Supplementary material for “Optimization of power consumption and user impact based on point process modeling of the request sequence”

Jean-Baptiste Durand, Stéphane Girard, Victor Ciriza and Laurent Donini

J.-B. Durand (corresponding author) and S. Girard are with Team Mistis, INRIA Rhône-Alpes and LJK, 655 avenue de l'Europe Montbonnot, 38334 Saint-Ismier Cedex, France (Jean-Baptiste.Durand@inrialpes.fr)

V. Ciriza and L. Donini are with Xerox Research Centre Europe, 6 chemin de Maupertuis, 38240 Meylan, France.

1. Mathematical results

LEMMA 1. *The expected consumption with one sleep mode between two successive printings given $X_{1:i-1}$ is:*

$$\begin{aligned}\mathbb{E}(h(X_i,\tau_i)|X_{1:i-1}) &= a\mathbb{E}(X_i|X_{1:i-1}) + (a-b)\bar{F}_i(\tau_i)(\Delta t + \tau_i) \\ &- (a-b)\int_{\tau_i}^{+\infty} x f_i(x)dx.\end{aligned}$$

LEMMA 2. *The expected consumption with multiple sleep modes between two successive print requests given $X_{1:i-1}$ is:*

$$\begin{aligned}&\mathbb{E}(h(X_i,\tau_i^{(1)},\ldots,\tau_i^{(m)})|X_{1:i-1}) = a\mathbb{E}(X_i|X_{1:i-1}) \\ &+ \sum_{j=1}^{m}(b_{j-1}-b_j)\bar{F}_i(\tau_i^{(j)})(\Delta t_j + \tau_i^{(j)}) \\ &- \sum_{j=1}^{m}(b_{j-1}-b_j)\int_{\tau_i^{(j)}}^{+\infty} x f_i(x)dx.\end{aligned}$$

It is remarkable that the expected energy consumption is expanded as the sum of m terms, each of them depending on one and only one timeout. Thus, the minimization of $\mathbb{E}(h(X_i,\tau_i^{(1)},\ldots,\tau_i^{(m)})|X_{1:i-1})$ with respect to $(\tau_i^{(1)},\ldots,\tau_i^{(m)})$ can be split into m optimization problems leading to explicit optimal timeouts.

PROPOSITION 1. *Two situations are examined, depending on the behavior of the printing rate function.*
a) Suppose that the printing rate function $z_i(x)$ is decreasing in x. For each $j = 1, \dots, m$ three cases occur:

- *If $1/\Delta t_j < \ell_i$, then $\hat{\tau}_i^{(j)} = +\infty$.*
- *If $\ell_i \leq 1/\Delta t_j \leq z_i(0)$, then $\hat{\tau}_i^{(j)}$ is the unique solution of $z_i(\hat{\tau}_i^{(j)}) = 1/\Delta t_j$.*
- *If $z_i(0) < 1/\Delta t_j$, then $\hat{\tau}_i^{(j)} = 0$.*

b) Suppose that z_i is increasing or constant. For each $j = 1, \dots, m$ four cases occur:

- *If $1/\Delta t_j < z_i(0)$, then $\hat{\tau}_i^{(j)} = +\infty$.*
- *If $z_i(0) \leq 1/\Delta t_j \leq \min(\ell_i, 1/\mathbb{E}(X_i|X_{1:i-1}))$, then $\hat{\tau}_i^{(j)} = +\infty$.*
- *If $\max(z_i(0), 1/\mathbb{E}(X_i|X_{1:i-1})) < 1/\Delta t_j \leq \ell_i$, then $\hat{\tau}_i^{(j)} = 0$.*
- *If $\ell_i < 1/\Delta t_j$, then $\hat{\tau}_j^{(1)} = 0$.*

LEMMA 3. *The expected consumption including user impact is*

$$\begin{aligned}\mathbb{E}(g(X_i, \tau_i)|X_{1:i-1}) &= a\mathbb{E}(X_i|X_{1:i-1}) + (a-b)\bar{F}_i(\tau_i)(\tilde{\Delta}t_+\tau_i) \\ &- (a-b)\int_{\tau_i}^{+\infty} x f_i(x)dx\end{aligned}$$

with $\tilde{\Delta}t = (c + d + \delta)/(a - b)$.

2. M step of the EM algorithm for Weibull HMCs.

This paragraph describes the M step of EM algorithm, dedicated to parameter re-estimation in HMCs, in the case of Weibull emission distributions.
Generally in HMCs, the re-estimation procedure for the π_k and $A_{k,l}$ parameters is not specific to the family of emission distributions. In particular, the usual formulae (6.14) and (6.15) in Ephraim and Merhav (2002) hold for Weibull emission distributions. For all $k = 1, .., K$ the new values of parameters $(\lambda_k^{(m+1)}, \alpha_k^{(m+1)})$ after m iterations of the EM algorithm cancel the partial derivatives of the $\mathbb{Q}$ function (formula (6.13) in Ephraim and Merhav, 2002), and thus satisfy the system:

$$\left\{\begin{array}{l}\sum_t \mathbb{P}_{\eta^{(m)}}(S_t = k|X_1^n = x_1^n)\dfrac{\partial}{\partial \lambda_k}\log f_{\lambda_k,\alpha_k}(x) = 0 \\ \sum_t \mathbb{P}_{\eta^{(m)}}(S_t = k|X_1^n = x_1^n)\dfrac{\partial}{\partial \alpha_k}\log f_{\lambda_k,\alpha_k}(x) = 0.\end{array}\right. \quad (1)$$

