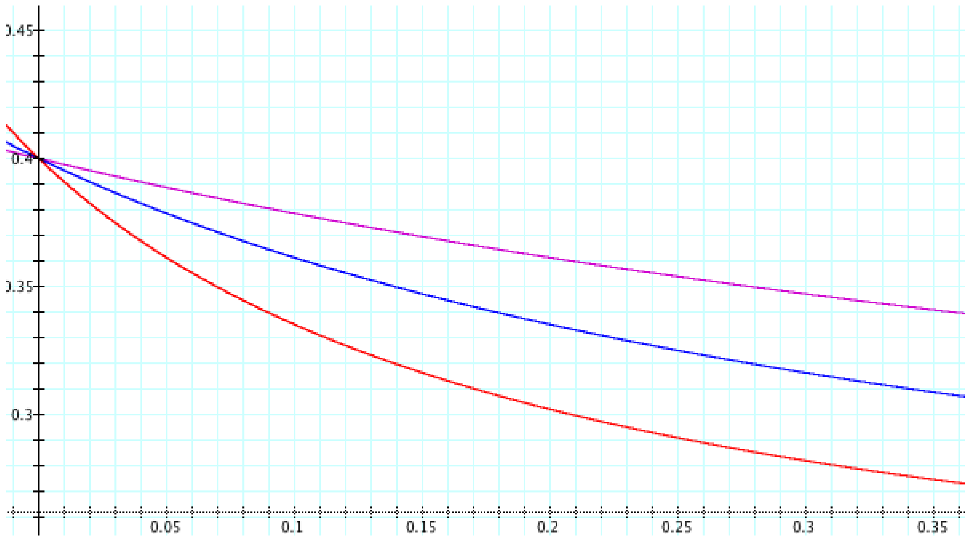

FIG. 4 – Unavailability factor $\mathcal{J}(h)$ (in ordinate) for balanced $\mathcal{D}$ with $\mu = 1$, $\lambda = \mathcal{J} = 0.2$, $\alpha = 1$ (i.e. first line of Eq.(15)) as a function of $\frac{h}{U}$ (in abscissa) for $U = 1$ (purple), $U = 2$ (blue) and $U = 4$ (red).

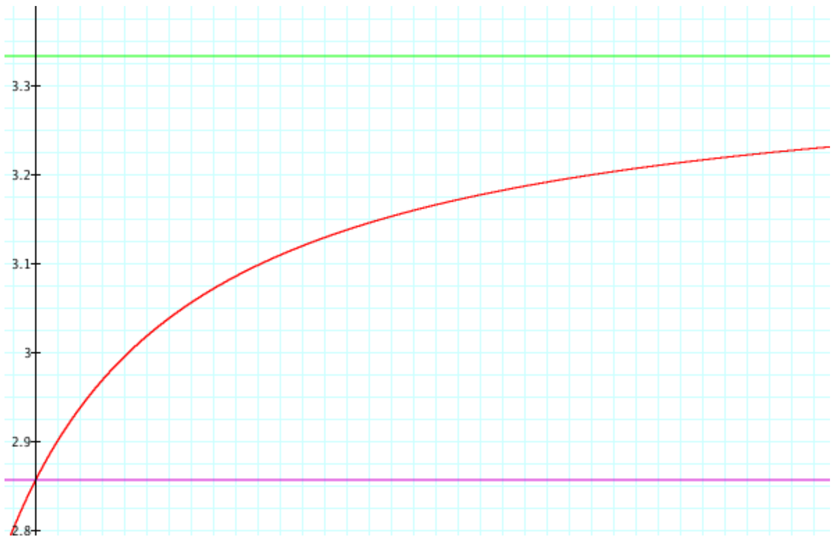

FIG. 5 – Throughput $Z(h)$ (in ordinate) for balanced $\mathcal{D}$ with $\mu = 1$, $\lambda = \mathcal{J} = 0.2$, $\alpha = 1$ (i.e. first line of Eq.(15)) as a function of $\frac{h}{U}$ (in abscissa) for $U = 4$ (red) ; (the green asymptote corresponds to $Z(\infty) = (4/1 + 0.2) = 3.33$ and the purple line stands for $Z(0) = (4/1 + 0.4) = 2.85$).

