Approximations to Stochastic Dynamic Programs via Information Relaxation Duality

Online Appendix

Santiago R. Balseiro
Graduate School of Business
Columbia University
srb2155@columbia.edu

David B. Brown
Fuqua School of Business
Duke University
dbbrown@duke.edu

May 22, 2018

B Effective value formulation for stochastic knapsack

In this section we show that all our results extend to the effective value formulation of Dean et al. (2008) in which $w_i = v_i \mathbb{P}\{s_i \leq \kappa\}$ denotes the *effective value* of item i and $\mu_i = \mathbb{E}[\min\{s_i, \kappa\}]$ denotes the mean truncated size of item i. This formulation takes into account the fact that, in the event that $s_i > \kappa$, the actual realization of the size is irrelevant because item i certainly overflows the knapsack, and the DM will never collect the item's value in this case. Let $\tilde{s}_i = \tilde{s}_i$ denote the truncated size of item i. The corresponding approximation to the continuation value is

$$Q_t(c, i, \tilde{s}_i) = w_i + \frac{w_i}{\mu_i}(c - \tilde{s}_i)\,, \tag{B-1}$$

Because the expected continuation value is $\mathbb{E}_{\tilde{s}_i}\left[Q_t(c, i, \tilde{s}_i)\right] = w_i/\mu_i c$, equation (3) suggests that the greedy policy induced by (B-1) sorts the items in decreasing order of effective value per expected size, w_i/μ_i, and inserts items in this order until the knapsack overflows or no items remain. We refer to this policy as the *effective-value* greedy policy. Without loss of generality we assume that items are sorted in decreasing order of this ratio, i.e., $w_1/\mu_1 \geq w_2/\mu_2 \geq \ldots \geq w_I/\mu_I$. The expected performance of this greedy policy is given by

$$V^{\tilde{\mathrm{G}}} \triangleq \mathbb{E}\left[\sum_{t=1}^{I \wedge (\tau^{\tilde{\mathrm{G}}}-1)} v_t\right],$$

where $\tau^{\tilde{\mathrm{G}}}$ is the first time that capacity overflows under the effective-value greedy policy.

Before proceeding, we first discuss a variation of the problem that will be helpful. Specifically, the effective greedy policy ranks items using their effective values, so it will be useful for us to work with a variation of the problem in which the values v_i are replaced by the effective value w_i. In this variation, we also need to include the value of the overflowing item. This leads us to the formulation

$$W^* = \max_{\alpha \in \mathcal{A}} \mathbb{E}\left[\sum_{t=1}^{I \wedge \tau^\alpha} w_{\alpha_t}\right],$$

where the stopping time τ^α is defined as before. We first show that this formulation provides an upper bound.

Proposition B.1. $V^* \leq W^*$.

Proof. Let $\mathcal{S}^\alpha$ be the (stochastic) set of items that policy α attempts to insert into the knapsack and let c_i^α be the capacity remaining before policy α attempts to insert an item $i \in \mathcal{S}^\alpha$ into the knapsack. For any $\alpha \in \mathcal{A}$,

$$V^\alpha = \mathbb{E}\left[\sum_{t=1}^{I\wedge(\tau^\alpha-1)} v_{\alpha_t}\right] = \sum_{i=1}^{I} v_i \mathbb{P}\{i \in \mathcal{S}^\alpha, s_i \leq c_i^\alpha\}$$
$$\leq \sum_{i=1}^{I} v_i \mathbb{P}\{i \in \mathcal{S}^\alpha, s_i \leq \kappa\} = \sum_{i=1}^{I} w_i \mathbb{P}\{i \in \mathcal{S}^\alpha\} = \mathbb{E}\left[\sum_{t=1}^{I\wedge\tau^\alpha} w_{\alpha_t}\right] = W^\alpha\,,$$

where the inequality follows because $s_i \leq c_i^\alpha$ implies $s_i \leq \kappa$ since $c_i^\alpha \leq \kappa$. The third equality follows because the events $i \in \mathcal{S}^\alpha$ and $s_i \leq \kappa$ are independent (for nonanticipative policies, recall that the size is not revealed until after an item is selected), and because $w_i = v_i \mathbb{P}\{s_i \leq \kappa\}$. Since this holds for all $\alpha \in \mathcal{A}$, this variation provides an upper bound on the optimal value in original formulation, i.e., $V^* \leq W^*$. □

We now return to the penalty, which we apply to this effective value formulation. Equation (6) implies that the penalty induced by (B-1) is $z_t(c, i, \tilde{s}_i) = w_i/\mu_i(\mu_i - \tilde{s}_i)$. Using the effective value formulation, we can write the problem with penalties included as

$$W_z^* = \max_{\alpha\in\mathcal{A}} \mathbb{E}\left[\sum_{t=1}^{I\wedge\tau^\alpha} w_{\alpha_t} + z_{\alpha_t}(\tilde{s}_{\alpha_t} - \mu_{\alpha_t})\right].$$

Because the penalty is dual feasible as described above, $W_z^* = W^*$. We let $W_z^{\mathrm{P}}(\mathbf{s})$ denote the optimal (deterministic) value of the penalized perfect information problem for sample path $\mathbf{s} \in \mathbb{R}_+^I$. Since the set of perfect information policies includes the set $\mathcal{A}$ of nonanticipative policies, and the penalty is dual feasible, we obtain an upper bound $W^* \leq W_z^{\mathrm{P}}$, where $W_z^{\mathrm{P}} = \mathbb{E}_{\mathbf{s}}[W_z^{\mathrm{P}}(\mathbf{s})]$ denotes the penalized perfect information bound. Below we discuss how to calculate $W_z^{\mathrm{P}}(\mathbf{s})$ by solving an integer program that includes additional variables representing which item, if any, overflows the knapsack.

