Economical Caching With Stochastic Prices*

Matthias Englert[1], Berthold Vöcking[2], Melanie Winkler[2]

[1] DIMAP and Department of Computer Science, University of Warwick
englert@dcs.warwick.ac.uk
[2] Department of Computer Science, RWTH Aachen University
{voecking,winkler}@cs.rwth-aachen.de

Abstract. In the *economical caching* problem, an online algorithm is given a sequence of prices for a certain commodity. The algorithm has to manage a buffer of fixed capacity over time. We assume that time proceeds in discrete steps. In step i, the commodity is available at price $c_i \in [\alpha, \beta]$, where $\beta > \alpha \geq 0$ and $c_i \in \mathbb{N}$. One unit of the commodity is consumed per step. The algorithm can buy this unit at the current price c_i, can take a previously bought unit from the storage, or can buy more than one unit at price c_i and put the remaining units into the storage.
In this paper, we study the economical caching problem in a probabilistic analysis, that is, we assume that the prices are generated by a random walk with reflecting boundaries α and β. We are able to identify the optimal online algorithm in this probabilistic model and analyze its expected cost and its expected *savings*, i.e., the cost that it saves in comparison to the cost that would arise without having a buffer. In particular, we compare the savings of the optimal online algorithm with the savings of the optimal offline algorithm in a probabilistic competitive analysis and obtain tight bounds (up to constant factors) on the ratio between the expected savings of these two algorithms.

1 Introduction

We study a stochastic version of the *economical caching* problem dealing with the management of a storage for a commodity whose price varies over time. An online algorithm is given a sequence of prices for a certain commodity. The algorithm has to manage a buffer of fixed capacity over time. We assume that time proceeds in discrete steps. In step i, the commodity is available at price $c_i \in [\alpha, \beta]$, where $\beta > \alpha \geq 0$ and $c_i \in \mathbb{N}$. One unit of the commodity is consumed per step. The algorithm can buy this unit at the current price c_i, can take a previously bought unit from the storage, or can buy more than one unit at price c_i and put the remaining units into the storage.

A slightly more general version of this problem has been introduced in [2], where it is assumed that the consume per step is not fixed to one but varies over time. The motivation for introducing this problem was the battery management of a hybrid car with two engines, one operated with petrol, the other

* Supported by the DFG GK/1298 "AlgoSyn" and EPSRC grant EP/F043333/1.

with electrical energy. An algorithm has to decide at which time the battery for the electrical energy should be recharged by the combustion engine. In this context, prices correspond to the combustion efficiency which, e.g., depends on the rotational speed. Obviously, the economical caching problem is quite generic and one can think of multiple other applications beyond hybrid cars, e.g., purchasing petrol at gas stations where prices vary on a daily basis, caching data streams in a mobile environment in which services that operate at different price levels are available in a dynamic fashion etc. The problem is related to the well-known *one-way trading* problem [1].

In [2], the economical caching problem has been analyzed in a worst case competitive analysis as introduced by [4]. Let $h = \beta/\alpha$ and assume $\beta > \alpha > 0$. It is shown that no memoryless algorithm (as formalized in [2]) can achieve a competitive factor better than $\sqrt{h}$. This competitive factor is guaranteed by a simple threshold scheme that purchases as much as possible units if the price is smaller than $\sqrt{\alpha\beta}$ and purchases as few as possible units if the price is above $\sqrt{\alpha\beta}$. This is, however, not the best possible competitive factor: It is shown that there is an online algorithm that takes into account the history in its buying decisions and, this way, achieves a competitive factor of $W\left(\frac{1-h}{eh} + 1\right)^{-1}$, where W denotes the LambertW function. This competitive factor beats $\sqrt{h}$ by a factor of about $\sqrt{2}$. Moreover, it is shown that this competitive factor is, in fact, best possible for any online algorithm.

The results from [2] determine the competitive factor for economical caching with respect to worst-case prices. The given bounds are tight. Nevertheless, the practical relevance of these results is questionable as the presented online algorithm optimizes against an adversary that generates a worst-case price sequence in response to the decisions of the algorithm. In contrast, in the aforementioned applications, the prices are set by a process that is (at least to a large extend) independent of the decisions of the online algorithm. We believe that a model in which prices are determined by a stochastic process can lead to more algorithms that perform better in practice.

In this paper, we assume that prices are generated by a random walk with reflecting barriers α and β. We do not claim that this is exactly the right process for any of the aforementioned applications. We hope, however, that our analysis is a first step towards analyzing more realistic, and possibly more complicated input distributions. Before stating our results, we give a formal description of the problem.