content h (hence of the dimensionless parameter $\mathcal{F}(h)$) and more precisely we can write :

$$Z(0) = \frac{U}{1+2\mathcal{J}_1} \leq Z(h) = \frac{U}{1+\mathcal{J}_{\text{dip}}(h)} \leq Z(\infty) = \frac{U}{1+2\mathcal{J}_1}, \tag{24}$$

where $\mathcal{J}_{\text{dip}}(h)$ is given by the first line of Eq.(15). Accordingly, **for balanced production dipoles $\mathcal{D}$, Eq.(24) enables to straightforwardly estimate the throughput increase that can be realized by installing a buffer between the couple of failure prone machines $\mathcal{M}_1$ and $\mathcal{M}_2$.**

Remark 3. It is also worth to mention that the efficiency factor $e := [1/1+\mathcal{J}_{\text{dip}}(h)] \in [0,1]$ effectively represents the average fraction of time that $\mathcal{M}_2$ is non-operating. Non-operation occurs when $\mathcal{M}_2$ is either failed or starved.

4.2 Unbalanced dipoles

Let us now analyze the unbalanced production dipole represented by the second line in Eq.(15) [11]. First we remark that $\mathcal{D}$ is indeed unbalanced since we have $U_1 = U_2 = U$ but $\mathcal{J}_1 \neq \mathcal{J}_2$ and therefore :

$$Z_1 := \frac{U_1}{1+\mathcal{J}_1} = \frac{U}{1+\mathcal{J}_1} \neq Z_2 := \frac{U_2}{1+\mathcal{J}_2} = \frac{U}{1+\mathcal{J}_2}. \tag{25}$$

Let us focus first on the case $\mathcal{J}_1 > \mathcal{J}_2$ which implies straightforwardly that :

$$\lambda_1/\lambda_2 = \alpha^{-1} > \beta^{-1} = \mu_1/\mu_2 \quad \Rightarrow \quad \alpha < \beta. \tag{26}$$

In this case, the average output of $\mathcal{M}_1$ is smaller than the the average output of $\mathcal{M}_2$ implying that $\mathcal{B}_{12}$ exhibits a net tendency to be empty (and hence $\mathcal{M}_2$ will, have a net tendency to be starved). Let us again analyze two limiting situations. As according to Eq. (26) $\alpha < \beta$, we have $\Gamma < 0$ in the second line of Eq.(15) so we shall have :

i) *Absence of buffer*. With $h = 0$, the second line of Eq.(15) yields :

[11] In high flows production engineering, we shall very barely encounter unbalanced $\mathcal{D}$ as they immediately lead to bottlenecks. However, other similar situations like electricity power storage in safety batteries are precisely of the unbalanced type. Indeed, we would systematically try to guarantee non-empty charge in a battery feeding a random electricity demand. Therefore, we would try to have an unbabalnaced electricity flow with a net tendency to stick to the filled boundary of the buffer

$$\mathcal{J}_{\text{dip}}(h=0) = \frac{\frac{\alpha^2}{\beta^2}-1}{\frac{\alpha}{\beta}-1}\mathcal{J}_1 = (1+\frac{\alpha}{\beta})\mathcal{J}_1 = \mathcal{J}_1 + \mathcal{J}_2 \tag{27}$$

and again we consistently observe that in absence of buffer, (i.e. when $h = 0$), the total $\mathcal{D}$-unavailability is the sum of the individual unavailabilities.

ii) *Very large buffer content.* For $h \to \infty$ when $\Gamma < 0$, Eq.(15) immediately implies :

$$\mathcal{J}_{\text{dip}}(\infty) = \mathcal{J}_1, \tag{28}$$

meaning that the less performant machine (in this case $\mathcal{M}_1$) "slaves" the production throughput [12].

Again, the second line of Eq.(15) implies that $\mathcal{J}_{\text{dip}}(h)$ is a monotonously decreasing function of the buffer maximal capacity h, (or the dimensionless factor $\mathcal{F}(h)$) and hence, $Z(h)$ will be an monotonously increasing function of h. In this case we shall write :

$$\mathcal{J}_{\text{dip}}(0) = \mathcal{J}_1 + \mathcal{J}_2 \geq \mathcal{J}_{\text{dip}}(h) \geq \mathcal{J}_{\text{dip}}(\infty) = \max\{\mathcal{J}_1, \mathcal{J}_2\}\,. \tag{29}$$

Accordingly, in view of Eqs. (14) and (23), the stationary average throughput $Z(h)$ can be written as :

$$Z(0) = \frac{U}{1+\mathcal{J}_1+\mathcal{J}_2} \leq Z(h) = \frac{U}{1+\mathcal{J}_{\text{dip}}(h)} \leq Z(\infty) = \frac{U}{1+\max\{\mathcal{J}_1,\mathcal{J}_2\}}, \tag{30}$$

4.3 How to relax our (over)simplifying hypotheses

To derive the quite elegant compact form Eq. (14), basically two simplifying hypotheses have been made : i) $U_1 = U_2$ and the r.v.'s t_{ON} and t_{OFF} are assumed to be exponential distributions for both machines $\mathcal{M}_1$ and $\mathcal{M}_2$, (see Eq.(13)). In actual contexts, this might not be realized and one should hence try to obtain approximations and information valid under relaxed.