It is instructive to see how this works on the example from Dean et al. (2008) as discussed in Section 3.2, with I symmetric items of value one and sizes that are Bernoulli with probability 1/2, scaled by $\kappa + \varepsilon$. Recall that a greedy policy is (trivially) optimal and the optimal value is $V^* = 1 - (1/2)^I$, but the perfect information bound without penalty provides the poor bound of $V^{\mathrm{P}} = I/2$. In this example, $w_i = 1/2$, $\mu_i = \kappa/2$, and $\tilde{s}_i$ is either 0 or κ, each with probability 1/2. In the penalized perfect information problem, the value for selecting an item is $(w_i/\mu_i)\tilde{s}_i$, which is 0 if $\tilde{s}_i = 0$ and 1 if $\tilde{s}_i = \kappa$: in particular, any items with realized sizes of zero provide zero value as well. Moreover, we can select at most one item with realized positive size of $\kappa + \epsilon$ - in particular, an item that overflows the knapsack. Thus, $W_z^{\mathrm{P}}(\mathbf{s}) = 1$ if $s_i > 0$ for any i, and $W_z^{\mathrm{P}}(\mathbf{s}) = 0$ otherwise. Because $\mathbb{P}\{s_i = 0 \;\forall\; i\} = (1/2)^I$, the penalized perfect information bound then is

$$W_z^{\mathrm{P}} \;=\; 1 - (1/2)^I \;=\; V^*,$$

i.e., we recover a tight bound for all values of I.

In the following, we denote the penalized performance of the effective-value greedy policy by

$$V_z^{\tilde{\mathrm{G}}}(\mathbf{s}) = \sum_{i=1}^{I\wedge(\tau^{\tilde{\mathrm{G}}}-1)} v_i + \sum_{i=1}^{I\wedge\tau^{\tilde{\mathrm{G}}}} w_i/\mu_i(\tilde{s}_i - \mu_i)\,.$$

We now provide an upper bound on the penalized perfect information value in terms of the penalized performance of the effective-value greedy policy.

Proposition B.2. *For every sample path* $\mathbf{s}$*, the performance of the effective value greedy policy satisfies:*

$$W_z^{\mathrm{P}}(\mathbf{s}) - V_z^{\mathrm{G}}(\mathbf{s}) \leq \max_{i\in\mathcal{I}} w_i + \max_{i\in\mathcal{I}} w_i/\mu_i\tilde{s}_i\,.$$

Proof. Let $u_i = w_i/\mu_i$. Relaxing constraints (B-2b) and (B-2d) we obtain that problem (B-2) decouples in terms of the decision variables $\mathbf{x}$ and $\mathbf{y}$. Thus, we obtain the upper bound

$$\bar{W}_z^{\mathrm{P}}(\mathbf{s}) \leq \underbrace{\max_{\mathbf{x}\in\{0,1\}^I} \sum_{i=1}^I u_i \tilde{s}_i x_i \quad \text{s.t.} \quad \sum_{i=1}^I \tilde{s}_i x_i \leq 1\,.}_{\clubsuit} + \underbrace{\max_{\mathbf{y}\in\{0,1\}^I} \sum_{i=1}^I u_i \tilde{s}_i y_i \quad \text{s.t.} \quad \sum_{i=1}^I y_i \leq 1\,.}_{\spadesuit},$$

where we also relaxed the knapsack constraint (B-2a) to $\sum_{i=1}^I \tilde{s}_i x_i \leq 1$ because $\tilde{s}_i \leq s_i$. We conclude the proof by bounding each term at a time.

For the first problem note that the ratio of value to (truncated) size of each item is u_i, as in the greedy policy. By considering the continuous relaxation to $x_i \in [0,1]$, we obtain that $x_i = 1$ for $i \leq \tau^{\tilde{\mathrm{G}}}$ and $x_i \in (0,1]$ for $i = \tau^{\tilde{\mathrm{G}}}$ whenever $\tau^{\tilde{\mathrm{G}}} \leq I$ (and $x_i = 1$ for all i when $\tau^{\tilde{\mathrm{G}}} > I$). Rounding up to one the last fractional item and using that $u_i\tilde{s}_i = w_i + u_i(\tilde{s}_i - \mu_i)$, we obtain the upper bound

$$\clubsuit \leq \sum_{i=1}^{I\wedge\tau^{\tilde{\mathrm{G}}}} u_i\tilde{s}_i = \sum_{i=1}^{I\wedge(\tau^{\tilde{\mathrm{G}}}-1)} w_i + \sum_{i=1}^{I\wedge\tau^{\tilde{\mathrm{G}}}} u_i(\tilde{s}_i - \mu_i) + w_{\tau^{\tilde{\mathrm{G}}}}\mathbf{1}\{\tau^{\tilde{\mathrm{G}}} \leq I\} \leq V_z^{\mathrm{G}}(\mathbf{s}) + \max_{i\in\mathcal{I}} w_i\,,$$

where the last inequality follows from $w_i = v_i\mathbb{P}\{s_i \leq \kappa\} \leq v_i$ and $w_{\tau^{\mathrm{G}}} \leq \max_{i\in\mathcal{I}} w_i$.

For the second problem note that the optimal solution selects the item with maximum objective and thus

$$\spadesuit = \max_{i\in\mathcal{I}} u_i\tilde{s}_i\,.$$

The result then follows. □

Proposition B.2 implies the following chain of inequalities

$$V^{\mathrm{G}} \leq V^* \leq W^* \leq W_z^{\mathrm{P}} \leq V^{\mathrm{G}} + \max_i w_i + \mathbb{E}[\max_i w_i\tilde{s}_i/\mu_i]\,.$$

The first inequality in the chain follow trivially, the second inequality was argued previously and the third inequality follows because the penalty $\mathbf{z}$ is dual feasible.

Integer programming formulation for $W_z^{\mathrm{P}}(\mathbf{s})$. To calculate $W_z^{\mathrm{P}}(\mathbf{s})$, because the item that overflows the knapsack now counts towards the objective, we need to explicitly account for the overflowing item, whenever it exists. We obtain an upper bound on the penalized perfect information problem for a fixed sample path $\mathbf{s}$ by solving the integer programming problem

$$\begin{aligned}
\bar{W}_z^{\mathrm{P}}(\mathbf{s}) \triangleq \max_{\mathbf{x},\mathbf{y}\in\{0,1\}^I} & \sum_{i=1}^I (w_i + u_i(\tilde{s}_i - \mu_i))(x_i + y_i) \\
\text{s.t.} & \sum_{i=1}^I s_i x_i \leq \kappa\,, && \text{(B-2a)}\\
& x_i + y_i \leq 1\,, \quad \forall i \in \mathcal{I}\,, && \text{(B-2b)}\\
& \sum_{i=1}^I y_i \leq 1\,, && \text{(B-2c)}\\
& \sum_{i=1}^I s_i(x_i + y_i) \geq \kappa(1 - x_i)\,, \quad \forall i \in \mathcal{I}\,. && \text{(B-2d)}
\end{aligned}$$