Let $\xi_k^{(t)} = \mathbb{P}_{\eta^{(m)}}(S_t = k | X_1^n = x_1^n)$. Since for Weibull emission distributions,

$$\log f_{\lambda_k,\alpha_k}(x) = \log(\alpha_k) + \alpha_k \log(\lambda_k) + (\alpha_k - 1)\log(x) - (\lambda_k x)^{\alpha_k},$$

we have

$$\frac{\partial \log f_{\lambda_k,\alpha_k}(x)}{\partial \lambda_k} = \frac{\alpha_k}{\lambda_k} - \alpha_k x^{\alpha_k} \lambda_k^{\alpha_k - 1}$$

and

$$\frac{\partial \log f_{\lambda_k,\alpha_k}(x)}{\partial \alpha_k} = \frac{1}{\alpha_k} + \log(\lambda_k) + \log x - (\log(\lambda_k x))(\lambda_k x)^{\alpha_k}.$$

The first equation of the system (1) can be rewritten as

$$\begin{aligned}
&\sum_t \xi_k^{(t)} \frac{\partial}{\partial \lambda_k} \log f_{\lambda_k,\alpha_k}(x_t) = \sum_t \xi_k^{(t)} \left[\frac{\alpha_k}{\lambda_k} - \alpha_k x_t^{\alpha_k} \lambda_k^{\alpha_k - 1}\right] = 0 \\
&\quad \Leftrightarrow \frac{\alpha_k}{\lambda_k} \sum_t \xi_k^{(t)} - \alpha_k \lambda_k^{\alpha_k - 1} \sum_t \xi_k^{(t)} x_t^{\alpha_k} = 0 \Leftrightarrow \sum_t \xi_k^{(t)} = \lambda_k^{\alpha_k} \sum_t \xi_k^{(t)} x_t^{\alpha_k} \quad (2) \\
&\quad \Leftrightarrow \lambda_k = \left[\frac{\sum_t \xi_k^{(t)} x_t^{\alpha_k}}{\sum_t \xi_k^{(t)}}\right]^{-\frac{1}{\alpha_k}}. \qquad (3)
\end{aligned}$$

Replacing the expression of λ_k obtained in equation (3) into the second equation of the system (1) yields

$$\begin{aligned}
0 &= \sum_t \xi_k^{(t)} \frac{\partial}{\partial \alpha_k} \log f_{\lambda_k,\alpha_k}(x_t) \\
&= \sum_t \xi_k^{(t)} \left[\frac{1}{\alpha_k} + \log \lambda_k + \log x_t - (\log(\lambda_k x_t))(\lambda_k x_t)^{\alpha_k}\right] \\
&= \sum_t \xi_k^{(t)} \left(\frac{1}{\alpha_k} + \log x_t + \log \lambda_k\right) - \lambda_k^{\alpha_k} \sum_t \xi_k^{(t)} x_t^{\alpha_k} (\log \lambda_k + \log x_t) \\
&= \sum_t \xi_k^{(t)} \left(\frac{1}{\alpha_k} + \log x_t\right) - \lambda_k^{\alpha_k} \sum_t \xi_k^{(t)} x_t^{\alpha_k} \log x_t \\
&\quad + \log \lambda_k \left[\sum_t \xi_k^{(t)} - \lambda_k^{\alpha_k} \sum_t \xi_k^{(t)} {x_t}^{\alpha_k}\right]. \qquad (4)
\end{aligned}$$

Using equations (2) and (3) in (4) yields

$$0 = \sum_t \xi_k^{(t)} \log x_t - \left[\frac{\sum_t \xi_k^{(t)}}{\sum_t \xi_k^{(t)} x_t^{\alpha_k}} \right] \sum_t \xi_k^{(t)} x_t^{\alpha_k} \log x_t + \frac{1}{\alpha_k} \sum_t \xi_k^{(t)}$$

$$\Leftrightarrow \quad 0 = \alpha_k \left[\frac{\sum_t \xi_k^{(t)} \log x_t}{\sum_t \xi_k^{(t)}} - \frac{\sum_t \xi_k^{(t)} x_t^{\alpha_k} \log x_t}{\sum_t \xi_k^{(t)} x_t^{\alpha_k}} \right] + 1. \qquad (5)$$

Equation (5) has no known solution; hence it has to be solved numerically, by the algorithm described in Forsythe et al. (1976) in the circumstances.

3. Markov decision processes

In this Section, a connection between our approach and the theory of Markov decision processes (MDPs) is established. More specifically, the problem of determining the optimal timeout period by minimizing the expected consumption up to following request is shown to be a particular case of an MDP with a continuous action space, if the times between printings are independent random variables. The value function with (finite) horizon 1 of the corresponding MDP is shown to be the opposite of the expected future cost. Moreover, this MDP has a single possible state, which explains why an explicit solution of this problem could be derived in Proposition 1.