1.1 Formal description of the problem

We consider a finite process of n steps. In step $i \in \{1, \ldots, n\}$ the algorithm can purchase the commodity at price c_i per unit. Prices $c_1, c_2, \ldots, c_n, c_{n+1}$ are natural numbers generated by a random walk with reflecting boundaries $\alpha \in \mathbb{N}$ and $\beta \in \mathbb{N}$, $\beta > \alpha \geq 0$. Price c_{n+1} is used to valuate the units in the storage at the end of the process.

We assume that the random walk starts at a position chosen uniformly at random from $\{\alpha, \ldots, \beta\}$. This is justified by the fact that this allocation corresponds to the steady state distribution of the random walk. If $c_i \in \{\alpha+1, \beta-1\}$ then $\mathbf{Pr}[c_{i+1} = c_i - 1] = \mathbf{Pr}[c_{i+1} = c_i + 1] = \frac{1}{2}$; if $c_i = \alpha$ then $\mathbf{Pr}[c_{i+1} = \alpha] = \mathbf{Pr}[c_{i+1} = \alpha + 1] = \frac{1}{2}$; and if $c_i = \beta$ then $\mathbf{Pr}[c_{i+1} = \beta - 1] = \mathbf{Pr}[c_{i+1} = \beta] = \frac{1}{2}$. It is well known that, under these conditions, for every $1 \leq i \leq n+1$, $\alpha \leq j \leq \beta$, $\mathbf{Pr}[c_i = j] = t^{-1}$, where $t = |\{\alpha, \ldots, \beta\}| = \beta - \alpha + 1$.

The capacity of the storage is denoted by $b \in \mathbb{N}$. Let $s_i \in [0, b]$ denote the amount stored in the storage at the end of step i. Initially, the storage is assumed to be empty, that is, $s_0 = 0$. Consider an algorithm A. Let $u_i \in \{0, \ldots, b - s_{i-1} + 1\}$ denote the number of units that A purchases in step i. In each step, one unit of the commodity is consumed. Thus, $s_i = s_{i-1} + u_i - 1 \leq b$. Furthermore, if $s_{i-1} = 0$ at least one unit has to be purchased in step i. The *cost* of the algorithm is $\sum_{i=1}^{n} u_i \cdot c_i - s_n \cdot c_{n+1}$. Observe that we value the units in the buffer at the end of the process with the price c_{n+1}.

An algorithm is called *online* when u_i, the amount purchased in step i depends only on the prices $c_1, \ldots, c_i$ but not on the prices $c_{i+1}, \ldots, c_{n+1}$. An algorithm that is allowed to take into account knowledge about the full price sequence is called *offline*.

1.2 Preliminaries

Algorithm *NoBuffer* is an example of a very simple online algorithm. It sets $u_i = 1$, regardless of how the prices are chosen. As its name suggest, this algorithm does not exploit the storage. For this reason, its expected cost can be easily calculated. In every step i, the expected price is $\frac{\alpha+\beta}{2}$ since $\mathbf{Pr}[c_i = j] = t^{-1}$. Hence, by linearity of expectation, the expected cost of algorithm NoBuffer are $n \cdot \frac{\alpha+\beta}{2}$.

The *savings* of an algorithm A are defined to be the cost of NoBuffer minus the cost of the algorithm A. The savings are of particular relevance as this quantity describes the reduction in cost due to the availability of a buffer. For an algorithm A, let $\phi_i(A)$, $0 \leq i \leq n$, denote the savings accumulated until step i, i.e., the savings of A assuming that A as well as NoBuffer terminate after step i. Let $\Delta\phi_i(A) = \phi_i(A) - \phi_{i-1}(A)$ be the savings achieved in step i.

Proposition 1. *For $1 \leq i \leq n$, $\Delta\phi_i = s_i(c_{i+1} - c_i)$.*

Proof. The increase in cost of the NoBuffer algorithm during step i is c_i. The increase in cost of A during the same step, taking into account the change in the value of the storage, is

$$u_i \cdot c_i - s_i \cdot c_{i+1} + s_{i-1} \cdot c_i = c_i - s_i(c_{i+1} - c_i) \ ,$$

where the last equation follows from $s_i = s_{i-1} + u_i - 1$. Subtracting the increase in cost of A from the increase in cost of NoBuffer gives the proposition. □

The proposition shows that the savings achieved in step i correspond to the (possibly negative) increase of the value of the storage due to the change of the cost from c_i to c_{i+1} during the step.

