Positivity CrossMark

A prophet inequality for L^p-bounded dependent random variables

Adam Osękowski

Abstract Let $X = (X_n)_{n \geq 1}$ be a sequence of arbitrarily dependent nonnegative random variables satisfying the boundedness condition

$$\sup_{\tau} \mathbb{E} X_{\tau}^p \leq t,$$

where $t > 0$, $1 < p < \infty$ are fixed numbers and the supremum is taken over all finite stopping times of X. Let $M = \mathbb{E} \sup_n X_n$ and $V = \sup_{\tau} \mathbb{E} X_{\tau}$ denote the expected supremum and the optimal expected return of the sequence X, respectively. We establish the prophet inequality

$$M \leq V + \frac{V}{p-1} \log \left(\frac{te}{V^p} \right)$$

and show that the bound on the right is the best possible. The proof of the inequality rests on Burkholder's method and exploits properties of certain special functions. The proof of the sharpness is somewhat indirect, but we also provide an indication how the extremal sequences can be constructed.

Keywords Prophet inequality · Optimal stopping · Bellman function · Best constants

A. Osękowski (✉)
Department of Mathematics, Informatics and Mechanics, University of Warsaw,
Banacha 2, 02-097 Warsaw, Poland
e-mail: ados@mimuw.edu.pl

Birkhäuser

1 Introduction

The purpose of the paper is to establish a sharp estimate between the expected supremum of a sequence $X = (X_n)_{n\geq 1}$ of L^p-bounded random variables and the optimal expected return (i.e., optimal stopping value) of X. Such comparisons are called "prophet inequalities" in the literature and play a distinguished role in the theory of optimal stopping, as evidenced in the papers of Allaart [1], Allaart and Monticino [2], Assaf et al. [4,5], Boshuizen [3,6], Hill [8,9], Hill and Kertz [10–12], Kennedy [13,14], Kertz [15,16], Krengel, Sucheston and Garling [17–19], Tanaka [24,25] and many others.

We start with the necessary background and notation. Assume that $X = (X_n)_{n\geq 1}$ is a sequence of (possibly dependent) random variables defined on the probability space $(\Omega, \mathcal{F}, \mathbb{P})$. With no loss of generality, we may assume that this probability space is the interval $[0, 1]$ equipped with its Borel subsets and Lebesgue measure. Let $(\mathcal{F}_n)_{n\geq 1} = (\sigma(X_1, X_2, \ldots, X_n))_{n\geq 1}$ be the natural filtration of X. The problem can be generally stated as follows: under some boundedness condition on X, find universal inequalities which compare $M = \mathbb{E}\sup_n X_n$, the expected supremum of the sequence, with $V = \sup_\tau \mathbb{E}X_\tau$, the optimal stopping value of the sequence; here τ runs over the class $\mathcal{T}$ of all finite stopping times adapted to $(\mathcal{F}_n)_{n\geq 1}$. The term "prophet inequality" arises from the optimal-stopping interpretation of M, which is the optimal expected return of a player endowed with complete foresight; this player observes the sequence X and may stop whenever he wants, incurring a reward equal to the variable at the time of stopping. With complete foresight, such a player obviously stops always when the largest value is observed, and on the average, his reward is equal to M. On the other hand, the quantity V corresponds to the optimal return of the non-prophet player.

Let us mention here several classical results in this direction; for an excellent exposition on the subject, we refer the interested reader to the survey by Hill and Kertz [12]. The first universal prophet inequality is due to Krengel, Sucheston and Garling [17,18]: if $X_1, X_2, \ldots$ are independent and nonnegative, then

$$M \leq 2V$$

and the constant 2 is the best possible. The next result, coming from [8] and [10], states that if $X_1, X_2, \ldots$ are independent and take values in $[0, 1]$, then

$$M - V \leq \frac{1}{4}$$

and

$$M \leq V - V^2.$$

Both estimates are sharp: equalities may hold for some non-trivial sequences X. Analogous inequalities for other types of variables $X_1, X_2, \ldots$ (e.g., arbitrarily-dependent

and uniformly bounded, i.i.d., averages of independent r.v.'s, exchangeable r.v.'s, etc.), as well as for other stopping options (for instance, stopping with partial recall, stopping several times, using only threshold stopping rules, etc.) have been studied intensively in the literature and have found many interesting applications. We refer the reader to the papers cited at the beginning.

The motivation for the results obtained in this paper comes from the following statement proved by Hill and Kertz [11]: if $X_1, X_2, \ldots$ are arbitrarily dependent and take values in [0, 1], then we have the sharp bound

$$M \le V - V \log V. \tag{1.1}$$

There is a very natural problem concerning the L^p-version of this result, where p is a fixed number between 1 and infinity. For example, consider the following interesting question. Suppose that $X_1, X_2, \ldots$ are nonnegative random variables satisfying $\sup_n \mathbb{E} X_n^p \le 1$. What is the analogue of (1.1)? Unfortunately, as we will see in Sect. 5 below, there is no non-trivial prophet inequality in this setting. More precisely, for any $K > 0$ one can construct a sequence X bounded in L^p which satisfies $V = 1$ and $M \ge K$.