3.1. General principle

Markov decision processes are a class of optimization problems for controlling the temporal evolution of an agent in a given environment characterized by a set of states $\mathcal{S}$. At each time step t, given the state S_t of the environment, the agent is allowed to perform an action A_t chosen from a set $\mathcal{A}$. The chosen action may modify the next state S_{t+1} of the environment, and brings a scalar reward R_{t+1} to the agent. All the quantities A_t, S_t and R_t constitute a homogeneous random process. The problem is to determine the distribution for A_t that maximizes the expected future rewards given the current state S_t.

The process $(A_t, S_t, R_t)_{t\in\mathbb{N}}$ is supposed to obey the *Markov property*. Moreover, under the three following assumptions:

(a) the action A_{t+1} is independent on the past of the three processes up to time t, and on R_{t+1}, given state S_{t+1};
(b) the reward R_{t+1} is independent on the past of the three processes up to time t given the states S_t and S_{t+1}, and given A_t;

(c) S_{t+1} is independent on the past of the three processes up to time t given S_t and A_t;

an MDP is totally specified by the following distributions:

- the transition probabilities $\mathcal{P}_{ss'}^{a} = \mathbb{P}(S_{t+1} = s'|S_t = s, A_t = a)$, which define how next state is affected by current state s and the chosen action a;

- the policy function $\pi(s, a) = \mathbb{P}(A_t = a|S_t = s)$, which defines what action to choose given current state s;

- the reward distribution, *i.e.* the distribution of R_{t+1} given $S_t = s$, $A_t = a$ and $S_{t+1} = s'$.

The optimization problem associated with this MDP consists in finding the policy $\pi : \mathcal{S} \times \mathcal{A} \to [0, 1]$ that maximizes the expected future rewards (under the constraints $\sum_a \pi(s, a) = 1$ and $\forall (a, s), \pi(a, s) \geq 0$). The future rewards are modelled through the random variable $\mathfrak{R}_t = \sum_{k=0}^{\infty} \gamma_k R_{t+k+1}$, where $\forall k$, γ_k represents the weight of the reward after $k+1$ time steps. The sequence $(\gamma_k)_{k\in\mathbb{N}}$ is referred to the *discount* sequence. The function to be maximized is called the *value function* and is denoted by V^π; it corresponds to the expectation of $\mathfrak{R}_t$ given the value s of current state. This leads to the following formal definitions:

$$V^\pi(s) = \mathbb{E}(\mathfrak{R}_t|S_t = s) = \sum_{k=0}^{\infty} \gamma_k \mathbb{E}(R_{t+k+1}|S_t = s) \qquad (6)$$

$$\text{and } \hat{\pi}(s, .) = \arg\max_{\pi(s,.)} V^\pi(s). \qquad (7)$$

The reward distribution $\mathbb{P}(R_{t+1}|S_t = s, A_t = a, S_{t+1} = s')$ is only involved through its expectation $\mathcal{R}_{ss'}^{a} = \mathbb{E}(R_{t+1}|S_t = s, A_t = a, S_{t+1} = s')$; consequently, only $\mathcal{R}_{ss'}^{a}$ needs to be defined explicitly.

There are two particular cases of interest for the sequence $(\gamma_k)_{k\in\mathbb{N}}$:

- $\forall k$, $\gamma_k = \gamma^k$, where $0 \leq \gamma < 1$. In this case, $V^\pi(s)$ satisfies a fixed point equation known as the *Bellman equation*. Generally, no closed form is available for the optimal policy.

- $\forall k$, $\gamma_k = \delta_0(k) = \begin{cases} 1 & \text{if } k = 0 \\ 0 & \text{otherwise} \end{cases}$, where δ denotes the Kronecker symbol. Then, it is easily shown that:

$$V^\pi(s) = \sum_a \pi(s, a) \sum_{s'} \mathcal{P}_{ss'}^{a} \mathcal{R}_{ss'}^{a}. \qquad (8)$$

LEMMA 4. *In the case where the sequence* $(\gamma_k)_{k\in\mathbb{N}}$ *is defined by* $\gamma_k = \delta_0(k)$, *the optimal policy is:*

$$\hat{\pi}(s, a') = \begin{cases} 1 & \textit{if } a' = \arg\max_a \sum_{s'} \mathcal{R}^a_{ss'} \mathcal{P}^a_{ss'} \\ 0 & \textit{otherwise.} \end{cases}$$

In the case where the state space is reduced to a singleton $\mathcal{S} = \{1\}$ *and where* $\forall k,\ \gamma_k = \gamma^k$, *the optimal policy is:*

$$\hat{\pi}(1, a') = \begin{cases} 1 & \textit{if } a' = \arg\max_a \mathcal{R}^a_{1,1} \\ 0 & \textit{otherwise.} \end{cases}$$

Thus, the optimal policy corresponds to a deterministic strategy. Given current state s, the chosen action systematically is the one that maximizes $\sum_{s'} \mathcal{R}^a_{ss'} \mathcal{P}^a_{ss'}$ or $\mathcal{R}^a_{1,1}$ (depending on the sequence $(\gamma_k)_{k\in\mathbb{N}}$), with respect to a.

3.2. Connection of our approach with MDPs

In this paragraph, our approach is shown to be a particular case of MDP. The proof is derived in the case of one single sleep mode printer for the sake of simplicity, but can easily be extended to an arbitrary number of sleep modes.