where $x_i \in \{0,1\}$ indicates whether the item is selected and fits the knapsack, and $y_i \in \{0,1\}$ indicates if the item overflows the knapsack. Constraint (B-2b) imposes that an item either fits the knapsack or overflows it. Constraint (B-2c) guarantees that there is at most one overflowing item. Constraint (B-2d) requires the overflowing item, if one exists, causes the selected capacity to exceed the capacity of knapsack. This

constraint is vacuous when all items fit the knapsack, i.e., if there is no overflow. Note that we can only be sure that $W_z^{\mathrm{P}}(\mathbf{s}) \leq \bar{W}_z^{\mathrm{P}}(\mathbf{s})$, because the "overflowing" item y_i can be chosen to exactly match the capacity of the knapsack. In order for item y_i to actually overflow the knapsack we need to make inequality (B-2d) strict. When the distribution of item sizes are absolutely continuous or lattice (i.e., there exists some $h > 0$ such that $\mathbb{P}\{s_i \in \{0, h, 2h, \ldots\}\} = 1$ for all $i \in \mathcal{I}$), replacing constraint (B-2d) by $\sum_{i=1}^{I} s_i(x_i + y_i) \geq (\kappa + \epsilon)(1 - x_i)$ for some $\epsilon > 0$ in problem (B-2) gives that $\bar{W}_z^{\mathrm{P}}(\mathbf{s}) = W_z^{\mathrm{P}}(\mathbf{s})$. The bound given by $\bar{W}_z^{\mathrm{P}}(\mathbf{s})$, however, suffices for our analysis.

C Alternative performance guarantee for stochastic scheduling

In this section, we show how to obtain Corollary 4.1 from Möhring et al. (1999) using penalized perfect information analysis. For a given sample path of processing times $\mathbf{p}$ and scheduling decisions $\mathbf{a}$, we consider the penalty

$$z(\mathbf{a}) = \sum_{j \in \mathcal{J}} r_j S_j^{\mathbf{a}}(p_j - \mathbb{E}[p_j]) \ , \tag{C-3}$$

where $S_j^{\mathbf{a}}$ denotes the start time of job j using $\mathbf{a}$. Dual feasibility of this penalty follows because, for any nonanticipative policy, the start time of a job is independent of the job's processing time.

Proposition C.1. *The performance of the WSEPT policy satisfies*

$$V^{\mathrm{G}} - V^* \leq V^{\mathrm{G}} - V_z^{\mathrm{P}} \leq \frac{M-1}{2M}(\rho + 1) \sum_{j \in \mathcal{J}} w_j \mathbb{E}[p_j] \ ,$$

with the penalty z in (C-3) and ρ is an upper bound on the squared coefficient of variation of all jobs.

Proof. The only inequality we need to show is the last inequality in (C-3), which we argue using three steps. First, we upper bound the performance of the WSEPT policy. Second, we lower bound the objective value of the penalized perfect information problem for a fixed realization of processing times. Finally, we combine the previous bounds to bound the performance of the WSEPT policy in terms of the *expected* penalized perfect information bound.

Step 1. We first upper bound the performance of the WSEPT policy. For any list scheduling policy α that sequences jobs according to $\alpha_1, \alpha_2, \ldots, \alpha_J$, it is well known that the completion time of job $C_{\alpha_j}^{\alpha}$ satisfies that

$$C_{\alpha_j}^{\alpha} \leq \frac{1}{M} \sum_{\ell=1}^{j-1} p_{\alpha_\ell} + p_{\alpha_j} \ .$$

See for example, Lemma 3.3 of Hall et al. (1997) for a proof. As a result we obtain that the performance of the WSEPT policy satisfies that

$$V^{\mathrm{G}} = \mathbb{E}\left[\sum_{j=1}^{J} w_j C_j^{\mathrm{G}}\right] \leq \mathbb{E}\left[\sum_{j=1}^{J} w_j \left(p_j + \frac{1}{M}\sum_{i=1}^{j-1} p_i\right)\right] = \sum_{j=1}^{J} w_j \mathbb{E}\left[p_j\right] + \frac{1}{M}\sum_{j=1}^{J}\sum_{i=1}^{j-1} w_j \mathbb{E}\left[p_i\right] \ , \tag{C-4}$$

where the inequality follows from the fact that WSEPT sequences jobs according to $1, \ldots, n$ because jobs are assumed to be sorted in decreasing order w.r.t the ratio of weight to expected processing time $r_i = w_i/\mathbb{E}[p_i]$.

Step 2. We next lower bound the objective value of the penalized perfect information problem for a fixed realization of processing times. Using the penalty in (C-3) and the fact that the completion times $C_j^{\mathbf{a}}$ satisfy $C_j^{\mathbf{a}} = S_j^{\mathbf{a}} + p_j$, the perfect information problem can be lower bounded by

$$V_z^{\mathrm{P}}(\mathbf{p}) \geq \underline{V}_z^{\mathrm{P}}(\mathbf{p}) + \sum_{j=1}^{J} w_j p_j - r_j p_j^2 \, , \tag{C-5}$$

where $\underline{V}_z^{\mathrm{P}}(\mathbf{p})$ is the objective value of the deterministic scheduling problem $PM//\sum_j w_j^z C_j$ with weights $w_j^z = w_j + r_j(p_j - \mathbb{E}[p_j]) = r_j p_j$. Using a know result from deterministic parallel machine scheduling, (Lemma C.2), we obtain the following lower bound on the objective value of the previous scheduling problem

$$\underline{V}_z^{\mathrm{P}}(\mathbf{p}) \geq \sum_{j=1}^{J} r_j \left(\frac{1}{M} p_j \sum_{i=1}^{j-1} p_i + \frac{M+1}{2M} p_j^2 \right). \tag{C-6}$$

