Time-Constrained Energy Minimization for Online Execution of a Stochastic DAG Task

Jonatha Anselmi, Bruno Gaujal, **Karl Gottlieb**

Univ. Grenoble Alpes, Inria, CNRS, Grenoble INP, LIG

February 24, 2026

Features of our problem

Objective

Minimize average energy consumption while ensuring that the task is completed before some hard real-time deadline.

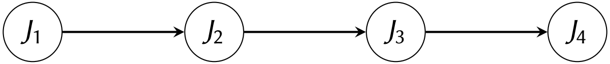
Processing Unit

- *Speed* $s \in \mathbb{R}_+$
- *Energy consumption* per unit of speed $s : Q(s) = s^{\alpha-1}$ (with $\alpha > 2$)

Deadline

- Computation *start* at time 0
- Task **must** be completed before *deadline* $D > 0$

Task



Example of a task made of four elementary jobs

Task

$$J_1 \longrightarrow J_2 \longrightarrow J_3 \longrightarrow J_4$$

Example of a task made of four elementary jobs

Hypothesis on jobs

- Size of job J_k in $[0, W_k^{\mathsf{max}}]$ with $W_k^{\mathsf{max}} < \infty$
- Size of jobs *unknown* until they are completed
- *Cumulative distribution* and *Tail* functions F_k and F_k^c
- *Independent*
- *Sequential* execution

Objective and existing approaches

Objective

Minimize average energy consumption while ensuring that the task is completed before some hard real-time deadline.

Objective and existing approaches

Objective

Minimize average energy consumption while ensuring that the task is completed before some hard real-time deadline.

- **offline** case : size of jobs *known*, computations *before* execution.
- **online** case : size of jobs *unknown*, discovered when a job is completed, and computation *during* execution.

Objective and existing approaches

Objective

Minimize average energy consumption while ensuring that the task is completed before some hard real-time deadline.

- **offline** case : size of jobs *known*, computations *before* execution.
- **online** case : size of jobs *unknown*, discovered when a job is completed, and computation *during* execution.

Our case: *Online*, and we use the *work* instead of *time*.

Problem Statement

Objective

Minimize average energy consumption while ensuring that the task is completed before the deadline.

Optimization problem

Minimize over all speed profile s the quantity $\mathcal{E}^s_{J_1,\ldots,J_K}(D)$ defined by

$$\mathcal{E}^s_{J_1,\ldots,J_K}(D) := \mathbb{E}\left[\int_0^{W_1+\ldots+W_K} Q(s(w))\mathrm{d}w\right] = \mathbb{E}\left[\int_0^{W_1+\ldots+W_K} s(w)^{\alpha-1}\mathrm{d}w\right]$$

$$\text{subject to: } \int_0^{W_1^{\max}+\ldots+W_K^{\max}} \frac{\mathrm{d}w}{s(w)} \leqslant D.$$

One job

The problem can be rewritten as

$$\mathcal{E}^*_{J_1}(D) := \inf_{s_1:[0,W_1^{\mathsf{max}}]\to\mathbb{R}_+} \int_0^{W_1^{\mathsf{max}}} s_1(w)^{\alpha-1} F_1^c(w)\mathrm{d}w.$$

One job

The problem can be rewritten as

$$\mathcal{E}^*_{J_1}(D) := \inf_{s_1:[0,W_1^{\max}]\to\mathbb{R}_+} \int_0^{W_1^{\max}} s_1(w)^{\alpha-1} F_1^c(w)\mathrm{d}w.$$

Known from the literature [Lorch and Smith, 2001]:

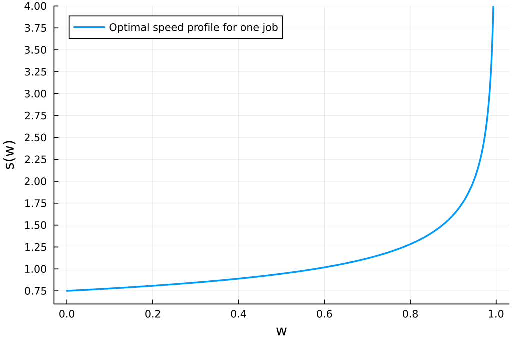
$$s_1^*(w) = \frac{C_1}{D\, F_1^c(w)^{\frac{1}{\alpha}}}$$

where

$$C_1 = \int_0^{W_1^{\max}} F_1^c(w)^{\frac{1}{\alpha}}\,\mathrm{d}w.$$

We get

$$\mathcal{E}^*_{J_1}(D) = \frac{C_1^\alpha}{D^{\alpha-1}}.$$



Optimal speed profile for a job given by a uniform distribution on $[0, 1]$.

Several jobs is challenging

- **Adaptive** policy : The jobs are considered as one, and the speed profile is refreshed when any job is completed.

- *d*-**Sweep** policy : A deadline is fixed for each job and compute the optimal partial deadline instead.
  It was used by [Lu, Zhang, Stan, Lach and Skadron, 2006] and claimed to be optimal without proof.

Several jobs is challenging

- **Adaptive** policy : The jobs are considered as one, and the speed profile is refreshed when any job is completed.
- *d*-**Sweep** policy : A deadline is fixed for each job and compute the optimal partial deadline instead.
  It was used by [Lu, Zhang, Stan, Lach and Skadron, 2006] and claimed to be optimal without proof.

Our approach: Compute explicitly using the optimality principle

Optimality principle for two jobs

$$\mathcal{E}^{s_1}_{J_1,J_2}(D) = \int_0^{W_1^{\max}} s_1(w)^{\alpha-1} F_1^c(w)\mathrm{d}w + \dots$$

Optimality principle for two jobs

$$\mathcal{E}^{s_1}_{J_1,J_2}(D) = \int_0^{W_1^{\mathsf{max}}} s_1(w)^{\alpha-1} F_1^c(w)\mathrm{d}w + \int_0^{W_1^{\mathsf{max}}} \mathcal{E}^*_{J_2}\left(D - \int_0^w \frac{\mathrm{d}u}{s_1(u)}\right)\mathrm{d}F_1(w)$$

Optimality principle for two jobs

Theorem

$$\mathcal{E}^*_{J_1,J_2}(D) = \inf_{s_1:[0,W_1^{\mathsf{max}}]\to\mathbb{R}^*_+} \left(\int_0^{W_1^{\mathsf{max}}} s_1(w)^{\alpha-1} F_1^c(w)\mathrm{d}w + \int_0^{W_1^{\mathsf{max}}} \mathcal{E}^*_{J_2}\left(D - \int_0^w \frac{\mathrm{d}u}{s_1(u)}\right) \mathrm{d}F_1(w) \right)$$

Two jobs : Optimal solution

Theorem (Optimal speed profile and mean energy consumption for two jobs)

$$s_1^*(w) = \frac{1 + \int_0^{W_1^{\max}} y(u)\mathrm{d}u}{D\, y(w)} \qquad \textit{and} \qquad s_2^*(w) = \frac{C_2}{\left(D - \int_0^{w_1} \frac{\mathrm{d}u}{s_1^*(u)}\right) F_2^c(w)^{\frac{1}{\alpha}}}$$

under $[W_1 \leqslant w_1]$ *and* y *is a solution of the following second order ODE (with* $\int y,\ y,\ y'$*):*

$$\frac{y'(w)}{y(w)} = \frac{1}{\alpha}\frac{\mathrm{d}F_1(w)}{F_1^c(w)}\left(C_2^{\alpha}\left(\frac{y(w)}{1 + \int_w^{W_1^{\max}} y(u)\mathrm{d}u}\right)^{\alpha} - 1\right)$$