In our context, the set of actions for the printer is the timeout period $a \in \mathcal{A} = \mathbb{R}^+$, which corresponds to $\tau^{(1)}$ in previous paragraphs. The decision is taken after each print job, after which the printer is necessarily in *idle* mode. Consequently, the state space is $\mathcal{S} = \{idle\}$. In our problem, the reward is minus the cost between two successive printings. It only depends on the time between printings X_i and on the action a, *i.e.* the timeout period.

The time index $i \in \mathbb{N}$ represents the number of past printings requests. Hence, even if the times of requests T_i take continuous values, the MDP is essentially a discrete-time problem, where decisions are taken after each print job only.

Let the expected reward be defined as $\mathcal{R}^a = -\mathbb{E}(h(X_i, a))$, where h is the cost between two successive print jobs defined by equation (10) in the case of multiple sleep modes. The transition probabilities are $\mathcal{P}^a_{ss'} = 1,\ \forall\ a,\ s$ and s', since the printer is always in *idle* state when a decision is taken.
As a consequence, from equation (8), the value function $V^\pi(s)$ for $\gamma_k = \delta_0(k)$ is:

$$V^\pi(s) = -\sum_a \pi(s, a)\mathbb{E}(h(X_i, a)),$$

and according to Lemma 4, the optimal policy is:

$$\pi(s,a') = \begin{cases} 1 & \text{if } a' = \arg\max_a -\mathbb{E}(h(X_i,a)) = \arg\min_a \mathbb{E}(h(X_i,a)) \\ 0 & \text{otherwise,} \end{cases}$$

which corresponds to the optimization problem $\hat{\tau}_i \in \arg\min_\tau \mathbb{E}(h(X_i,\tau)|X_{1:i-1})$, in the case where the times between printings X_i are independent random variables.

To conclude, our approach is a degenerate case of an MDP problem with a continuous action space and with one single state. Using the particular discount sequence $\gamma_k = \delta_0(k)$, the expected future cost coincides with the value function of the MDP with (finite) horizon 1. Since the state space is reduced to a single state, an explicit solution of this problem can be derived. This solution is given in Proposition 1.

4. Proofs

Proof of Lemma 1

In view of Section 2, the consumption between two successive printings is

$$h(X_i,\tau_i) = (aX_i)(1 - \mathbb{1}_{\{X_i>\tau_i\}}) + (a\tau_i + c + b(X_i - \tau_i) + d)\mathbb{1}_{\{X_i>\tau_i\}}.$$

Introducing $\Delta t = (c+d)/(a-b)$, the consumption can be rewritten as

$$h(X_i,\tau_i) = aX_i + (a-b)(\Delta t + \tau_i - X_i)\mathbb{1}_{\{X_i>\tau_i\}}.$$

As a consequence, the expected consumption between two successive printings given $X_{1:i-1}$ is

$$\begin{aligned} \mathbb{E}(h(X_i,\tau_i)|X_{1:i-1}) &= a\mathbb{E}(X_i|X_{1:i-1}) + (a-b)\bar{F}_i(\tau_i)(\Delta t + \tau_i) \\ &- (a-b)\int_{\tau_i}^{+\infty} x f_i(x)dx. \end{aligned}$$

and the result is proved. ■

Proof of Lemma 2.
The consumption h between two successive print requests is given by:

$$\begin{aligned} & h(X_i, \tau_i^{(1)}, \dots, \tau_i^{(m)}) \qquad (9) \\ =& a X_i \mathbb{1}_{\{X_i \le \tau_i^{(1)}\}} \\ &+ \sum_{r=1}^{m-1} \sum_{j=1}^{r-1} \left(c_j + b_j(\tau_i^{(j+1)} - \tau_i^{(j)}) \right) \mathbb{1}_{\{\tau_i^{(r)} < X_i \le \tau_i^{(r+1)}\}} \\ &+ \sum_{r=1}^{m-1} \left(a\tau_i^{(1)} + b_r(X_i - \tau_i^{(r)}) + c_r + d_r \right) \mathbb{1}_{\{\tau_i^{(r)} < X_i \le \tau_i^{(r+1)}\}} \\ &+ \sum_{j=1}^{m-1} \left(c_j + b_j(\tau_i^{(j+1)} - \tau_i^{(j)}) \right) \mathbb{1}_{\{X_i > \tau_i^{(m)}\}} \\ &+ \left(a\tau_i^{(1)} + c_m + b_m(X_i - \tau_i^{(m)}) + d_m \right) \mathbb{1}_{\{X_i > \tau_i^{(m)}\}}. \qquad (10) \end{aligned}$$

Letting $a = b_0$ yields

$$\begin{aligned} & \mathbb{E}\left(h(X_i, \tau_i^{(1)}, \dots, \tau_i^{(m)}) | X_{1:i-1} \right) \\ &= b_0 \mathbb{E}\left(X_i \mathbb{1}_{\{X_i < \tau_i^{(1)}\}} | X_{1:i-1} \right) + \sum_{r=1}^{m-1} b_r \mathbb{E}\left(X_i \mathbb{1}_{\{\tau_i^{(r)} < X_i \le \tau_i^{(r+1)}\}} | X_{1:i-1} \right) \\ &\quad + b_m \mathbb{E}\left(X_i \mathbb{1}_{\{X_i > \tau_i^{(m)}\}} | X_{1:i-1} \right) \\ &\quad + \sum_{r=1}^{m-1} \mathbb{P}\left(\tau_i^{(r)} < X_i \le \tau_i^{(r+1)} | X_{1:i-1} \right) \left(\sum_{j=1}^{r} \left(\tau_i^{(j)}(b_{j-1} - b_j) + c_j \right) + d_r \right) \\ &\quad + \mathbb{P}\left(X_i > \tau_i^{(m)} | X_{1:i-1} \right) \left(\sum_{j=1}^{m} \left(\tau_i^{(j)}(b_{j-1} - b_j) + c_j \right) + d_m \right). \end{aligned}$$