[12] Observe conversely, when $\alpha > \beta$ implying $\Gamma > 0$, Eq.(15) also implies the fully consistent behavior, namely : $\mathcal{J}_{\text{dip}}(\infty) = \frac{\alpha}{\beta}\mathcal{J}_1 = \mathcal{J}_2$. Summarizing, we can write that $\mathcal{J}_{\text{dip}}(\infty) = \max\{\mathcal{J}_1, \mathcal{J}_2\}$

4.3.1 Unbalance production rate $U_1 \neq U_2$

In this case, we can approximate the dynamics by re-normalizing the parameters λ's and μ's to effectively correct the production rate balance : Specifically, we shall write :

$$\frac{U_2}{1+\mathcal{J}_2} = \frac{U_1}{1+\mathcal{J}_{2,\mathrm{eff}}} \quad \Rightarrow \quad \mathcal{J}_{2,eff} := \frac{\lambda_{\mathrm{eff}}}{\mu_{\mathrm{eff}}} = \frac{(1+\mathcal{J}_2)U_1}{U_2}. \tag{31}$$

Using Eq.(31), we may now directly use Eq.(15) with $\alpha = \frac{\mu_{\mathrm{eff}}}{\mu_1} = \frac{\lambda_{\mathrm{eff}}}{\lambda_1}$. Choosing arbitrarily $\mu_{\mathrm{eff}} = \mu_2$, using Eq.(31), we end with [13] :

$$\lambda_{\mathrm{eff}} = \mu_2 \frac{(1+\mathcal{J}_2)U_1}{U_2}. \tag{32}$$

As to re-establish the balance, via the λ_{eff} factor, we effectively introduce, extra fluctuations in the flow. Hence, for the approximate throughput $Z_{\mathrm{eff}}(h)$ obtain by using $\mathcal{J}_{\mathrm{eff}}$ in Eq.(31), we shall have $Z_{\mathrm{eff}}(h) < Z_{\mathrm{exact}}(h)$, where $Z_{\mathrm{exact}}(h)$ stands for the (unknown) exact average throughput value for $\mathcal{D}$ with $U_1 \neq U_2$.

4.3.2 Arbitrary probability distributions for t_{ON} and t_{OFF}

For arbitrary distributions of t_{ON} and t_{OFF}, the dynamics becomes non-Markovian and this makes all analytical calculations more tedious, (and eventually impossible!). However, we safely express the general qualitative principle : *"the more fluctuations the less average throughput"*. Based on this, we may conclude that whenever the coefficient of variations $CV^2_{\mathrm{ON}} > 1 (< 1)$ and $CV^2_{\mathrm{OFF}} > 1 (< 1)$, the exact throughput $Z(h)$ given by Eq.(15) is derived for the situations with $CV^2 = 1$. Hence Eq.(15) will actually deliver an optimistic, (pessimistic) approximation.

Additional remarks. So far we were only interested in calculating the average throughput $Z(h)$ of $\mathcal{D}$. This is and average quantity. Obviously the throughput itself is also a r.v. and another very relevant performance parameter will be given by the variance (equivalently the standard deviation) of the throughput. This question is directly addressed in [14]

[13] With this choice, we actually adapt the balance of the production rate by a modification of the average of $t_{\mathrm{ON}}\lambda_2 \mapsto \lambda_{\mathrm{eff}}$ of $\mathcal{M}_2$. One could alternatively use the modification $t_{\mathrm{OFF}} \mapsto \mu_{\mathrm{eff}}$ and keep λ_2 unchanged.

[14] Ph. Ciprut, M.-O. Hongler and Y. Salama, *"On the Variance of the Production Output of Transfer Lines"* published in IEEE Trans. on Robotics and Automation **15**, (1999), 33-43.