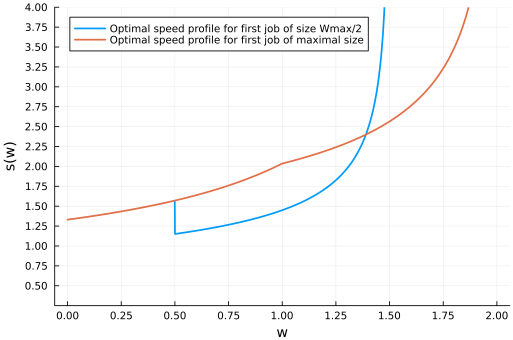
with initial condition $y(W_1^{\max}) = C_2^{-\frac{1}{\alpha}}, y'(W_1^{\max}) = 0$. *In particular,*

$$\mathcal{E}^*_{J_1,J_2}(D) = \frac{C_{1,2}}{D^{\alpha-1}}$$

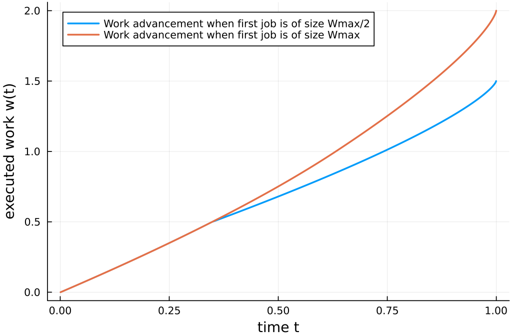
where $C_{1,2}$ *is a constant that does not depend on* D.

Online example with two jobs

Job sizes W_1, W_2 follows uniform distribution on $[0, 1]$ and $D = 1$.

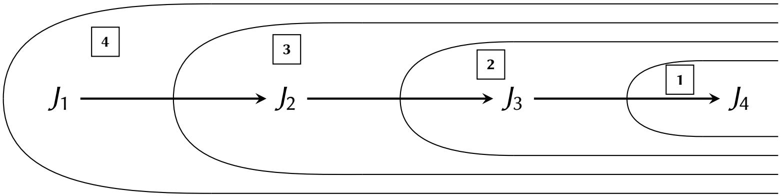


Speed profile given work executed

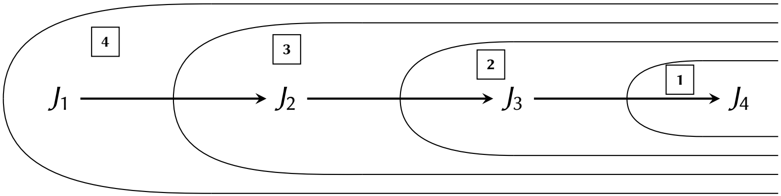


Work advancement given time

Extension to tasks of $K \geq 2$ sequential jobs



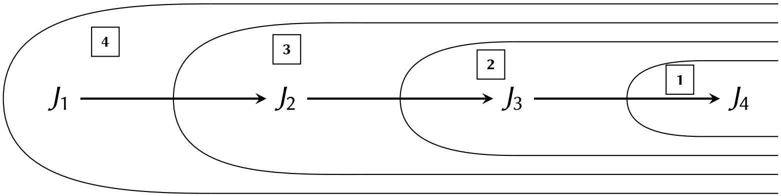
Extension to tasks of $K \geq 2$ sequential jobs



Computing $C_{k,\dots,K}$ an *offline constant* by *Backward Induction*

$$
\begin{aligned}
\mathcal{E}^*_{J_k,\dots,J_K}(D) &= \frac{C_{k,\dots,K}}{D^{\alpha-1}} \\
&= \inf_{s_k} \left\{ \int_0^{W_k^{\max}} s_k(w)^{\alpha-1} F_k^c(w)\mathrm{d}w + \int_0^{W_k^{\max}} \frac{C_{k+1,\dots,K}}{\left(D - \int_0^w \frac{\mathrm{d}u}{s_k(u)}\right)^{\alpha-1}} \mathrm{d}F_k(w) \right\}
\end{aligned}
$$