1.3 Our results

We are able to identify the *optimal online algorithm*, i.e., the algorithm that achieves the smallest possible expected cost among all online algorithms. This algorithm works as follows: It fills the storage completely if $c_i = \alpha$ and uses units from the storage if $c_i > \alpha$, cf. Algorithm 1. The algorithm is called *Cautious* as it

---

Algorithm 1 (*Cautious*)

Input: c_i
1: **if** $c_i = \alpha$ **then**
2: $\quad u_i := b - s_{i-1} + 1$
3: **else**
4: $\quad u_i := 1 - \min\{1, s_i\}$
5: **end if**

---

only stores units bought at the lowest possible price. Obviously, there is no risk in buying these units. Somewhat surprisingly, this approach is the best possible online strategy.

We estimate the expected cost of the Cautious algorithm in the limiting case, i.e., assuming that n is sufficiently large. Per step, the expected cost of Cautious are

$$\alpha + t \cdot 2^{-\Theta(b/t^2)} \quad .$$

Observe, if b is of order $\omega(t^2)$ then $2^{-\Theta(b/t^2)} = o(1)$. That is, the expected cost per step approach the lower boundary α. If $b = O(t^2)$ then $2^{-\Theta(b/t^2)} = \Theta(1)$. Thus, in this case, the expected cost of Cautious are $\alpha + \Theta(t)$, which, however, is not a very significant result as it does not show the dependence of the expected cost on b, for $b = O(t^2)$.

To get more meaningful results we investigate the expected savings rather than the expected cost of the Cautious algorithm and compare it with the expected savings of an optimal offline algorithm. We get tight lower and upper bounds on the expected savings for both of these algorithms. The expected savings of the Cautious algorithm are $\Theta(\min\{\frac{b}{t}, t\})$ per step, while the expected savings of the optimal offline algorithm are $\Theta(\min\{\sqrt{b}, t\})$ per step. Thus the *competitive factor*, i.e., the ratio between the expected savings of Cautious and the expected savings of the optimal online algorithm, is of order

$$\frac{\min\left\{\frac{b}{t}, t\right\}}{\min\left\{\sqrt{b}, t\right\}} = \Theta\left(\min\left\{\frac{\sqrt{b}}{t}, 1\right\}\right) \quad .$$

2 Cautious is the optimal online algorithm

In this section, we show that Cautious achieves the smallest possible expected cost among all online algorithms.

Theorem 1. *For every $n \in \mathbb{N}$, there is no online algorithm achieving smaller expected cost than Cautious.*

Proof. Consider any online algorithm A. We assume that A is *reasonable*, i.e., it completely fills the storage in every step in which the price per unit is α. We can make this assumption w.l.o.g. since buying units at cost α instead of buying them later at possibly higher cost or buying additional units that are valuated at a price of $c_{n+1} \geq \alpha$ after step n cannot increase the cost of the algorithm.

We will show, for every $1 \leq i \leq n$, Cautious achieves larger expected savings in step i than A, that is, $\mathbf{E}[\Delta\phi_i(A)] \leq \mathbf{E}[\Delta\phi_i(\text{Cautious})]$. By linearity of expectation, this implies $\mathbf{E}[\phi_n(A)] \leq \mathbf{E}[\phi_n(\text{Cautious})]$. In words, the expected savings of A over all steps are not larger than the expected savings of Cautious, which implies the theorem.

Towards this end, we study the expected savings of A per step. Firstly, we analyze $\mathbf{E}[\Delta\phi_i(A)]$ conditioned on c_i being fixed to some value. For $k \in \{\alpha + 1, \ldots, \beta - 1\}$, Proposition 1 gives

$$\mathbf{E}[\Delta\phi_i(A) \mid c_i = k] = (\mathbf{E}[c_{i+1} \mid c_i = k] - k) \cdot \mathbf{E}[s_i \mid c_i = k] = 0$$

because $\mathbf{E}[c_{i+1} \mid c_i = k] = k$. In contrast,

$$\mathbf{E}[\Delta\phi_i(A) \mid c_i = \alpha] = (\mathbf{E}[c_{i+1} \mid c_i = \alpha] - \alpha) \cdot \mathbf{E}[s_i \mid c_i = \alpha] = \frac{1}{2} \cdot \mathbf{E}[s_i \mid c_i = \alpha]$$

because $\mathbf{E}[c_{i+1} \mid c_i = \alpha] = \alpha + \frac{1}{2}$, and

$$\mathbf{E}[\Delta\phi_i(A) \mid c_i = \beta] = (\mathbf{E}[c_{i+1} \mid c_i = \beta] - \beta) \cdot \mathbf{E}[s_i \mid c_i = \beta] = -\frac{1}{2} \cdot \mathbf{E}[s_i \mid c_i = \beta]$$

because $\mathbf{E}[c_{i+1} \mid c_i = \beta] = \beta - \frac{1}{2}$.