5 More complex production networks - Aggregation methodology

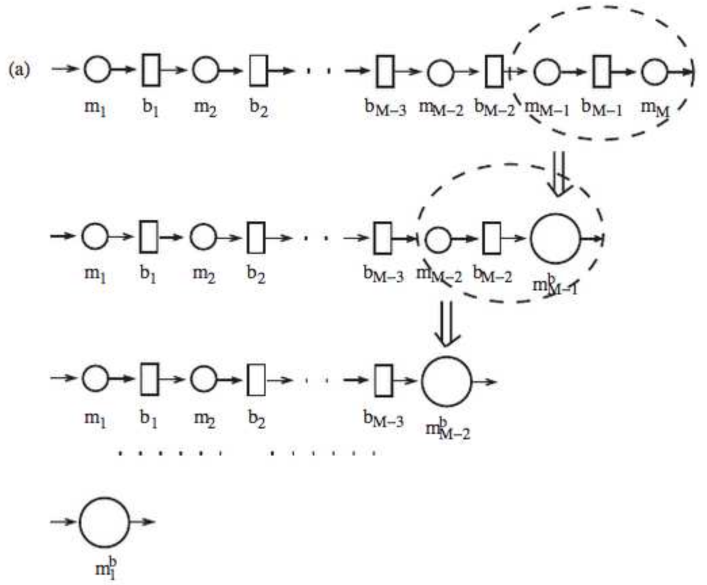

FIG. 6 – Basic principle of the aggregation method of a production line formed by machines m_k separated by buffers b_k. Successive aggregation steps lead ultimately to a single effective machine here denoted m_1^b.

The production dipole configuration $\mathcal{D}$ can be consider as the cornerstone of more complex production networks. In this section, we shall focus on serial production lines characterized by a serial succession of N machines $\mathcal{M}_j$, $j = 1, 2, \cdots, N$ separated by $M < N$ buffers of different capacities h_k, $k = 1, 2, \cdots, M$. By a set of successive aggregations in which we assimilate a $\mathcal{D}$ to a single effective aggregated $\mathcal{M}_{\text{ag}} \equiv \mathcal{D}$, we can ultimately **aggregate** the whole production line to end with a global effective machine. This approach obviously introduces errors. Indeed, the r.v. $t_{\text{ON,ag}}$ and $t_{\text{OFF,ag}}$ entering into the stochastic dynamics of $\mathcal{M}_{\text{ag}}$ are not exponentially distributed. The use of Eq.(15) to calculate the average throughput $Z_{\text{ag}}(h)$ resulting from the aggregation of dipoles formed by a couple of $\mathcal{M}_{\text{ag}}$ separated by a buffer, introduces therefore an error [15].

To show how the aggregation method operates, let us focus on a simple

[15]In practice, it is worth to mention that, fortunately, this error remains usually relatively small.

illustration in which we have a succession of four identical machines $U_k = U$, $\lambda_k = \lambda$ and $\mu_k = \mu$ for $k = 1, 2, 3, 4$ separated by three buffer with maximal capacities h_1, h_2 and h_3. The total buffer capacity of the line will be denoted by $H := h_1 + h_2 + h_3$. We may now address the question :

Assuming that we can actually dispose of a global buffer H, how to optimally distribute H between the four $\mathcal{M}_k$'s ?

A simple way to address the question is to calculate the average line throughput $\mathcal{Z}_{\mathcal{L}}$ for two extreme configurations, namely :

i) *Central buffer configuration.*

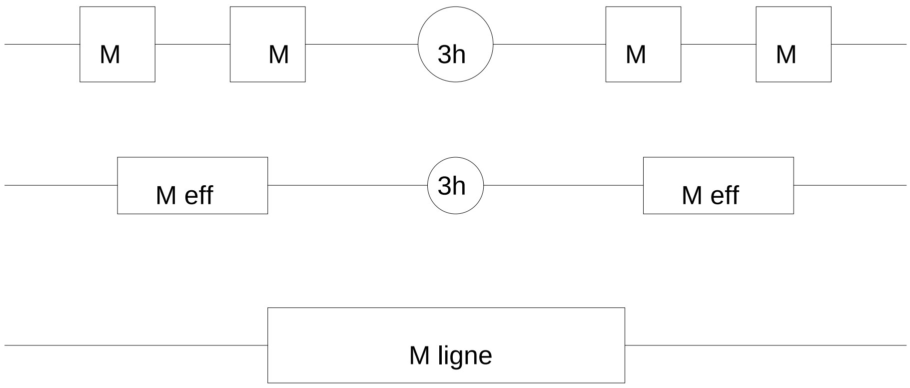

FIG. 7 – Central buffer configuration with $H = 3h$.