Extension to tasks of $K \geq 2$ sequential jobs

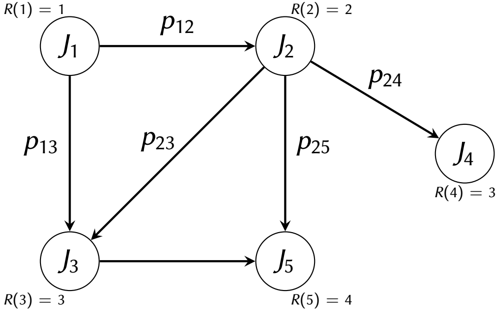


Computing $C_{k,\dots,K}$ an *offline constant* by *Backward Induction*

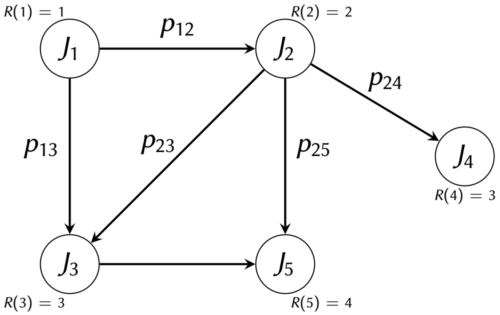
$$
\begin{aligned}
\mathcal{E}^*_{J_k,\dots,J_K}(D) &= \frac{C_{k,\dots,K}}{D^{\alpha-1}} \\
&= \inf_{s_k} \left\{ \int_0^{W_k^{\max}} s_k(w)^{\alpha-1} F_k^c(w)\mathrm{d}w + \int_0^{W_k^{\max}} \frac{C_{k+1,\dots,K}}{\left(D - \int_0^w \frac{\mathrm{d}u}{s_k(u)}\right)^{\alpha-1}} \mathrm{d}F_k(w) \right\}
\end{aligned}
$$

Then computing the *online* optimal speed by *Forward induction.*

Solving flowcharts

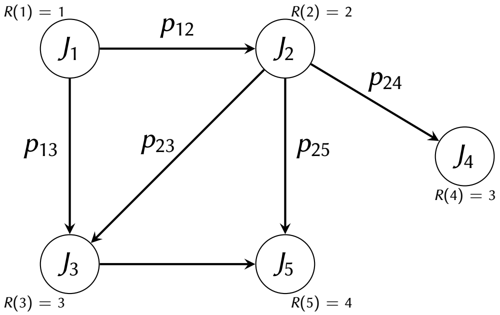


Solving flowcharts



$$\mathcal{E}^*_{G_1}(D) = \inf_{s_1} \left(\int_0^{W_1^{\max}} s_1(w)^{\alpha-1} F_1^c(w) \mathrm{d}w + \int_0^{W_1^{\max}} \frac{\sum_{j \in \mathrm{Succ}(1)} p_{1j}(w) C_{G_j}(W_j^{\max})^{\alpha}}{\left(D - \int_0^w \frac{\mathrm{d}u}{s_1(u)}\right)^{\alpha-1}} \mathrm{d}F_1(w) \right)$$

Solving flowcharts



$$\mathcal{E}^*_{G_1}(D) = \inf_{s_1} \left(\int_0^{W_1^{\max}} s_1(w)^{\alpha-1} F_1^c(w) \mathrm{d}w \right.$$
$$\left. + \int_0^{W_1^{\max}} \frac{\sum_{j \in \mathrm{Succ}(1)} p_{1j}(w) C_{G_j}(W_j^{\max})^{\alpha}}{\left(D - \int_0^w \frac{\mathrm{d}u}{s_1(u)}\right)^{\alpha-1}} \mathrm{d}F_1(w) \right)$$

Global complexity is of order of $O(AN)$ where A denotes the number of arcs of the DAG and N denotes the partition size of the discretization.