An $\Omega(\Delta^{1/2})$ Gap example on the (W)SEPT Policy

– unpublished note –

Wang Chi Cheung* Felix Fischer† Jannik Matuschke† Nicole Megow†

August 2014

We consider scheduling jobs with stochastic processing times on identical parallel machines. The task is to find an adaptive scheduling policy that minimizes the sum of (weighted) expected completion times. We denote this problem in the standard notation as $\mathrm{P}||\mathbb{E}\left[\sum(w_j)C_j\right]$. The stochastic problem has received quite some attention in the literature. The single machine problem can be solved optimally by ordering jobs in non-increasing order of the ratio of weight over expected processing time [4]. This generalization of the well-known deterministic *Smith rule* [6] is called *Weighted Shortest Processing Time* (WSEPT) rule. For two or more machines, WSEPT is not optimal anymore. Several approximation algorithms have been proposed. To date, all algorithms have an approximation guarantee in the order $\mathcal{O}(\Delta)$, where Δ is an upper bound on the squared coefficients of variation of the processing time distributions P_j, that is, $\mathrm{Var}[P_j]/\mathbb{E}\left[P_j\right]^2 \leq \Delta$ for all jobs j. It is a major open question whether there is a constant factor approximation algorithm for this problem. For proper definitions and related work, we refer to [2, 3, 5].

Recently, Benjamin Labonte [1] showed in his master thesis a lower bound of $\Omega(\Delta^{1/4})$ on the approximation ratio for WSEPT. Here, we give a stronger example and demonstrate a gap of $\Omega(\Delta^{1/2})$ for the unweighted setting, i.e., the performance of SEPT.

Theorem 1. *There is a lower bound of* $\Omega(\Delta^{1/2})$ *on the approximation guarantee of* SEPT *for* $\mathrm{P}||\mathbb{E}\left[\sum C_j\right]$.

Proof. Consider the following scheduling instance on m parallel identical machines and two types of jobs, denoted by A and B. We denote the number of type-A (type-B) jobs by n_A (n_B), and their length by L_A (L_B). Let $\epsilon \in (0,1]$. The parameters are set as follows:

$$\text{Jobs of type } A\text{: } n_A = m^2, L_A = \frac{1}{m},$$

$$\text{Jobs of type } B\text{: } n_B = \frac{m^3}{4}, L_B = \begin{cases} 0 & \text{with prob } 1-\frac{1}{m^2} \\ m^{1+\epsilon} & \text{with prob } \frac{1}{m^2} \end{cases}.$$

While jobs of type A are deterministic, B-jobs are stochastic with high coefficient of variation. Note that

$$\mathbb{E}\left[L_B\right] = m^{-1+\epsilon} > L_A, \text{ and } \Delta = CV[L_B] = \frac{\mathbb{E}\left[L_B^2\right]}{\mathbb{E}\left[L_B\right]^2} - 1 = \frac{m^{2+2\epsilon}m^{-2}}{m^{-2+2\epsilon}} - 1 = m^2 - 1.$$

WESPT will schedule jobs of type A first and then jobs of type B. Let $Cost(A \prec B)$ denote the expected cost of WSEPT and let $Cost(B \prec A)$ denote the expected of list scheduling according to the priority order in which B-jobs precede A-jobs. To demonstrate the gap in this example, we will show that

$$\frac{Cost(A \prec B)}{Cost(B \prec A)} = \Omega(m^{1-\epsilon}) = \Omega(\Delta^{\frac{1-\epsilon}{2}}).$$

First, we show that $Cost(A \prec B) = \Theta(m^3)$. Let $Cost_0(J)$ denote the expected cost for scheduling job set J from time 0 on using m machines. Furthermore, let $Delay(A,B) := Cost(A \prec$

*Massachussetts Institute of Technology, Cambridge, US. Visitor at Technische Universität Berlin, Germany.
†Department of Mathematics, Technical University of Berlin, Germany.