Applying these equations, the expected savings of A in step i can be calculated as

$$\begin{aligned}\mathbf{E}[\Delta\phi_i(A)] &= \mathbf{Pr}[c_i = \alpha] \cdot \mathbf{E}[\Delta\phi_i(A) \mid c_i = \alpha] + \mathbf{Pr}[c_i = \beta] \cdot \mathbf{E}[\Delta\phi_i(A) \mid c_i = \beta] \\ &= \frac{1}{2t}(\mathbf{E}[s_i \mid c_i = \alpha] - \mathbf{E}[s_i \mid c_i = \beta]) \ .\end{aligned}$$

Now observe that $\mathbf{E}[s_i \mid c_i = \alpha] = b$ because of our initial assumption that A is reasonable. Thus, the expected savings are maximized if $\mathbf{E}[s_i \mid c_i = \beta]$ is minimized, which is achieved if A has the property that for every step j with $c_j > \alpha$, $u_j \leq \beta - c_j - s_{j-1} + 1$. Cautious has this property. Thus Theorem 1 is shown. □

3 Expected cost of Cautious

We now calculate the expected cost of Cautious.

Theorem 2. *As n tends to infinity, the expected cost per step of Cautious tend to $\alpha + t \cdot 2^{-\Theta(b/t^2)}$.*

Proof. To simplify notation we will assume that $(\beta+\alpha)/2$ is integral. Fix an arbitrary step $i > b$. We will show that $\mathbf{E}\left[c_i - (s_{i-1} - s_i)(c_i - \alpha)\right] = \alpha + t \cdot 2^{-\Theta(b/t^2)}$ which corresponds to the expected cost of Cautious per step in the limiting case, that is, if we ignore an additive error that does not depend on the sequence length n (but may depend on b, α, and β). To see that $\mathbf{E}\left[c_i - (s_{i-1} - s_i)(c_i - \alpha)\right]$ is indeed related to the expected cost per step we take the sum over all steps $j > b$ and obtain

$$\begin{aligned}\sum_{j=b+1}^{n} \mathbf{E}\left[c_j - (s_{j-1} - s_j)(c_j - \alpha)\right] &= \mathbf{E}\left[\sum_{j=b+1}^{n} (c_j - (s_{j-1} - s_j)(c_j - \alpha))\right] \\ &= \mathbf{E}\left[\sum_{j=b+1}^{n} (u_j(c_j - \alpha) + \alpha)\right] \\ &= \mathbf{E}\left[\sum_{j=b+1}^{n} u_j c_j - (s_n - s_b) \cdot \alpha\right]\end{aligned}$$

which approximates the expected total cost of Cautious $\mathbf{E}\left[\sum_{j=1}^{n} u_j c_j - s_n c_{n+1}\right]$ within an additive error of $2b \cdot \beta$.

From the definition of the Cautious algorithm it follows immediately that $c_i - (s_{i-1} - s_i)(c_i - \alpha)$ equals α if $s_{i-1} > 0$ and equals c_i otherwise. Therefore $\mathbf{E}\left[c_i - (s_{i-1} - s_i)(c_i - \alpha)\right]$ is equal to

$$\begin{aligned}&\sum_{x=\alpha}^{\beta} \mathbf{Pr}\left[c_i = x\right] \cdot \left(\mathbf{Pr}\left[s_{i-1} = 0 \mid c_i = x\right] \cdot x + (1 - \mathbf{Pr}\left[s_{i-1} = 0 \mid c_i = x\right]) \cdot \alpha\right) \\ &\quad = \alpha + \sum_{x=\alpha}^{\beta} \mathbf{Pr}\left[c_i = x\right] \cdot \mathbf{Pr}\left[s_{i-1} = 0 \mid c_i = x\right] \cdot (x - \alpha) \\ &\quad = \alpha + \frac{1}{t} \sum_{x=\alpha}^{\beta} \mathbf{Pr}\left[s_{i-1} = 0 \mid c_i = x\right] \cdot (x - \alpha) \ ,\end{aligned}$$