This situation is ketched in Figure 7. We aggregate first $\mathcal{M}_1$ with $\mathcal{M}_2$ to form a single machine $\mathcal{M}_{\text{ag},1}$ and similarly, we aggregate $\mathcal{M}_3$ with $\mathcal{M}_4$ to form a single machine $\mathcal{M}_{\text{ag},2}$. Then we consider the production dipole formed by $\mathcal{M}_{\text{ag},1}$ and $\mathcal{M}_{\text{ag},2}$ separated by a buffer of maximal capacity H. Using the result Eq.(17), we immediately have that $\mathcal{M}_{\text{ag},1}$ and $\mathcal{M}_{\text{ag},2}$ are identical machines with the common unavailability $\mathcal{J}_{\text{ag}} = 2\frac{\lambda}{\mu}$. Finally, the first line of Eq.(15) enables to derive the average throughput $Z_{\mathcal{L}}(H)$ as[16] :

$$\begin{cases} \mathcal{J}_{\mathcal{L},center}(H) = 2\mathcal{J}\left\{1 + \frac{1}{1+\frac{\mu H}{2U}(1+2\mathcal{J})}\right\}, \\ \\ Z_{\mathcal{L},center}(H) = \frac{U}{1+\mathcal{J}_{\mathcal{L},center}(H)}. \end{cases} \tag{33}$$

Observe from Eq.(33) that when $H = 0$ we consistently find $Z\mathcal{L} = 4\mathcal{J}$.

[16]For this configuration, the buffer maximal content is simply H and since all machines are identical, we simply have $\alpha = 1$. Hence from Eq.(15), we conclude that $\mathcal{F}(H) = \frac{\mu H}{2U}$.

ii) *Evenly distributed buffers configuration.*

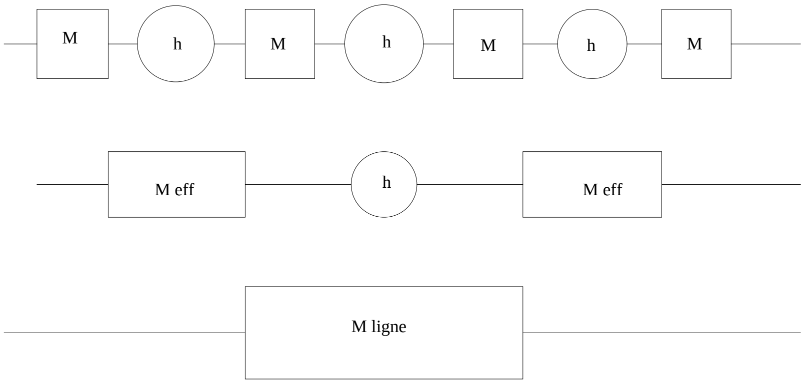

FIG. 8 – Evenly distributed buffer configuration with a total buffer $H = 3h$.

This configuration is sketched in Figure 8. We assimilate the dipole formed by $\mathcal{M}_1$ with $\mathcal{M}_2$ separated by a buffer with maximal capacity $\frac{1}{3}H$ and similarly for $\mathcal{M}_3$ with $\mathcal{M}_4$ also separated by a buffer with capacity $\mathcal{M}_3$ with $\mathcal{M}_4$. For both cases, using the first line of Eq.(15) gives :

$$\mathcal{J}_{\rm ag}(\frac{H}{3}) = \mathcal{J}\left\{1 + \frac{1}{1 + \frac{\mu H}{6U}(1+\mathcal{J})}\right\}. \qquad (34)$$

Finally, to calculate the global unavailability $\mathcal{J}_{\mathcal{L},dist}(H)$, we aggregate once more to end with [17] :

$$\left\{\begin{array}{l} \mathcal{J}_{\mathcal{L},dist}(H) = \mathcal{J}_{\rm ag}(\frac{H}{3})\left\{1 + \frac{1}{1+\frac{\mu H}{6U}\left[1+\mathcal{J}_{\rm ag}(\frac{H}{3})\right]}\right\}, \\ \\ Z_{\mathcal{L},dist}(H) = \frac{U}{1+\mathcal{J}_{\mathcal{L},dist}(H)}. \end{array}\right. \qquad (35)$$

Now, we have to compare the resulting average throughputs $Z_{\mathcal{L},center}(H)$ and $Z_{\mathcal{L},dist}(H)$. To do that analytically, let us consider two extremes situations :

a) *Very limited global capacity content,* $H \gtrsim 0$.
Here, we Taylor expand Eqs.(33) and (35) to first order in H to allow the direct comparison. We obtain :