Taking account of

$$\begin{aligned} \mathbb{E}\left(X_i \mathbb{1}_{\{\tau_i^{(r)} < X_i \le \tau_i^{(r+1)}\}} | X_{1:i-1} \right) &= \mathbb{E}\left(X_i \mathbb{1}_{\{X_i > \tau_i^{(r)}\}} | X_{1:i-1} \right) \\ &- \mathbb{E}\left(X_i \mathbb{1}_{\{X_i > \tau_i^{(r+1)}\}} | X_{1:i-1} \right), \\ \mathbb{E}\left(X_i \mathbb{1}_{\{X_i < \tau_i^{(1)}\}} | X_{1:i-1} \right) &= \mathbb{E}\left(X_i | X_{1:i-1} \right) \\ &- \mathbb{E}\left(X_i \mathbb{1}_{\{X_i > \tau_i^{(1)}\}} | X_{1:i-1} \right) \end{aligned}$$

the expected consumption can be rewritten as

$$\begin{aligned}
&\mathbb{E}\left(h(X_i,\tau_i^{(1)},\ldots,\tau_i^{(m)})|X_{1:i-1}\right)\\
&=-\sum_{j=1}^{m}(b_{j-1}-b_j)\mathbb{E}\left(X_i\mathbb{1}_{\{X_i>\tau_i^{(j)}\}}|X_{1:i-1}\right)+b_0\mathbb{E}(X_i|X_{1:i-1})\\
&\quad+\left(1-F_{X_i|X_{1:i-1}}(\tau_i^{(m)})\right)\left(\sum_{j=1}^{m}\left(\tau_i^{(j)}(b_{j-1}-b_j)+c_j\right)+d_m\right)\\
&\quad+\sum_{r=1}^{m-1}\left(\left(F_i(\tau_i^{(r+1)})-F_i(\tau_i^{(r)})\right)\left(\sum_{j=1}^{r}\left(\tau_i^{(j)}(b_{j-1}-b_j)+c_j\right)+d_r\right)\right).
\end{aligned}$$

Splitting the second right-hand term into two parts yields

$$\begin{aligned}
&\mathbb{E}\left(h(X_i,\tau_i^{(1)},\ldots,\tau_i^{(m)})|X_{1:i-1}\right)\\
&=-\sum_{j=1}^{m}(b_{j-1}-b_j)\mathbb{E}\left(X_i\mathbb{1}_{\{X_i>\tau_i^{(j)}\}}|X_{1:i-1}\right)+\sum_{j=1}^{m}\left(\tau_i^{(j)}(b_{j-1}-b_j)+c_j\right)+d_m\\
&\quad+b_0\mathbb{E}(X_i|X_{1:i-1})+\sum_{r=2}^{m}F_i(\tau_i^{(r)})\left(\sum_{j=1}^{r-1}\left(\tau_i^{(j)}(b_{j-1}-b_j)+c_j\right)+d_{r-1}\right)\\
&\quad-\sum_{r=1}^{m}F_i(\tau_i^{(r)})\left(\sum_{j=1}^{r}\left(\tau_i^{(j)}(b_{j-1}-b_j)+c_j\right)+d_r\right)
\end{aligned}$$

and collecting the two last right-hand terms we obtain

$$\begin{aligned}
&\mathbb{E}\left(h(X_i,\tau_i^{(1)},\ldots,\tau_i^{(m)})|X_{1:i-1}\right)\\
&=-\sum_{j=1}^{m}(b_{j-1}-b_j)\mathbb{E}\left(X_i\mathbb{1}_{\{X_i>\tau_i^{(j)}\}}|X_{1:i-1}\right)+b_0\mathbb{E}(X_i|X_{1:i-1})\\
&\quad+\left(\sum_{j=1}^{m}\left(\tau_i^{(j)}(b_{j-1}-b_j)+c_j\right)+d_m\right)-F_i(\tau_i^{(1)})\left(\tau_i^{(1)}(b_0-b_1)+d_1\right)\\
&\quad-\sum_{j=2}^{m}F_i(\tau_i^{(j)})\left(\tau_i^{(j)}(b_{j-1}-b_j)+c_j+d_j-d_{j-1}\right).
\end{aligned}$$