Note that $\mathbf{Pr}\left[s_{i-1} = 0 \mid c_i = x\right]$ is the probability that the price was strictly larger than α for the last b steps, i.e., $c_{i-b} > \alpha, \ldots, c_{i-1} > \alpha$. In the following we will show that $\mathbf{Pr}\left[s_{i-1} = 0 \mid c_i = x\right] = 2^{-\Theta(b/t^2)}$, for $x \geq (\alpha + \beta)/2$, which implies the theorem since on one hand

$$
\begin{aligned}
&\alpha + \frac{1}{t}\sum_{x=\alpha}^{\beta} \mathbf{Pr}\left[s_{i-1}=0 \mid c_i = x\right] \cdot (x-\alpha) \\
&\qquad \le \alpha + \frac{1}{t}\sum_{x=\alpha}^{\beta} \mathbf{Pr}\left[s_{i-1}=0 \mid c_i = \beta\right] \cdot (x-\alpha) \\
&\qquad = \alpha + \frac{2^{-\Theta(b/t^2)}}{t}\sum_{x=\alpha}^{\beta}(x-\alpha) \\
&\qquad = \alpha + t \cdot 2^{-\Theta(b/t^2)}
\end{aligned}
$$

and on the other hand

$$
\begin{aligned}
&\alpha + \frac{1}{t}\sum_{x=\alpha}^{\beta} \mathbf{Pr}\left[s_{i-1}=0 \mid c_i = x\right] \cdot (x-\alpha) \\
&\qquad \ge \alpha + \frac{1}{t}\sum_{x=(\alpha+\beta)/2}^{\beta} \mathbf{Pr}\left[s_{i-1}=0 \mid c_i = (\alpha+\beta)/2\right] \cdot (x-\alpha) \\
&\qquad = \alpha + \frac{2^{-\Theta(b/t^2)}}{t}\sum_{x=(\alpha+\beta)/2}^{\beta}(x-\alpha) \\
&\qquad = \alpha + t \cdot 2^{-\Theta(b/t^2)} \quad .
\end{aligned}
$$

First we show $\mathbf{Pr}\left[s_{i-1}=0 \mid c_i = x\right] = 2^{-\Omega(b/t^2)}$ for any x. For simplicity we assume that b is a multiple of $2(t-1)^2$ and we divide the b previous steps into $b/(2(t-1)^2)$ phases of length $2(t-1)^2$. It is well-known that if the random walk starts with a price x' in the beginning of a phase, the expected time until the random walk will reach a price of α will be $(x'-\alpha)\cdot(2\beta-\alpha-x'-1)$ [3, p. 349]. Using Markov's inequality we know that the probability not to reach α in two times as many steps is at most $1/2$. The number of steps is maximized for $x'=\beta$ which gives us that the probability not to reach α in $2\cdot(\beta-\alpha)\cdot(\beta-\alpha-1) \le 2(t-1)^2$ steps is at most $1/2$. In other words, if we fix one of the phases, the probability that the price is strictly larger than α in every step of this phase is at most $1/2$ and because our argument does not depend on the price x' a phase starts with, this holds independently for every phase. Therefore, the probability not to have a price of α in any of the $b/(2(t-1)^2)$ phases is bounded from above by $2^{-b/(2(t-1)^2)} = 2^{-\Omega(b/t^2)}$.

It only remains to show that $\mathbf{Pr}\left[s_{i-1}=0 \mid c_i = x\right] = 2^{-O(b/t^2)}$ for $x \ge (\alpha+\beta)/2$. For this we divide the b previous steps into $\lceil 16\cdot b/(t-1)^2\rceil$ phases of length at most $(t-1)^2/16$. We call a phase starting with some price above $(\alpha+\beta)/2$ successful if the price at the end of the phase is above $(\alpha+\beta)/2$ as well and no step of the phase has a price of α. If we can show that the success probability for a fixed phase can be bounded from below by some constant $c > 0$, we can conclude that the probability for all $\lceil 16 \cdot b/(t-1)^2 \rceil$ phases to be successful is at least $c^{\lceil 16 \cdot b/(t-1)^2 \rceil} = 2^{-O(b/t^2)}$ which is the desired result and concludes the proof of the theorem.

Fix a phase, let A denote the event that no step in the phase has a price of α, and let B denote the event that the price at the end of the phase is above $(\alpha+\beta)/2$. We will show that the probability for the phase to be successful $\mathbf{Pr}[A \wedge B] = \mathbf{Pr}[B] \cdot \mathbf{Pr}[A \mid B]$ is at least $1/64$. Clearly $\mathbf{Pr}[B] \geq 1/2$. Now assume for contradiction that $\mathbf{Pr}[A \mid B] \leq 1/32$. Let X be the random variable describing the number of steps it takes a random walk starting from $(\alpha+\beta)/2$ to hit a price of α. We already know that $\mathbf{E}[X] = 3(t-1)^2/4$ (as before this follows from [3, p. 349]) and that $\mathbf{Pr}\left[X \geq j \cdot 2(t-1)^2\right] \leq 2^{-j}$, for any integral $j > 1$. Furthermore, the assumption $\mathbf{Pr}[A \mid B] \leq 1/32$ implies $\mathbf{Pr}\left[X \leq (t-1)^2/16\right] \geq 31/32$, which gives $\mathbf{Pr}\left[X > (t-1)^2/16\right] \leq 1/32$. Thus, we can conclude

- $\mathbf{Pr}\left[X \in [0, (t-1)^2/16]\right] \leq 1$,
- $\mathbf{Pr}\left[X \in ((t-1)^2/16, 16 \cdot (t-1)^2\right] \leq 1/32$, and
- for any integral $j \geq 1$, $\mathbf{Pr}\left[X \in (2 \cdot j \cdot (t-1)^2, 2 \cdot (j+1) \cdot (t-1)^2\right] \leq 2^{-j}$.