[17] After one aggregation, we again get a dipole with identical machines $\mathcal{M}_{\rm ag}$ (thus again implying that $\alpha = 1$ in Eq.(15)) separated by a buffer of capacity $\frac{H}{3}$.

$$\begin{cases} \mathcal{J}_{\mathcal{L},center}(H) \simeq 4\mathcal{J}\left(1 - \frac{\mu H}{4U}\right) + \mathcal{O}(H^2), \\ \\ \mathcal{J}_{\mathcal{L},dist}(H) \simeq 4\mathcal{J}\left(1 - \frac{\mu H}{6U}\right) + \mathcal{O}(H^2)\right). \end{cases} \tag{36}$$

From Eq.(36), we observe that for small H, we have $\mathcal{J}_{\mathcal{L},center}(H) < \mathcal{J}_{\mathcal{L},dist}(H)$. This directly implies that $Z_{\mathcal{L},dist}(H) < Z_{\mathcal{L},center}(H)$ and therefore, we conclude that **for very limited H, it is more advantageous to adopt a central configuration**.

b) *Very large global capacity content, $H \rightarrow \infty$.*
From Eq.(33), we see that $\lim_{H\rightarrow\infty} \mathcal{J}_{\mathcal{L},center}(H) = 2\mathcal{J}$ and conversely, from Eq.(35), we have $\lim_{H\rightarrow\infty} \mathcal{J}_{\mathcal{L},dist}(H) = \mathcal{J}$. Hence conclude that in this limit $Z_{\mathcal{L},dist}(H) > Z_{\mathcal{L},center}(H)$ and therefore, contrary that for the $H \gtrsim 0$ case, we conclude that **for very large buffers capacities it becomes more advantageous to adopt a distributed configuration**.

Remark. There are obviously several combinatorial possibilities to aggregate a single production line. The best results (i.e the smallest approximation errors) are obtaind by aggregating by first the dipoles with the smaller buffers and then by increasing buffer capacities.

6 Optimal hedging stocks and Just-In-Time production policy

As exposed in the previous section, we consider that production systems are intrinsically failure prone ant therefore their output flows exhibit a random component. On the other hand finish goods demand are also often fluctuating. In competitive markets, it is imperious to match the demand immediately and for this one has to introduce an inventory of finished goods $\mathcal{H}$, also called a hedging stock. (HS).The existence of a HS helps the management to serve the customers on the spot. Obviously HS incurs additional costs which are monotonously increasing with the available capacity. In the sequel, we shall denote by $c^+(y)$ the storage cost incurred when y items are present in the HS. Conversely, $c^-(y)$ will denote the cost incurred when y items are backlogged. As in the previous sections, we shall again assume that production flows delivered by a machine $\mathcal{M}$ can be approximated by an "hydrodynamic" picture and similarly to Eq.(11). Contrary to the section 1, we now assume that we have the freedom to **control the production rate $\mathcal{U}(t)$** and we shall write the $\mathcal{U}(t)$**-dependent production flow dynamics** as :

$$\begin{cases} \frac{dY_{\mathcal{U}}(t)}{dt} = \mathcal{U}(t)\chi(t) - D(t), \\ \\ Y_0 = 0. \end{cases} \tag{37}$$

where $\mathcal{U}(t)$ now stands for a controllable production rate, (in section 1, $\mathcal{U}(t) \equiv U$ was assumed to be a constant), as in section 1, $\chi(t)$ describes the operating state of $\mathcal{M}$ and $D(t)$ is the demand rate. In Eq.(37), the stochastic nature of the flows enters via $\chi(t)$ which models the failure prone character of $\mathcal{M}$ and via $D(t)$ which can be a random process. In this rest of this section, we will address two questions :

a) *What is the optimal production policy* $\mathcal{U}^*(t)$ *which ensures minimal hedging/backlog costs ?*

b) *What is the optimal capacity of* $\mathcal{H}$ *? Does it exists conditions under which no hedging stock is optimal, (in this last case we shall speak of Just-In-Time (JIT) production policy) ?*