Step 3. We conclude by combining our previous results to bound the performance of the WSEPT policy in terms of the *expected* penalized perfect information bound. Taking expectations w.r.t. the sample path $\mathbf{p}$ we obtain that

$$\begin{aligned}
V_z^{\mathrm{P}} &= \mathbb{E}_{\mathbf{p}}\left[V_z^{\mathrm{P}}(\mathbf{p})\right] \geq \mathbb{E}_{\mathbf{p}}[\underline{V}_z^{\mathrm{P}}(\mathbf{p})] + \sum_{j=1}^{J} w_j \mathbb{E}[p_j] - r_j \mathbb{E}[p_j^2] \\
&\geq \sum_{j=1}^{J} w_j \mathbb{E}[p_j] + \frac{1}{M} \sum_{j=1}^{J} r_j \mathbb{E}\left[p_j \sum_{i=1}^{j-1} p_i \right] - \frac{M-1}{2M} \sum_{j=1}^{J} r_j \mathbb{E}[p_j^2] \\
&= \sum_{j=1}^{J} w_j \mathbb{E}[p_j] + \frac{1}{M} \sum_{j=1}^{J} \sum_{i=1}^{j-1} w_j \mathbb{E}\left[p_i\right] - \frac{M-1}{2M} \sum_{j=1}^{J} w_j \frac{\mathbb{E}[p_j^2]}{\mathbb{E}[p_j]} \\
&\geq V^{\mathrm{G}} - \frac{M-1}{2M} \sum_{j=1}^{J} w_j \frac{\mathbb{E}[p_j^2]}{\mathbb{E}[p_j]} \geq V^{\mathrm{G}} - \frac{M-1}{2M} (\rho+1) \sum_{j=1}^{J} w_j \mathbb{E}[p_j],
\end{aligned}$$

where the first inequality follows from (C-5); the second inequality from (C-6); the second equation from the fact that $r_j = w_j/\mathbb{E}[p_j]$ and using that the processing times are independent; the third inequality from the bound on the performance of the WSEPT policy given in (C-4); and the last inequality because $\mathbb{E}[p_j^2]/\mathbb{E}[p_j]^2 = \mathrm{Var}[p_j]/\mathbb{E}[p_j]^2 + 1 \leq \rho + 1$. □

The following result is a strengthened version of Lemma 3.2 from Hall et al. (1997) and provides a lower bound on the objective value of the deterministic scheduling problem $PM//\sum_j r_j p_j C_j$. We reproduce the result for the sake of completeness.

Lemma C.2. *The objective value of the deterministic scheduling problem $PM//\sum_j r_j p_j C_j$, denoted by $\underline{V}_z^{\mathrm{P}}(\mathbf{p})$, is lower bounded by*

$$\underline{V}_z^{\mathrm{P}}(\mathbf{p}) \geq \sum_{j=1}^{J} r_j \left(\frac{1}{M} p_j \sum_{i=1}^{j-1} p_i + \frac{M+1}{2M} p_j^2 \right).$$

Proof. A lower bound on the objective value can be obtained from the observation that for any feasible schedule on M machines the completion times should satisfy the following inequalities

$$\sum_{j \in A} p_j C_j \geq \frac{1}{2M} \left(\sum_{j \in A} p_j \right)^2 + \frac{1}{2} \sum_{j \in A} p_j^2,$$

for every subset $A \in \mathcal{J}$ (see, e.g., Hall et al. (1997)). Lemma 3.2 from Hall et al. (1997) proves a similar result under a weaker class of valid inequalities.

Optimizing over the completion times we obtain that the latter deterministic scheduling problem can be lower bounded by the following linear program

$$\begin{aligned}
\underline{V}_z^{\mathrm{P}}(\mathbf{p}) \geq \min_{\mathbf{C} \in \mathbb{R}^J} & \sum_{j=1}^{J} r_j p_j C_j \\
& \text{s.t.} \sum_{j \in A} p_j C_j \geq f(A), \quad \forall A \subseteq \mathcal{J},
\end{aligned} \tag{C-7}$$

$$C_j \geq 0\,,$$

where the set function $f : 2^{\mathcal{J}} \to R$ is given by $f(A) = \frac{1}{2M}p(A)^2 + \frac{1}{2}p^2(A)$, where we denote by $p(A) = \sum_{j\in A} p_j$ and $p^2(A) = \sum_{j\in A} p_j^2$. It is not hard to see that the set function is super-modular, that is, for every $k, \ell \notin A$ we have that

$$f(A+k) + f(A+\ell) = f(A) + f(A+k+\ell) - 2p_k p_\ell \leq f(A) + f(A+k+\ell)\,.$$

Setting $y_j = p_j C_j$ we obtain that the feasible set of problem (C-7) is a polymatroid. Because the objective is linear in y_j, we can apply the greedy algorithm of Edmonds to characterize its optimal solution. Since the objective's coefficients satisfy $r_1 \geq r_2 \geq \ldots \geq r_J$, we obtain that the optimal solution is given by $y_j^* = f(\{1,\ldots,j\}) - f(\{1,\ldots,j-1\})$, or equivalently that

$$\begin{aligned} y_j^* &= f(\{1,\ldots,j\}) - f(\{1,\ldots,j-1\}) \\ &= \frac{1}{2M}\left(p(\{1,\ldots,j-1\}) + p_j\right)^2 - \frac{1}{2M}p(\{1,\ldots,j-1\})^2 + \frac{1}{2}p_j^2 \\ &= \frac{1}{M}p_j p(\{1,\ldots,j-1\}) + \frac{M+1}{2M}p_j^2\,. \end{aligned}$$

As a result the objective value is given by $\sum_{j=1}^J r_j y_j^*$, implying that

$$\underline{V}_z^{\mathrm{P}}(\mathbf{p}) \geq \sum_{j=1}^J r_j y_j^* = \sum_{j=1}^J r_j \left(\frac{1}{M}p_j \sum_{i=1}^{j-1} p_i + \frac{M+1}{2M}p_j^2\right)\,. \qquad \square$$

D Sequential search with large capacity

To obtain reservation prices, we consider a Lagrangian relaxation in which we relax the constraint that the decision maker can select at most K alternatives. Introducing a dual variable $\lambda \geq 0$ for the capacity constraint $\sum_{t=1}^{N\wedge\tau^\alpha} |\mathcal{S}_t^\alpha| \leq K$ we obtain the Lagrangian:

$$\mathcal{L}^\alpha(\lambda) = \lambda K + \mathbb{E}\left[\sum_{t=1}^{N\wedge\tau^\alpha}\left(\sum_{r\in\mathcal{S}_t^\alpha}\delta^t r - \lambda\right) - \delta^{t-1}s\right]\,.$$