This gives us the following contradiction

$$\mathbf{E}[X] \leq \frac{(t-1)^2}{16} + \frac{(t-1)^2}{2} + \sum_{j=8}^{\infty} \frac{(j+1) \cdot (t-1)^2}{2^{j-1}} < \frac{3(t-1)^2}{4} = \mathbf{E}[X]$$

and completes the proof of the theorem. □

Note that the above arguments in the proof of Theorem 2 also directly imply the following lemma, which will be useful later in the proof of Theorem 3.

Lemma 1. *The probability that a random walk starting at a price of β will not see a price of α for at least $(t-1)^2$ steps is at least $(1/64)^{16}$*

Proof. The proof is analogous to the above arguments if we consider only 16 instead of $\lceil 16 \cdot b/(t-1)^2 \rceil$ phases of length at most $(t-1)^2/16$. □

4 Expected savings of Cautious

Theorem 3. $\mathbf{E}[\phi_n(\text{Cautious})] = \Theta(n \cdot \min\{\frac{b}{t}, t\})$.

Proof. Fix an arbitrary time step i. In Theorem 1 it was already shown that $\mathbf{E}[\Delta\phi_i(\text{Cautious})] = (b - \mathbf{E}[s_i | c_i = \beta])/(2t)$.

The number of units s_i in the storage only depend on the number of steps $\ell := i - i'$ that passed since the last previous step i' that had a price of α. More precisely, the storage is empty if $\ell \geq b$ and otherwise the storage contains $b - \ell$ units, where we define $\ell := b$ if there was no previous step i' with a price of α. Hence, $s_i = b - \min\{b, \ell\}$ and thus,

$$\mathbf{E}[\Delta\phi_i(\text{Cautious})] = \frac{\mathbf{E}[\min\{b, \ell\} \mid c_i = \beta]}{2t} \quad .$$

Clearly $\mathbf{E}[\min\{b,\ell\} \mid c_i = \beta]/(2t) \le \frac{b}{2t} = O(b/t)$. Since the total saving over all steps is obviously bounded by $n \cdot t$,

$$\mathbf{E}[\phi_n(\text{Cautious})] \le \min\left\{ \sum_{i=1}^{n} \mathbf{E}[\Delta\phi_i(\text{Cautious})], n \cdot t \right\} = O\left(n \cdot \min\left\{\frac{b}{t}, t\right\}\right) .$$

It remains to show that $\mathbf{E}[\min\{b,\ell\} \mid c_i = \beta]/(2t) = \Omega(\min\{\frac{b}{t}, t\})$. The probability that a random walk starting at a price of β does not reach a price of α for at least $(t-1)^2$ steps is equal to $\mathbf{Pr}\left[\ell > (t-1)^2 \mid c_i = \beta\right]$. Hence, due to Lemma 1 $\mathbf{Pr}\left[\ell > (t-1)^2 \mid c_i = \beta\right] \ge (1/64)^{16}$. It follows

$$\frac{\mathbf{E}[\min\{b,\ell\} \mid c_i = \beta]}{2t} \ge \frac{\min\{b,(t-1)^2\}}{64^{16} \cdot 2t} = \Omega\left(\min\left\{\frac{b}{t}, t\right\}\right) .$$

Taking the sum over all steps gives $\mathbf{E}[\phi_n(\text{Cautious})] = \sum_{i=1}^{n} \mathbf{E}[\Delta\phi_i(\text{Cautious})] = \Omega\left(n \cdot \min\left\{\frac{b}{t}, t\right\}\right)$. □

5 Expected savings of the optimal offline algorithm

The cost of an optimal offline algorithm is $\sum_{i=1}^{n} \min\{c_j \mid j \ge 0, i-b \le j \le i\} + \sum_{i=n-b+1}^{n}(\min\{c_i,\ldots,c_n,c_{n+1}\} - c_{n+1})$ (which can be formally proven with arguments analogous to [2, Lemma 2.1]). To get a measurement of the quality of this savings of our optimal online algorithm Cautious, we calculate $E[\phi_n(\text{Off})]$ for the optimal offline algorithm.