To answer the previous questions with fully general dynamics Eq.(37) is a very ambitious task. To become very explicit, we shall now simplify the dynamics Eq.(37) by focusing on situations where $D(t) \equiv d$ (i.e. constant demand rate) and $\chi(t)$ with the characteristics defined by Eq.(13). The controllable production rate $U^*(t)$ is obviously limited in a range $\mathbb{U} := [0, U_{\max}]$ and we assume that the production objective is feasible for the demand rate d, namely :

$$\frac{U_{\max}}{1+\mathcal{J}} > d, \qquad (\mathcal{J} := \lambda/\mu), \tag{38}$$

meaning that at full production rate $\mathcal{M}$ is indeed able to satisfy the (constant) demand rate d. For production policy U, the average hedging costs $J(U)$ on an infinite operation time horizon, can be written as :

$$\begin{cases} J(\mathcal{U}) := \lim_{T\to\infty} \mathbb{E}\left\{\int_0^T \left[c^+ Y_{\mathcal{U}}^+(s) + c^- Y_{\mathcal{U}}^-(s)\right] ds\right\}, \\ \\ \frac{dY_{\mathcal{U}}(t)}{dt} = \mathcal{U}(t)\chi(t) - d, \end{cases} \tag{39}$$

where we have used the notation : $Y_{\mathcal{U}}^+ := \max\{0, Y_{\mathcal{U}}\}$ and $Y_{\mathcal{U}}^- := \max\{0, -Y_{\mathcal{U}}\}$. The mathematical formulation of the optimal control problem can be summarized as follows : *Find an admissible production policy* $U^*(t) \in \mathbb{U}$ *such that* :

$$\min_{\mathcal{U}\in\mathbb{U}} J(\mathcal{U}) = J(\mathcal{U}^*). \tag{40}$$

The formulation of Eq.(40) as a dynamic programming problem and its full solution have been first proposed by T. Bielecki and P. R. Kumar[18] and the results can be summarized as follows ; (see also Fir 9) :

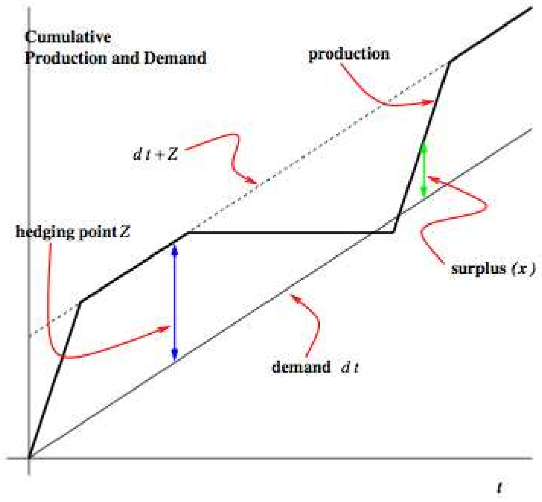

FIG. 9 – Dynamic of the HS content under the optimal policy $\mathcal{U}^*$. The cumulative production is follows the cumulative demand $d \cdot t$ when the optimal hedging level Z is attained. This picture is directly taken from S. B. Gershwin, *"Manufacturing System Engineering"*. Prentice Hall.

$$\mathcal{U}^*(t) = \begin{cases} 0 & \text{if} \quad Y_{\mathcal{U}^*} > Z \;\text{ and }\; \chi(t) = 1, \\ d & \text{if} \quad Y_{\mathcal{U}^*} = Z, \;\text{ and }\; \chi(t) = 1 \\ U_{\max} & \text{if} \quad Y_{\mathcal{U}^*} < Z \;\text{ and }\; \chi(t) = 1, \end{cases} \tag{41}$$

where Z stands for the optimal HS. It reads :

$$Z = \max\left\{0, \frac{1}{A}\log(B)\right\}, \tag{42}$$

[18]Under the present hypothesis, the full mathematical details of this section can be read in T. Bielecki and P. R. Kumar, *Optimality of zero-inventory policies for unreliable manufacturing systems*. Operations Research **36**(4), (1988), 532-541.

with the definitions :

$$\begin{cases} A = \frac{\mu}{d} - \frac{\lambda}{(U_{\max}-d)}, \qquad (\text{Eq.}(38) \Rightarrow A > 0)\,, \\ \\ B = \frac{U_{\max}\lambda\left[c^{+} + c^{-}\right]}{c^{+}(U_{\max}-d)(\lambda+\mu)}. \end{cases} \tag{43}$$

From Eqs.(42) and (43), we immediately conclude that the condition for the Just-In-Time production policy reads as $B \leq 1$. We sees that this last condition explicitly depends on the 6 parameters $(\lambda, \mu, U_{\max}, D, c^{+}, c^{-})$ which enter into the optimal control problem.