Note that, in the absence of a capacity constraint, every explored alternative with "adjusted" reward $\delta^t r - \lambda \geq 0$ should be immediately selected. Using the fact that alternatives are a priori identical we obtain that the dual function is given by:

$$\begin{aligned} \Psi(\lambda) = \max_{\alpha\in\mathcal{A}} \mathcal{L}^\alpha(\lambda) &= \lambda K + \max_{\alpha\in\mathcal{A}} \mathbb{E}\left[\sum_{t=1}^{N\wedge\tau^\alpha}\left(\delta^t r_t - \lambda\right)^+ - \delta^{t-1}s\right] \\ &= \lambda K + \sum_{t=1}^N \delta^{t-1}\left(\delta\mathbb{E}_{\tilde{r}}\left[\left(\tilde{r} - \lambda\delta^{-t}\right)^+\right] - s\right)^+\,, \end{aligned} \tag{D-8}$$

where the last equality follows because DM should continue exploring alternatives while the expected profit of exploring is positive, i.e., $\delta\mathbb{E}\left[(r_t - \lambda\delta^{-t})^+\right] - s > 0$, since the expected profit is non-increasing with time. Weak duality implies that $V^* \leq \Psi(\lambda)$ for all $\lambda \geq 0$. The previous derivation suggests time-dependent reservation prices $v_t = \lambda^*\delta^{-t}$ with $\lambda^* \in \arg\min_{\lambda\geq 0}\Psi(\lambda)$, i.e., the Lagrange multiplier is chosen to minimize the upper bound. Because the dual function is convex, the best Lagrange multiplier can be computed efficiently.

The following result compares, for every sample path, the penalized performance of the greedy policy to the performance of the penalized perfect information problem. In the following results we adopt the convention that $\sum_{i=a}^b x_i = 0$ if $b < a$.

Proposition D.1. *For every sample path* $\mathbf{r}$ *and* $\lambda \geq 0$*, the penalized performance of the greedy policy with* $v_t = \lambda\delta^{-t}$ *satisfies:*

$$V_z^{\mathrm{P}}(\mathbf{r}) - V_z^{\mathrm{G}}(\mathbf{r}) \leq \lambda \bar{k}^{\mathrm{G}} + \sum_{t=1}^{N(\lambda)} \delta^{t-1}\pi_t(\lambda) - \sum_{t=1}^{N\wedge\tau^{\mathrm{G}}} \delta^{t-1}\pi_t(\lambda)\,, \tag{D-9}$$

where $\pi_t(\lambda) = \delta\mathbb{E}_{\tilde{r}}\left[(\tilde{r} - \lambda\delta^{-t})^+\right] - s$ *is the adjusted expected payoff at time* t *and* $N(\lambda) = \sup\left\{t \in \{1,\dots,N\} : \pi_t(\lambda) \geq 0\right\}$ *is the (deterministic) last time period in which it is profitable to explore an alternative in the Lagrangian relaxation.*

Although we use a Lagrangian relaxation in this derivation, as we show in the proof below, the penalized perfect information value is tighter (smaller) than $\Psi(\lambda)$ for every sample path $\mathbf{r}$ for all $\lambda \geq 0$.

Proof. We first analyze the penalized perfect information problem. Recall that, because of discounting, the DM with penalized perfect information never delays the selection of an explored alternative. Using (30), we can write the penalized payoff of the alternative explored and selected at time t as

$$\delta r_t - s + z_t = \delta \min(r_t, v_t) + \delta\mathbb{E}_{\tilde{r}}[(\tilde{r} - v_t)^+] - s \leq \delta v_t + \delta\mathbb{E}_{\tilde{r}}[(\tilde{r} - v_t)^+] - s\,,$$

where we used that $\min(r_t, v_t) \leq r_t$. Similarly, we can write the penalized payoff of an alternative explored, but not selected, at time t as

$$-s + z_t = \delta \min(r_t, v_t) - \delta r_t + \delta\mathbb{E}_{\tilde{r}}[(\tilde{r} - v_t)^+] - s \leq \delta\mathbb{E}_{\tilde{r}}[(\tilde{r} - v_t)^+] - s\,,$$

where we used that $\min(r_t, v_t) \leq v_t$. Therefore, for every sample path $\mathbf{r}$, we can upper bound $V_z^{\mathrm{P}}(\mathbf{r})$ by the objective value of an alternative problem in which the reward of selecting an alternative is $\hat{r}_t = v_t$ and the cost of exploring an alternative is $\hat{s}_t = s - \delta\mathbb{E}_{\tilde{r}}[(\tilde{r} - v_t)^+]$. Because this adjusted cost of exploring an alternative can be negative for some periods, the DM may choose to explore but not select alternatives. Therefore, the total cost incurred from exploring alternatives is at most

$$-\sum_{t=1}^{N} \min(\hat{s}_t, 0) = \sum_{t=1}^{N} \delta^{t-1}\left(\delta\mathbb{E}_{\tilde{r}}\left[(\tilde{r} - v_t)^+\right] - s\right)^+ .$$

Letting $x_t \in \{0,1\}$ indicate whether an alternative is selected at time t, we can bound the total rewards from selected alternatives by

$$\sum_{t=1}^{N} \delta^t \hat{r}_t x_t = \lambda \sum_{t=1}^{N} x_t \leq \lambda K\,.$$

where we used that $r_t = v_t = \lambda\delta^{-t}$ and $\sum_{t=1}^{N} x_t \leq K$ from the capacity constraint. Combining both bounds and using (D-8) we conclude that $V_z^{\mathrm{P}}(\mathbf{r}) \leq \Psi(\lambda)$ and thus $V_z^{\mathrm{P}}(\mathbf{r})$ is tighter than $\Psi(\lambda)$ in every sample path.

We next analyze the penalized performance of the greedy policy. Using that the greedy policy selects an alternative whenever $r_t \geq v_t$, we obtain that the penalized payoff at time t is given by

$$\delta r_t \mathbf{1}\{r_t \geq v_t\} - s + z_t = \delta v_t \mathbf{1}\{r_t \geq v_t\} + \delta\mathbb{E}_{\tilde{r}}\left[(\tilde{r} - v_t)^+\right] - s = \delta v_t \mathbf{1}\{r_t \geq v_t\} + \pi_t(\lambda)\,.$$