$B) - Cost_0(B)$ be the expected delay cost that the job set A incurs on job set B (which is deterministic in our case). Then

$$Cost(A \prec B) = Cost_0(A) + Cost_0(B) + Delay(A, B). \tag{1}$$

It is clear that

$$Cost_0(A) = m \times \frac{\frac{n_A}{m}(\frac{n_A}{m} + 1)}{2} L_A = \Theta(m^2),$$

and

$$Delay(A, B) = \Theta(n_B \times \frac{n_A}{m} L_A) = \Theta(m^3).$$

We analyze $Cost_0(B)$ by conditioning on the number of B-jobs that have a non-zero processing time. Fix a realization of processing times and let X_i be a binary random variable that is 1 if and only if the i-th B-job is long. And let $X = \sum_{i=1}^{n_B} X_i$ be the total number of long B-jobs in this realization. Obviously

$$\mathbb{E}\left[\, Cost_0(B) | X < m \,\right] \leq m^{2+\epsilon}.$$

Now, we bound $\Pr\left[X \geq m\right]$ using Chernoff's inequality. Given iid Bernoulli random variables $\{X_i\}_{i=1}^n$, then for each $\delta \in (0, 1]$, it states that

$$\Pr\left[X \geq (1 + \delta)\mathbb{E}\left[\, X \,\right]\right] \leq e^{-\delta^2 \mathbb{E}[\, X \,]/3}.$$

The expected number of long B-jobs is $\mathbb{E}\left[\, X \,\right] = m/4$. Now choosing $\delta = 1$ yields $\Pr\left[X \geq m - 1\right] \leq \Pr\left[X \geq m/2\right] \leq e^{-m/4}$, which becomes arbitrarily small for sufficiently large m.

Thus,

$$\begin{aligned} Cost_0(B) &= \Pr\left[X < m\right] \mathbb{E}\left[\, Cost_0(B) | X < m \,\right] + \Pr\left[X \geq m\right] \mathbb{E}\left[\, Cost_0(B) | X \geq m \,\right] \\ &= O(m^{2+\epsilon}). \end{aligned}$$

Since $\epsilon \in (0, 1)$, these bounds applied to (1) show that the expected total cost of WSEPT is

$$Cost(A \prec B) = \Theta(m^3).$$

Next, we show that $Cost(B \prec A) = O(m^{2+\epsilon})$. Recall from previous calculation that $Cost_0(B) = O(m^{2+\epsilon})$. In particular, the analysis shows that with high probability, at least $m/2$ machines will still be empty after assigning all B-jobs. The cost of scheduling all the A-jobs can be upper bounded by the cost of scheduling them on $m/2$ machines, which is equal to $\Theta(m^2)$. Altogether, a gap of $\Omega(m) = \Omega(\Delta^{1/2})$ is established. □

References

[1] Benjamin Labonté. Ein Simulationssystem für stochastische Scheduling-Probleme und empirische Untersuchung zur Approximationsgüte von Politiken. Master's thesis, Technische Universität Berlin, 2013.

[2] Nicole Megow, Marc Uetz, and Tjark Vredeveld. Models and algorithms for stochastic online scheduling. *Math. Oper. Res.*, 31(3):513–525, 2006.

[3] R.H. Möhring, A.S. Schulz, and M. Uetz. Approximation in stochastic scheduling: The power of LP-based priority policies. *Journal of the ACM*, 46(6):924–942, 1999.

[4] Michael H. Rothkopf. Scheduling with random service times. *Management Science*, 12(9):707–713, 1966.

[5] Andreas S. Schulz. Stochastic online scheduling revisited. In *COCOA*, volume 5165 of *Lecture Notes in Computer Science*, pages 448–457, Berlin, 2008. Springer.

[6] W. E. Smith. Various optimizers for single–stage production. *Naval Research Logistics Quarterly*, 3(1-2):59–66, 6 1956.