Theorem 4. $\mathbf{E}[\phi_n(\text{Off})] = \Theta(n \cdot \min\{\sqrt{b}, t\})$.

Proof. We will show that $\mathbf{E}[c_i - \min\{c_{i-b},\ldots,c_i\}] = \Theta(\min\{\sqrt{b}, t\})$, for every step $i > b$. The theorem follows because $\sum_{i=1}^{b}(c_i - \min\{c_j \mid j \ge 1, j \in [i-b,i]\}) \le b(\beta - \alpha)$ and $\sum_{i=n-b+1}^{n}(\min\{c_i,\ldots,c_n,c_{n+1}\} - c_{n+1}) \in [b(\alpha-\beta), 0]$ and therefore

$$\begin{aligned}
\mathbf{E}[\phi_n(\text{Off})] &= \mathbf{E}\left[\sum_{i=1}^{n} c_i\right] - \mathbf{E}\left[\sum_{i=1}^{n} \min\{c_j \mid j \ge 1, j \in [i-b,i]\}\right] \\
&\quad - \mathbf{E}\left[\sum_{i=n-b+1}^{n} (\min\{c_i,\ldots,c_n,c_{n+1}\} - c_{n+1})\right] \\
&= \sum_{i=b+1}^{n} \mathbf{E}[c_i - \min\{c_{i-b},\ldots,c_i\}] \\
&\quad + \mathbf{E}\left[\sum_{i=1}^{b}(c_i - \min\{c_j \mid j \ge 1, j \in [i-b,i]\})\right] \\
&\quad + \mathbf{E}\left[\sum_{i=n-b+1}^{n} (\min\{c_i,\ldots,c_n,c_{n+1}\} - c_{n+1})\right] \\
&= \Theta(n \cdot \min\{\sqrt{b}, t\}) .
\end{aligned}$$

In this proof we will make use of the well-known fact that it is possible to map a symmetric unrestricted random walk $c'_1, \dots, c'_{n+1}$ to the random walk $c_1, \dots, c_{n+1}$ with reflecting boundaries at α and β. More precisely, let c'_1 be chosen uniformly at random from $\{0, \dots, t-1\}$ and set $\mathbf{Pr}\left[c'_{i+1} = c'_i + 1\right] = \mathbf{Pr}\left[c'_{i+1} = c'_i - 1\right] = \frac{1}{2}$, for any c'_i. Setting $c_i := \alpha + \min\{(c'_i \mod 2t), 2t-1-(c'_i \mod 2t)\}$ for each i, maps this unrestricted random walk to the reflected random walk.

Fix a time step $i > b$. We have to show that $\mathbf{E}\left[c_i - \min\{c_{i-b}, \dots, c_i\}\right] = \Theta(\min\{\sqrt{b}, t\})$ and we start with the upper bound. It is easy to see that $c_i - c_j \leq |c'_i - c'_j|$ for every j and it is known [3, Theorem III. 7.1] that

$$\mathbf{Pr}\left[|c'_i - \min\{c'_{i-b}, \dots, c'_i\}| = \ell\right] = \frac{1}{2^{b-1}} \cdot \left(\binom{b}{\frac{1}{2}(b+\ell)} + \binom{b}{\frac{1}{2}(b+\ell+1)}\right) .$$

Thus,

$$\begin{aligned}
\mathbf{E}\left[c_i - \min\{c_{i-b}, \dots, c_i\}\right] &\leq \mathbf{E}\left[|c'_i - \min\{c'_{i-b}, \dots, c'_i\}|\right] \\
&= \sum_{\ell=1}^{b} \frac{\ell}{2^{b-1}} \cdot \left(\binom{b}{\frac{1}{2}(b+\ell)} + \binom{b}{\frac{1}{2}(b+\ell+1)}\right) \\
&\leq 2 \cdot \sum_{\ell=1}^{b} \frac{\ell}{2^{b-1}} \cdot \binom{b}{\frac{1}{2}(b+\ell)} = O\big(\sqrt{b}\big) .
\end{aligned}$$

Since obviously $c_i - \min\{c_{i-b}, \dots, c_i\} \leq t$, we obtain $\mathbf{E}\left[c_i - \min\{c_{i-b}, \dots, c_i\}\right] = O(\min\{\sqrt{b}, t\})$.