Since the reservation prices v_t are nondecreasing in t, we obtain, by discarding the rewards of alternatives potentially recalled at time $t = N$, that the performance of the greedy policy is

$$\begin{aligned} V_z^{\mathrm{G}}(\mathbf{r}) &\geq \sum_{t=1}^{N\wedge\tau^{\mathrm{G}}} \delta^t v_t \mathbf{1}\{r_t \geq v_t\} + \sum_{t=1}^{N\wedge\tau^{\mathrm{G}}} \delta^{t-1}\pi_t(\lambda) = \lambda \sum_{t=1}^{N\wedge\tau^{\mathrm{G}}} \mathbf{1}\{r_t \geq v_t\} + \sum_{t=1}^{N\wedge\tau^{\mathrm{G}}} \delta^{t-1}\pi_t(\lambda) \\ &= \Psi(\lambda) - \lambda\bar{k}^{\mathrm{G}} + \sum_{t=1}^{N\wedge\tau^{\mathrm{G}}} \delta^{t-1}\pi_t(\lambda) - \sum_{t=1}^{N(\lambda)} \delta^{t-1}\pi_t(\lambda)\,, \end{aligned}$$

where the second equality follows from $v_t = \lambda\delta^{-t}$ and the last because $\sum_{t=1}^{N\wedge\tau^{\mathrm{G}}} \mathbf{1}\{r_t \geq v_t\} = K - \bar{k}^{\mathrm{G}}$ is the number of alternatives selected. □

Unlike the optimal policy in the Lagrangian relaxation, the greedy policy continues exploring alternatives even when the expected payoff of exploring an alternative is negative. A stronger performance guarantee can be provided for a modified greedy policy that stops exploring at time $N(\lambda)$ and, at that point, selects all available alternatives that were not previously collected. Both policies turn out to be asymptotically optimal in the regime we consider.

To simplify exposition, we will use the following assumption in the asymptotic result that follows; the boundedness assumption can be relaxed (e.g., by instead assuming boundedness of appropriate tail moments) at the expense of additional notation and lengthier arguments.

Assumption D.2. *We have $\delta \in (0,1)$, $\delta\mathbb{E}[\tilde{r}] > s$, and the distribution of rewards is absolutely continuous and has bounded support $[0,\bar{x}]$ with $\bar{x} < \infty$.*

Taking expectations in (D-9) and using the duality results of Section 2 we can provide the following guarantees on the performance of the greedy policy.

Corollary D.3. *The performance of the greedy policy satisfies:*

(i) **Performance guarantee.** *For every $\lambda \geq 0$,*

$$V^* - V^{\mathrm{G}} \leq V_z^{\mathrm{P}} - V^{\mathrm{G}} \leq \mathbb{E}\left[\lambda\bar{k}^{\mathrm{G}} + \sum_{t=1}^{N(\lambda)} \delta^{t-1}\pi_t(\lambda) - \sum_{t=1}^{N\wedge\tau^{\mathrm{G}}} \delta^{t-1}\pi_t(\lambda)\right].$$

(ii) **Asymptotic optimality.** *If $K = \Theta(N)$, $\limsup_{N\to\infty} (1-\delta)N < \infty$, and $\lambda^* \in \arg\min_{\lambda\geq 0} \Psi(\lambda)$, and Assumption D.2 holds, then*

$$\lim_{N\to\infty} \frac{1}{K}\left(V^* - V^{\mathrm{G}}\right) = 0\,.$$

Proof. Part (i) follows directly taking expectations of (D-9) and applying Proposition 2.1.

For part (ii), we use the following lemma, which establishes some useful properties of λ^*, an optimal solution to the Lagrangian dual problem.

Lemma D.4. *If Assumption D.2 holds, then $\lambda^* \in [0,\bar{x}]$, $\underline{p} \triangleq \min_{i=1}^{N(\lambda^*)} \mathbb{P}\{\tilde{r} \geq \lambda^*\delta^{-i}\} \geq s/\bar{x}$, and*

$$K \leq \sum_{t=1}^{N(\lambda^*)} \mathbb{P}\{\tilde{r} \geq \lambda^*\delta^{-t}\} \leq K+1.$$

Proof. We write $\pi_t(\lambda) = \mathbb{E}\left[(\tilde{r} - \lambda\delta^{-t})^+\right] - s = h(\lambda\delta^{-t}) - s$, where $h(x) = \mathbb{E}\left[(\tilde{r}-x)^+\right] = \int_x^{\bar{x}} \bar{F}(y)\mathrm{d}y$ is continuous and strictly decreasing for $x \in [0,\bar{x}]$. The left and right derivative of the dual function are $\partial_-\Psi(\lambda) = K - \sum_{t=1}^N \mathbb{P}\{\tilde{r} \geq \lambda\delta^{-t}\}\mathbf{1}\{\pi_t(\lambda) \geq 0\}$ and $\partial_+\Psi(\lambda) = K - \sum_{t=1}^N \mathbb{P}\{\tilde{r} > \lambda\delta^{-t}\}\mathbf{1}\{\pi_t(\lambda) > 0\}$, respectively.

First note that for $t \leq N(\lambda^*)$, we have $\pi_t(\lambda^*) \geq 0$. Because $(\tilde{x} - v_t)^+ \leq \bar{x}\{\tilde{x} \geq v_t\}$, we obtain that $\pi_t(\lambda) \leq \bar{x}\mathbb{P}\{\tilde{x} \geq v_t\} - s$ and thus $\mathbb{P}\{\tilde{x} \geq v_t\} \geq s/\bar{x}$. This implies that $\underline{p} \geq s/\bar{x}$. Because $\partial_-\Psi(\bar{x}) = K$, we obtain that $\lambda^* \leq \bar{x}$.

We first consider the case $K = N$. Note that $\partial_+\Psi(0) = K - N$ because $\delta\mathbb{E}[\tilde{r}] > s$ and rewards have no atom at zero. Because $\Psi(\lambda)$ is convex, we have that $\lambda^* = 0$. The result follows trivially.