Finally, letting $\Delta t_j = (c_j + d_j - d_{j-1})/(b_{j-1} - b_j)$ with the convention $d_0 = 0$, we have

$$
\begin{aligned}
&\mathbb{E}\left(h(X_i, \tau_i^{(1)}, \dots, \tau_i^{(m)})|X_{1:i-1}\right) \\
&= -\sum_{j=1}^{m}(b_{j-1} - b_j)\mathbb{E}\left(X_i \mathbb{1}_{\{X_i > \tau_i^{(j)}\}}|X_{1:i-1}\right) + \sum_{j=1}^{m}(\tau_i^{(j)} + \Delta t_j)(b_{j-1} - b_j) \\
&\quad - \sum_{j=1}^{m}(b_{j-1} - b_j)\, F_i(\tau_i^{(r)})(\tau_i^{(j)} - \Delta t_j) + b_0\mathbb{E}(X_i|X_{1:i-1}) \\
&= \sum_{j=1}^{m}(b_{j-1} - b_j)\left[\left(1 - F_i(\tau_i^{(j)})\right)\left(\Delta t_j + \tau_i^{(j)}\right) - \int_{\tau_i^{(j)}}^{+\infty} x f_i(x)dx\right] \\
&\quad + b_0\mathbb{E}(X_i|X_{1:i-1}), \qquad (11)
\end{aligned}
$$

and the conclusion follows. ■

Proof of Proposition 1

Differentiating the above expected consumption with respect to τ_i yields

$$\frac{d\mathbb{E}(h(X_i, \tau_i)|X_{1:i-1})}{d\tau_i} = (a - b_1)\bar{F}_i(\tau_i)(1 - \Delta t z_i(\tau_i)). \qquad (12)$$

Let us recall that $a > b$. Two main cases are considered:
(i) Suppose that z_i is decreasing. Three situations occur:
- If $1/\Delta t < \lim_{x\to+\infty} z_i(x)$, then the derivative (12) is negative, the expected consumption is a stricly decreasing function of τ_i and thus $\hat{\tau}_i = +\infty$.
- If $\lim_{x\to+\infty} z_i(x) \le 1/\Delta t \le z_i(0)$, then the following equation

$$z_i(\tau_i^{(1)}) = 1/\Delta t \qquad (13)$$

has an unique root τ_i in $(0, +\infty)$ which is the unique minimum of the expected consumption.
- Finally, if $z_i(0) < 1/\Delta t$, then the derivative (12) is positive, the expected consumption is a strictly increasing function of τ_i and thus $\hat{\tau}_i = 0$.
(ii) Suppose that z_i is increasing or constant. Three situations occur:
- If $1/\Delta t \le z_i(0)$, then the derivative (12) is non-positive, the expected consumption is a non-increasing function of τ_i and thus $\hat{\tau}_i = +\infty$.
- If $z_i(0) < 1/\Delta t < \lim_{x\to+\infty} z_i(x)$ then equation (13) has an unique root in $(0, +\infty)$

and the expected consumption is a concave function of τ_i. As a consequence, $\hat{\tau}_i = 0$ if $\mathbb{E}\left(h(X_i,0)|X_{1:i-1}\right) < \lim_{x\to\infty} \mathbb{E}\left(h(X_i,x)|X_{1:i-1}\right)$ and $\hat{\tau}_i = +\infty$ otherwise. Since

$$\mathbb{E}\left(h(X_i,0)|X_{1:i-1}\right) - \lim_{x\to\infty} \mathbb{E}\left(h(X_i,x)|X_{1:i-1}\right) = (a-b)(\Delta t - \mathbb{E}(X_i|X_{1:i-1})),$$

the conclusion follows.
- Finally, if $\lim_{x\to+\infty} z_i(x) \leq 1/\Delta t$, then the derivative (12) is non-negative, the expected consumption is a non-decreasing function of τ_i and $\hat{\tau}_i = 0$. ■

Proof of Proposition 1 is quite similar to that of Proposition 1, and thus is omitted. This is also the case for Lemma 3, which proof is similar to that of Lemma 1.

Proof of Lemma 4. In the case where $\gamma_k = \begin{cases} 1 & \text{if } k = 0 \\ 0 & \text{else} \end{cases}$, then

$$\begin{aligned} & V^\pi(s) \\ & = \mathbb{E}(R_{t+1}|S_t = s) \\ & = \sum_a \pi(s,a) \sum_{s'} \mathbb{E}(R_{t+1}|S_t = s, S_{t+1} = s', A_t = a)\mathbb{P}(S_{t+1} = s'|S_t = s, A_t = a) \\ & = \sum_a \pi(s,a) \sum_{s'} \mathcal{R}^a_{ss'} \mathcal{P}^a_{ss'}. \end{aligned} \tag{14}$$

The optimal policy is the policy π maximizing the value function V^π:

$$\hat{\pi} = \arg\max_\pi \left(\sum_a \pi(s,a) \sum_{s'} \mathcal{R}^a_{ss'} \mathcal{P}^a_{ss'} \right)$$

with $\sum_a \pi(s,a) = 1$ and $\pi(s,a) \geq 0 \ \forall (s,a)$.
In the case where the state space is reduced to a singleton $\mathcal{S} = \{1\}$, and if $\gamma_k = \gamma^k \quad \forall k$, the optimal policy is, from Bellman equation (see Sutton and Barto (1998)):

$$\begin{aligned} V^\pi(1) &= \sum_a \pi(s,a) \sum_{s'} \mathcal{P}^a_{ss'} \left[\mathcal{R}^a_{ss'} + \gamma V^\pi(s')\right] \\ &= \sum_a \pi(1,a)\mathcal{R}^a_{1,1} + \gamma V^\pi(1) \sum_a \pi(1,a) \end{aligned}$$

since $\mathcal{S} = \{1\}$ and $\forall a$, $\mathcal{P}^a_{ss'} = 1$. Remarking that $\sum_a \pi(1,a) = 1$, we have

$$\begin{aligned}(1-\gamma)V^\pi(1) &= \sum_a \pi(1,a)\mathcal{R}^a_{1,1} \\ \Leftrightarrow \quad V^\pi(1) &= \frac{1}{(1-\gamma)}\sum_a \pi(1,a)\mathcal{R}^a_{1,1}. \end{aligned} \tag{15}$$