To show the lower bound we, once again, start by considering the symmetric unrestricted random walk and obtain

$$\begin{aligned}
\mathbf{Pr}\left[c'_i - c'_{i-b} > \frac{\sqrt{b}}{4}\right] &= \frac{1}{2} \cdot \mathbf{Pr}\left[|c'_i - c'_{i-b}| > \frac{\sqrt{b}}{4}\right] \\
&= \frac{1}{2} \cdot \left(1 - \mathbf{Pr}\left[|c'_i - c'_{i-b}| \leq \frac{\sqrt{b}}{4}\right]\right) \\
&= \frac{1}{2} \cdot \left(1 - 2^{-b} \cdot \sum_{\ell=-\sqrt{b}/4}^{\sqrt{b}/4} \binom{b}{\frac{1}{2}(b+\ell)}\right) \\
&\geq \frac{1}{2} \cdot \left(1 - 2^{-b} \cdot \sum_{\ell=-\sqrt{b}/4}^{\sqrt{b}/4} \binom{b}{\frac{b}{2}}\right) \geq \frac{1}{4} .
\end{aligned}$$

Since the step size of the random walk is one, $c'_0 - c'_{i-b} > \sqrt{b}/4$ implies that there exists a $0 \leq k \leq b$ such that $c'_i - c'_{i-k} = \lceil \sqrt{b}/4 \rceil$. Therefore we can conclude for our reflecting random walk that

$$\mathbf{Pr}\left[c_i - \min\{c_{i-b}, \dots, c_i\} \geq \left\lceil \frac{\sqrt{b}}{4} \right\rceil \mid c_i = j\right] \geq \frac{1}{4}$$

for any $j \geq \alpha + \lceil \sqrt{b}/4 \rceil$.

By exploiting this lower bound on the probability that there is a large difference between the current price c_i and the minimum price over the last b steps $\min\{c_{i-b}, \ldots, c_i\}$ we get

$$
\begin{aligned}
&\mathbf{E}\left[c_i - \min\{c_{i-b}, \ldots, c_i\}\right] \\
&\quad = \sum_{j=\alpha}^{\beta} \mathbf{Pr}\left[c_{i-b} = j\right] \cdot \mathbf{E}\left[c_i - \min\{c_{i-b}, \ldots, c_i\} \mid c_i = j\right] \\
&\quad \geq \frac{1}{t} \sum_{j=\alpha+\lceil \sqrt{b}/4 \rceil}^{\beta} \mathbf{E}\left[c_i - \min\{c_{i-b}, \ldots, c_i\} \mid c_i = j\right] \\
&\quad \geq \frac{\lceil \sqrt{b}/4 \rceil}{t} \sum_{j=\alpha+\lceil \sqrt{b}/4 \rceil}^{\beta} \mathbf{Pr}\left[c_i - \min\{c_{i-b}, \ldots, c_i\} \geq \left\lceil \frac{\sqrt{b}}{4} \right\rceil \mid c_i = j\right] \\
&\quad \geq \frac{1}{t} \sum_{j=\alpha+\lceil \sqrt{b}/4 \rceil}^{\beta} \frac{\sqrt{b}}{16} \\
&\quad = \frac{\sqrt{b}}{16t} \cdot \left(t - \left\lceil \frac{\sqrt{b}}{4} \right\rceil\right) .
\end{aligned}
$$

Hence, for $b \leq t^2$, $\mathbf{E}\left[c_i - \min\{c_{i-b}, \ldots, c_i\}\right] = \Omega(\sqrt{b})$. Since the amount $c_i - \min\{c_{i-b}, \ldots, c_i\}$ is monotonic in b, we also have $\mathbf{E}\left[c_i - \min\{c_{i-b}, \ldots, c_i\}\right] = \Omega(t)$ for $b > t^2$, which concludes the proof of the lower bound. □

6 Conclusions

We considered the Economical Caching problem with stochastic prices generated by a random walk with reflecting boundaries at α and β. We identified an optimal online algorithm for this setting and gave bounds on its expected cost and savings. However, modeling the price development as a random walk might be a serious simplification of many real world scenarios. Therefore this work should be seen as a first step towards an analysis of more realistic input models.

In reality, prices may increase or decrease by more than one unit in each time step and the probabilities for an increase and a decrease might differ and depend on the current price. Furthermore, the assumption that the upper and lower bound α and β is known to the algorithm, is unrealistic in most applications. In fact, in general, these strict bounds do not even exist.

References

1. A. Borodin and R. El-Yaniv. *Online computation and competitive analysis.* Cambridge University Press, New York, NY, USA, 1998.

2. M. Englert, H. Röglin, J. Spönemann, and B. Vöcking. Economical caching. In *Proceedings of the 26th Symposium on Theoretical Aspects of Computer Science (STACS)*, pages 385–396, 2009.
3. W. Feller. *An Introduction to Probability Theory and Its Applications.* Wiley, 3rd edition, 1971.
4. D. D. Sleator and R. E. Tarjan. Amortized efficiency of list update and paging rules. *Commun. ACM*, 28(2):202–208, 1985.