We now consider the case when $1 \leq K < N$. Because $\partial_+\Psi(0) > 0$, we have that $\lambda^* > 0$. Since the dual function is convex, at an optimal solution, it holds that $\partial_+\Psi(\lambda^*) \geq 0$ and $\partial_-\Psi(\lambda^*) \leq 0$. Therefore, $\partial_-\Psi(\lambda^*) \leq 0$ implies that $K \leq \sum_{t=1}^{N(\lambda^*)} \mathbb{P}\{\tilde{r} \geq \lambda^*\delta^{-t}\}$because $\pi_t(\lambda^*) < 0$ for $t > N(\lambda^*)$. Since rewards are

absolutely continuous, we get that $\partial_+\Psi(\lambda^*) = K - \sum_{t=1}^{N} \mathbb{P}\{\tilde{r} \geq \lambda^*\delta^{-t}\}\mathbf{1}\{\pi_t(\lambda^*) > 0\}$. If $\pi_t(\lambda^*) = 0$ for $t = N(\lambda^*)$ we have that $\partial_+\Psi(\lambda^*) = K - \sum_{t=1}^{N(\lambda^*)-1} \mathbb{P}\{\tilde{r} \geq \lambda^*\delta^{-t}\}$ because $\pi_t(\lambda^*)$ is strictly decreasing in t for $t \leq N(\lambda^*)$ since $\delta \in (0,1)$, $\lambda^* > 0$, and $\underline{p} > 0$. Because $\partial_+\Psi(\lambda^*) \geq 0$, this implies that $\sum_{t=1}^{N(\lambda^*)} \mathbb{P}\{\tilde{r} \geq \lambda^*\delta^{-t}\} \leq K+1$ since the summands are at most one. If $\pi_t(\lambda^*) > 0$ for $t = N(\lambda^*)$, then $\pi_t(\lambda^*) < 0$ for $t = N(\lambda^*)+1$ by definition and $\partial_+\Psi(\lambda^*) = \partial_-\Psi(\lambda^*)$. The result follows. □

We now prove part (ii) of Proposition D.3. Note that for $t \leq N(\lambda)$ we have that $\pi_t(\lambda) \leq \mathbb{E}[\tilde{r}]$ because $\lambda \geq 0$, while for $t > N(\lambda)$ we have that $\pi_t(\lambda) \geq -s \geq -\delta\mathbb{E}[\tilde{r}]$ because $\delta\mathbb{E}[\tilde{r}] \geq s$ by assumption. Therefore, we can write the loss term as

$$\begin{aligned}\sum_{t=1}^{N(\lambda)} \delta^{t-1}\pi_t(\lambda) - \sum_{t=1}^{N\wedge\tau^{\mathrm{G}}} \delta^{t-1}\pi_t(\lambda) &= \sum_{t=N\wedge\tau^{\mathrm{G}}+1}^{N(\lambda)} \delta^{t-1}\pi_t(\lambda) - \sum_{t=N(\lambda)+1}^{N\wedge\tau^{\mathrm{G}}} \delta^{t-1}\pi_t(\lambda) \\ &\leq \mathbb{E}[\tilde{r}]\Bigg(\underbrace{\sum_{t=N\wedge\tau^{\mathrm{G}}+1}^{N(\lambda)} \delta^{t-1}}_{(I)} + \underbrace{\sum_{t=N(\lambda)+1}^{N\wedge\tau^{\mathrm{G}}} \delta^{t}}_{(II)}\Bigg).\end{aligned}$$

For the first term, by taking expectations over τ^{G}, we have that

$$\mathbb{E}\left[(I)\right] = \sum_{i=1}^{N(\lambda)-1} \mathbb{P}\{\tau^{\mathrm{G}} = i\} \sum_{t=i+1}^{N(\lambda)} \delta^{t-1} = \sum_{t=2}^{N(\lambda)} \delta^{t-1} \sum_{i=1}^{t-1} \mathbb{P}\{\tau^{\mathrm{G}} = i\} = \sum_{t=1}^{N(\lambda)-1} \delta^{t}\mathbb{P}\{\tau^{\mathrm{G}} \leq t\},$$

where the second equality follows from exchanging the order of summations. Set $\lambda = \lambda^*$ and $v_t = \lambda^*\delta^{-t}$. Using that $M_t = \sum_{i=1}^{t} \mathbf{1}\{r_i \geq v_i\} - \bar{F}(v_i)$ with $\bar{F}(x) = \mathbb{P}\{\tilde{r} \geq x\}$ is a martingale with differences bounded by $|M_{t+1} - M_t| \leq 1$, we obtain from Azuma's inequality that for $t \leq N(\lambda^*) - 1$:

$$\begin{aligned}\mathbb{P}\left\{\tau^{\mathrm{G}} \leq t\right\} &= \mathbb{P}\left\{\sum_{i=1}^{t} \mathbf{1}\{r_i \geq v_i\} \geq K\right\} = \mathbb{P}\left\{M_t \geq K - \sum_{i=1}^{t} \bar{F}(v_i)\right\} \leq \mathbb{P}\left\{M_t \geq \sum_{i=t+1}^{N(\lambda^*)} \bar{F}(v_i) - 1\right\} \\ &\leq \mathbb{P}\left\{M_t \geq (N(\lambda^*) - t - \underline{p}^{-1})\underline{p}\right\} \leq \exp\left(-(N(\lambda^*) - t - \underline{p}^{-1})^2\underline{p}^2/(2t)\right), \qquad \text{(D-10)}\end{aligned}$$

where the first inequality follows because $\sum_{i=1}^{N(\lambda^*)} \bar{F}(\lambda^*\delta^{-i}) \leq K+1$ from Lemma D.4 and $t+1 \leq N(\lambda^*)$, and the second inequality because $\sum_{i=t+1}^{N(\lambda^*)} \bar{F}(\lambda^*\delta^{-i}) \geq (N(\lambda^*) - t)\underline{p}$ where $\underline{p} = \min_{i=1}^{N(\lambda^*)} \bar{F}(\lambda^*\delta^{-i})$. This implies that

$$\begin{aligned}\sum_{t=1}^{N(\lambda^*)-1} \delta^{t}\mathbb{P}\{\tau^{\mathrm{G}} \leq t\} &\leq \sum_{t=1}^{N(\lambda^*)-1} \exp\left(-(N(\lambda^*) - t - \underline{p}^{-1})^2\underline{p}^2/(2t)\right) \\ &\leq \int_{1}^{N(\lambda^*)} \exp\left(-(N(\lambda^*) - t - \underline{p}^{-1})^2\underline{p}^2/(2t)\right) \mathrm{d}t \\ &\leq \int_{-\infty}^{\infty} \exp\left(-(t - N(\lambda^*) + \underline{p}^{-1})^2\underline{p}^2/(2N(\lambda^*))\right) \mathrm{d}t = \left((\pi N(\lambda^*))/(2\underline{p}^2)\right)^{1/2},\end{aligned}$$