In both equations (14) and (15), for each s, $V^\pi(s)$ is a linear function with respect to $\pi(s,a)$. Thus, a classical result of the optimization theory states that the maximum of $V^\pi(s)$ is achieved on an endpoint of an edge of the simplex $\left\{\pi(s,a)|a \in \mathbb{R}, \pi(s,a) \geq 0 \text{ and } \sum_a \pi(s,a) = 1\right\}$.
As a consequence, in the case where the sequence $(\gamma_k)_{k\in\mathbb{N}}$ is defined by $\gamma_k = \delta_0(k)$, the optimal policy is:

$$\hat{\pi}(s,a') = \begin{cases} 1 & \text{if } a' = \arg\max_a \sum_{s'} \mathcal{R}^a_{ss'}\mathcal{P}^a_{ss'} \\ 0 & \text{otherwise.} \end{cases}$$

In the case where the state space is reduced to a singleton $\mathcal{S} = \{1\}$ and where $\forall k$, $\gamma_k = \gamma^k$, the optimal policy is:

$$\hat{\pi}(1,a') = \begin{cases} 1 & \text{if } a' = \arg\max_a \mathcal{R}^a_{1,1} \\ 0 & \text{otherwise.} \end{cases}$$

■

References

Ephraim, Y. and Merhav, N. (2002) Hidden Markov processes. *IEEE Transactions on Information Theory*, **48**, 1518–1569.

Forsythe, G. E., Malcolm, M. A. and Moler, C. B. (1976) *Computer Methods for Mathematical Computations.* Prentice-Hall.

Sutton, R. S. and Barto, A. G. (1998) *Reinforcement Learning: An Introduction.* MIT Press, Cambridge, Massachusetts.

5. Figures

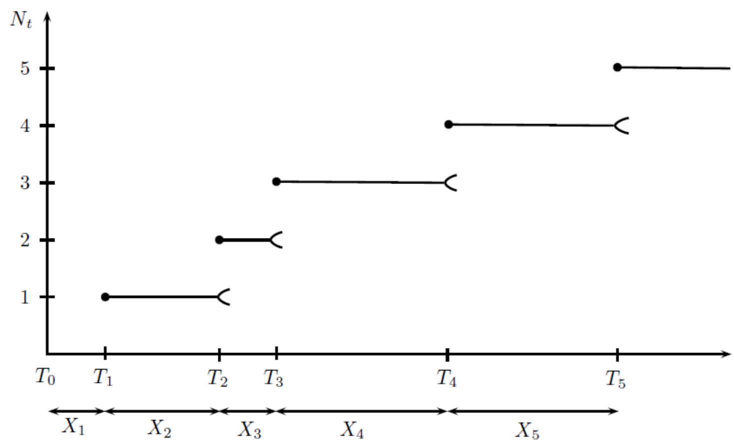
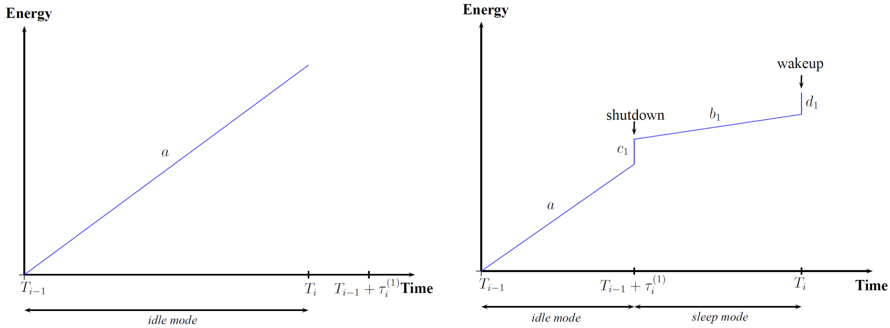


Figure 1. Print Process. Process $\{N_t\}_{t\geq 0}$ refers to the counting process of print requests, which is defined as the cumulative number of print requests between 0 and t, that is, $N_t = \max\{i \in \mathbb{N}; T_i \leq t\}$. On the x-axis, the times of print requests T_i and the times between requests X_i are also depicted. The three processes can be deduced from each other.