7 List of symbols

τ_c : cycle time (sec/item),
U : production rate (item/sec) ; $(U = \tau_c^{-1})$,
T : production time horizon (sec),
$S(T)$: cumulative production (item),
$\Theta(\cdot)$: Heaviside step function ; (definition given in Eq.(2)),
$\chi(t)$: operation state of the machine $\mathcal{M}$; $(\chi(t) \in \{\{0\}, \{1\}\})$,
t_{ON} : r.v. characterizing the length of the ON operation state of $\mathcal{M}$,
$f_{\mathrm{ON}}(t)$: probability density characterizing the r.v. t_{ON},
t_{OFF} : r.v. characterizing the length of the OFF operation state of $\mathcal{M}$,
$f_{\mathrm{OFF}}(t)$: probability density characterizing the r.v. t_{OFF},
$\mathbb{E}\{X\}$: average of the r.v. X,
λ : is the average rate of failure (1/sec),
$\lambda^{-1} := \mathbb{E}\{t_{\mathrm{ON}}\}$: average of t_{ON} (sec),
μ : is the average rate of reparation (1/sec),
$\mu^{-1} := \mathbb{E}\{t_{\mathrm{OFF}}\}$: average t_{OFF} (sec),
σ^2_{ON} : variance of t_{ON} (sec^2),
σ^2_{OFF} : variance of t_{OFF} (sec^2),
$\mathcal{J}$: the unavailability factor ; $(\mathcal{J} = (\lambda/\mu) = 0 \;\Rightarrow \mathcal{M}$ never fails),
CV^2_{ON} : the coefficient of variation of the ON r.v. operation time,
CV^2_{OFF} : the coefficient of variation of the OFF r.v. reparation time,
B : batch size (item),
τ_b : random time needed to complete a production batch B (sec),
$\mathcal{N}(\mathrm{mean}, \mathrm{var}; \zeta)d\zeta := \frac{1}{\sqrt{2\pi\,\mathrm{var}}} \exp\left\{-\frac{(\zeta-\mathrm{mean})^2}{2\,\mathrm{var}}\right\} d\zeta, \qquad \zeta \in \mathbb{R}$,
$\mathcal{G}(\mathrm{mean}, \mathrm{var}; \zeta)d\zeta := \frac{(\mathrm{mean})^3}{\sqrt{2\pi\,(\mathrm{var})^2\zeta^3}} \exp\left\{-\frac{(\zeta-\mathrm{mean})^2}{2\,(\mathrm{mean})^2\zeta}\right\} d\zeta, \qquad \zeta \in \mathbb{R}^+$,
$\mathcal{D}$: production dipole,
$\mathcal{B}_{12}$: buffer separating two failure prone machines $\mathcal{M}_1$ and $\mathcal{M}_2$,
λ_1^{-1} : average time between failure for $\mathcal{M}_1$ (sec),
μ_1^{-1} : average reparation time for $\mathcal{M}_1$ (sec),
λ_2^{-1} : average time between failure for $\mathcal{M}_2$ (sec),
μ_2^{-1} : average reparation time for $\mathcal{M}_2$ (sec),
$\mathcal{J}_1$: the unavailability factor for $\mathcal{M}_1$; $(\mathcal{J}_1 = (\lambda_1/\mu_1)$,
$\mathcal{J}_2$: the unavailability factor for $\mathcal{M}_2$; $(\mathcal{J}_2 = (\lambda_2/\mu_2)$,
$\alpha := \lambda_2/\lambda_1$,
$\beta := \mu_2/\mu_1$,
h : capacity of $\mathcal{B}_{12}$ between the failure prone $\mathcal{M}_1$ and $\mathcal{M}_2$ (item),
$Z(h)$: throughput of $\mathcal{D}$ equipped with $\mathcal{B}_{12}$ of capacity h (item/sec,)
$\mathcal{J}_{\mathrm{dip}}(h)$: h-dependent dipole unavailability,
$\mathcal{U}(t)$: controllable production rate (item/sec),

$c^+(y)$ storage cost of y items in the hedging stock (dollars/item$\cdot$ sec),
$c^-(y)$ backlog cost for y items (dollars/item$\cdot$ sec),
$D(t)$: demand rate for finished items (item/sec),
d : constant demand rate
$Y_{\mathcal{U}}^+ := \max\{0, Y_{\mathcal{U}}\}$
$Y_{\mathcal{U}}^- := \max\{0, -Y_{\mathcal{U}}\}$
Z optimal hedgins stocks.