where the first inequality follows from (D-10) and $\delta \leq 1$, the second inequality follows from the integral bound for summations since the summand in non-decreasing in t, the third inequality because $t \leq N(\lambda^*)$ and the fact that the integrand is positive, and the last equality from integrating the density of a normal distribution with mean $N(\lambda^*) + \underline{p}^{-1}$ and variance $N(\lambda^*)/\underline{p}^2$. We can use Lemma D.4 to bound $N(\lambda^*)$ as follows: $K+1 \geq \sum_{t=1}^{N(\lambda^*)} \mathbb{P}\left\{\tilde{r} \geq \lambda^*\delta^{-t}\right\} \geq \underline{p}N(\lambda^*)$. Therefore,

$$\mathbb{E}[(I)] \leq \left(\frac{\pi(K+1)}{2\underline{p}^3}\right)^{1/2}.$$

For the second term we have that

$$\mathbb{E}\left[(II)\right] = \sum_{i=N(\lambda)+1}^{N} \mathbb{P}\{\tau^{\mathrm{G}} = i\} \sum_{t=N(\lambda)+1}^{i} \delta^t = \sum_{t=N(\lambda)+1}^{N} \delta^t \sum_{i=t}^{N} \mathbb{P}\{\tau^{\mathrm{G}} = i\} = \sum_{t=N(\lambda)}^{N-1} \delta^{t+1} \mathbb{P}\{\tau^{\mathrm{G}} > t\}\,,$$

where the second equality follows from exchanging the order of summations. From Azuma's inequality for $t \geq N(\lambda^*)$ we obtain:

$$\begin{aligned}\mathbb{P}\left\{\tau^{\mathrm{G}} > t\right\} &= \mathbb{P}\left\{\sum_{i=1}^{t} \mathbf{1}\{r_i \geq v_i\} < K\right\} = \mathbb{P}\left\{M_t < K - \sum_{i=1}^{t} \bar{F}(v_i)\right\} \leq \mathbb{P}\left\{M_t < -\sum_{i=N(\lambda^*)+1}^{t} \bar{F}(v_i)\right\} \\ &\leq \mathbb{P}\left\{M_t < -(t - N(\lambda^*))\right\} \leq \exp\left(-(t-N(\lambda^*))^2/(2t)\right)\,, \end{aligned} \tag{D-11}$$

where the first inequality follows because $\sum_{i=1}^{N(\lambda^*)} \bar{F}(\lambda^* \delta^{-i}) \geq K$ from Lemma D.4 and $t \geq N(\lambda^*)$, and the second inequality because $\sum_{i=N(\lambda^*)+1}^{t} \bar{F}(v_i) \leq t - N(\lambda^*)$ since $\bar{F}(v_i) \leq 1$. This implies that

$$\begin{aligned}\sum_{t=N(\lambda^*)}^{N-1} \delta^{t+1} \mathbb{P}\{\tau^{\mathrm{G}} > t\} &\leq \sum_{t=N(\lambda^*)}^{N-1} \exp\left(-(t-N(\lambda^*))^2/(2t)\right) \leq \int_{N(\lambda^*)}^{N} \exp\left(-(t-N(\lambda^*))^2/(2t)\right) \mathrm{d}t \\ &\leq \int_{-\infty}^{\infty} \exp\left(-(t-N(\lambda^*))^2/(2N(\lambda^*))\right) \mathrm{d}t = \left(\pi N(\lambda^*)/2\right)^{1/2}\,,\end{aligned}$$

where the first inequality follows from (D-11) and $\delta \leq 1$, the second inequality follows from the integral bound for summations since the summand in non-decreasing in t, the third inequality because $t \geq N(\lambda^*)$ and the fact that the integrand is positive, and the last equality from integrating the density of a normal distribution with mean $N(\lambda^*)$ and variance $N(\lambda^*)$. Therefore,

$$\mathbb{E}[(II)] \leq \left(\frac{\pi(K+1)}{2\underline{p}}\right)^{1/2}\,.$$

Finally, by setting $\lambda = \lambda^*$ and $v_t = \lambda^* \delta^{-t}$ we have that

$$\bar{k}^{\mathrm{G}} = \left(K - \sum\nolimits_{t=1}^{N} \mathbf{1}\{r_t \geq v_t\}\right)^+ \leq \left(\sum\nolimits_{i=1}^{N(\lambda^*)} \left(\bar{F}(\lambda^* \delta^{-i}) - \mathbf{1}\{r_t \geq v_t\}\right)\right)^+ = \left(-M_{N(\lambda^*)}\right)^+\,,$$

where the first inequality follows because the summands are positive together with $N(\lambda^*) \leq N$ and $K \leq \sum_{i=1}^{N(\lambda^*)} \bar{F}(\lambda^* \delta^{-i})$ from Lemma D.4, and the last equality from our definition of M_t. Taking expectations and using that M_t is a martingale, we obtain that

$$\mathbb{E}\left[\left(-M_{N(\lambda^*)}\right)^+\right] = \frac{1}{2}\mathbb{E}[|M_{N(\lambda^*)}|] \leq \frac{1}{2}\left(\mathbb{E}[M^2_{N(\lambda^*)}]\right)^{1/2} = \frac{1}{2}\left(\sum\nolimits_{t=1}^{N(\lambda^*)} \mathrm{Var}[\mathbf{1}\{r_t \geq v_t\}]\right)^{1/2} \leq \frac{1}{4} N(\lambda^*)^{1/2}\,,$$

where the first equality follows because $\mathbb{E}[X^+] = \mathbb{E}|X|/2$ for any mean-zero random variable, the inequality from Lyapunov's inequality, the next equality because martingale differences are orthogonal, and the last inequality because the variance of a Bernoulli random variable is at most $1/4$. Therefore,

$$\mathbb{E}[\bar{k}^{\mathrm{G}}] \leq \frac{1}{4}\left(\frac{K+1}{\underline{p}}\right)^{1/2}\,.$$

Putting everything together, we conclude by Lemma D.4 that there exists a constant C independent of K such that $K^{-1}(V^* - V^{\mathrm{G}}) \leq CK^{-1/2}$. □

References

Dean, B. C., Goemans, M. X. and Vondrák, J. (2008), 'Approximating the stochastic knapsack problem: The benefit of adaptivity', *Mathematics of Operations Research* **33**(4), 945–964.

Hall, L. A., Schulz, A. S., Shmoys, D. B. and Wein, J. (1997), 'Scheduling to minimize average completion time: Off-line and on-line approximation algorithms', *Mathematics of Operations Research* **22**(3), 513–544.

Möhring, R. H., Schulz, A. S. and Uetz, M. (1999), 'Approximation in stochastic scheduling: The power of lp-based priority policies', *J. ACM* **46**(6), 924–942.