(a) When $X_i \leq \tau_i^{(1)}$ (b) When $X_i > \tau_i^{(1)}$

Figure 2. Energy consumption between T_{i-1} and T_i according to the position of T_{i-1}, T_i and $T_{i-1} + \tau_i^{(1)}$

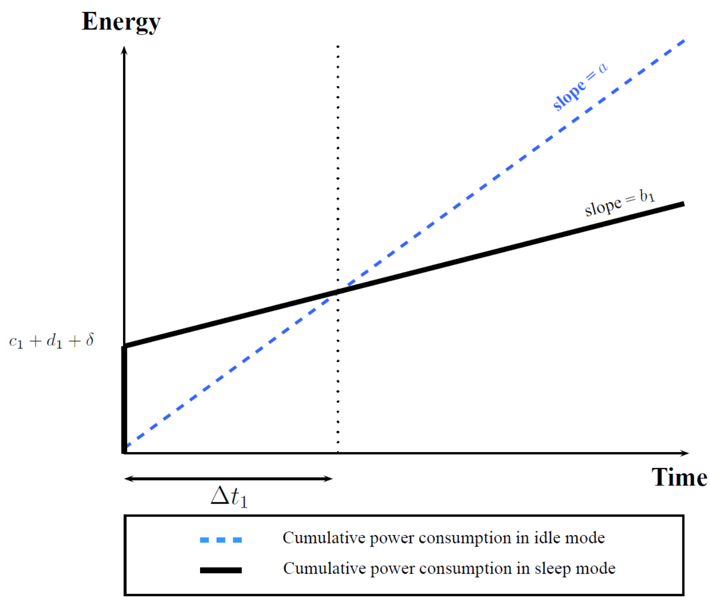


Figure 3. Graphical interpretation of Δt_1

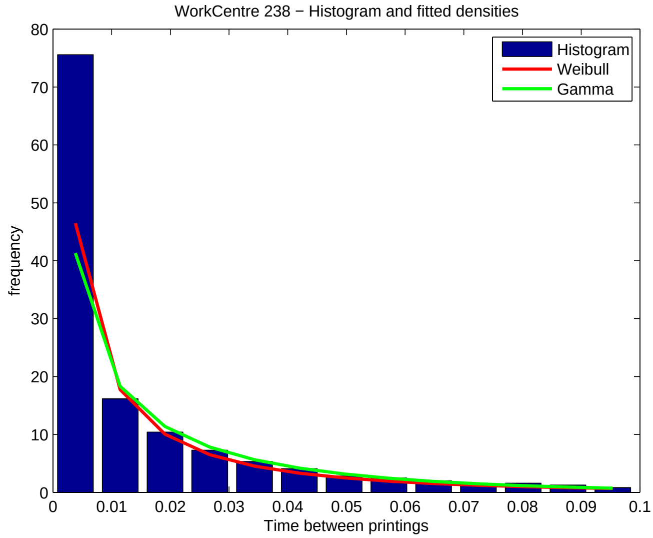


Figure 4. Histogram and fitted Weibull and Gamma pdfs for the WorkCentre 238 dataset

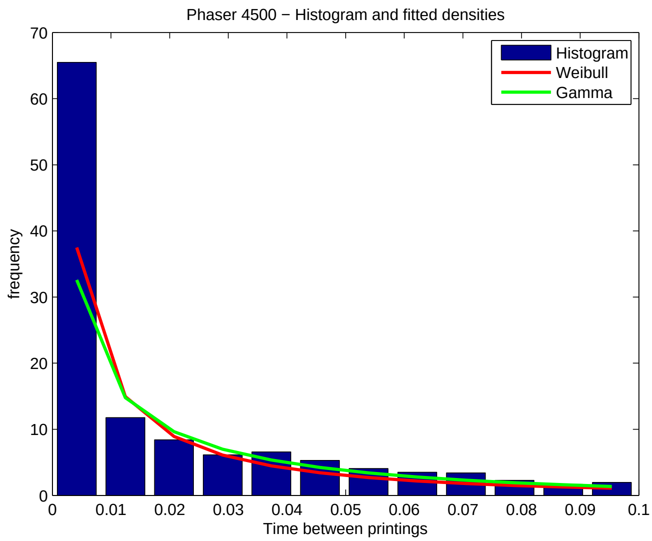


Figure 5. Histogram and fitted Weibull and Gamma pdfs for the Phaser 4500 dataset

6. Tables

Table 1. Empirical skewness and kurtosis for the WorkCentre 238 and Phaser 4500 datasets

	WorkCentre 238	Phaser 4500
Skewness	2.98	2.39
Kurtosis	14.7	9.5

Table 2. Energy consumption and mean computation time associated to the different strategies (including variants based on filtering and full conditional distribution in HMC models).

	Total consumption (kWh)			Standard deviation of consumption			Mean computation time by sample (ms)		
Sample size	361	121	61	361	121	61	361	121	61
Energy Star	500	500	500	7.99	4.73	3.04	1.2e+00	1.0e+00	2.0e+00
$\tau^{(1)} = \tau^{(2)} = 0$	498	498	498	6.77	3.76	2.55	2.0e+00	1.0e+00	1.0e+00
Exhaustive search	446	446	447	6.86	4.17	2.76	6.6e+04	1.5e+05	2.7e+05
Oracle	399	399	399	7.13	4.11	2.72	2.0e+00	2.0e+00	2.0e+00
c-competitive	471	471	471	7.66	4.54	2.94	5.0e-01	1.0e+00	2.0e+00
Static	446	446	446	7.02	4.13	2.77	2.0e+01	5.0e+01	9.0e+01
Sliding window	445	445	444	7.02	4.18	2.77	5.2e+05	1.5e+05	6.6e+04
Viterbi	471	464	462	8.07	3.97	2.64	1.0e+04	4.3e+03	2.8e+03
Filtering	472	462	462	8.29	3.93	2.62	1.3e+03	1.7e+02	1.9e+02
Conditional	456	454	450	5.62	4.04	2.72	5.8e+04	5.4e+04	6.0e+04