We will work under the more restrictive assumption

$$X_1, \ X_2, \ \ldots \ \text{are nonnegative and satisfy} \ \sup_{\tau \in \mathcal{T}} \mathbb{E} X_\tau^p \le t, \tag{1.2}$$

where t is a given positive number. For instance, this holds true, if the sequence X possesses a majorant ξ which satisfies $\mathbb{E} \xi^p \le t$.

The main result of the paper is the following.

Theorem 1.1 *Let $t > 0$, $1 < p < \infty$ be fixed and suppose that $X_1, X_2, \ldots$ satisfy* (1.2). *Then*

$$M \le V + \frac{V}{p-1} \log\left(\frac{te}{V^p}\right) \tag{1.3}$$

and the inequality is sharp: the bound on the right cannot be replaced by a smaller number.

Note that this statement generalizes the inequality (1.1) of Hill and Kertz: it suffices to take $t = 1$ and let p go to ∞ to recover the bound. On the other hand, the expression on the right of (1.3) explodes as $p \downarrow 1$, which indicates that there is no prophet inequality in the limit case $p = 1$.

A few words about the proof. Our approach is based on the following two-step procedure: first we show that it suffices to establish (1.3) under the additional assumption that X is a nonnegative supermartingale; second, we prove that in the supermartingale setting, the validity of (1.3) is equivalent to the existence of a certain special function which enjoys appropriate majorization and convexity properties. In the literature this equivalence is often referred to as Burkholder's method or Bellman function method, and it has turned out to be extremely efficient in numerous problems in probability and analysis: consult e.g. [7,20,21,26] and references therein.

We have organized the paper as follows. In the next section we reduce the problem to the supermartingale setting. Section 3 contains the description of Burkholder's method (or rather its variant which is needed in the study of (1.3) for supermartingales). In Sect. 4 we apply the method and provide the proof of Theorem 1.1. In the final part of the paper we show that there are no interesting prophet inequalities in the case when the variables $X_1, X_2, \ldots$ are only assumed to be bounded in L^p.

2 A reduction

In this section we show how to relate (1.3) to a certain inequality for nonnegative supermartingales. Throughout, we use the notation $X^* = \sup_{n \geq 1} X_n$ and $X_m^* = \sup_{1 \leq n \leq m} X_n$ for the maximal and the truncated maximal function of X. Recall that $V = \sup_{\tau \in \mathcal{T}} \mathbb{E} X_\tau$. We start with the observation that it is enough to deal with finite sequences only (in this paper we say that X is finite, if it is of the form $(X_1, X_2, \ldots, X_{N-1}, X_N, X_N, X_N, \ldots)$ for some deterministic N). This is straightforward: suppose we have successfully established the prophet inequality in this special case, and pick an arbitrary, possibly non-finite X. Then for any fixed N, the truncated sequence $X^N = (X_1, X_2, \ldots, X_{N-1}, X_N, X_N, X_N, \ldots)$ is finite, inherits (1.2) and its optimal expected return does not exceed V. Since the function $V \mapsto V + \frac{V}{p-1} \log \left(\frac{te}{V^p} \right)$ is nondecreasing on $(0, t^{1/p}]$, an application of (1.3) gives

$$\mathbb{E}(X^N)^* \leq V + \frac{V}{p-1} \log \left(\frac{te}{V^p} \right).$$

It remains to let $N \to \infty$ and use Lebesgue's monotone convergence theorem.

Lemma 2.1 *Suppose that $X = (X_n)_{n \geq 1}$ is an arbitrarily dependent finite sequence of random variables satisfying $X_1 \equiv 0$ and (1.2). Then there is a finite supermartingale $Y = (Y_n)_{n \geq 1}$ adapted to the filtration of X, which satisfies*

$$Y_n \geq X_n \text{ almost surely for all } n, \tag{2.1}$$
$$\mathbb{P}(Y_1 = V) = 1 \tag{2.2}$$

and

$$\sup_{\tau \in \mathcal{T}} \mathbb{E} Y_\tau^p \leq t, \qquad \sup_{\tau \in \mathcal{T}} \mathbb{E} Y_\tau \leq V. \tag{2.3}$$

Note that the additional assumption $X_1 \equiv 0$ is not restrictive at all: we can always replace the initial sequence $X_1, X_2, \ldots$ with $0, X_1, X_2, \ldots$, and the prophet inequality remains the same. In the proof of the above lemma we will need the notion of essential supremum, a well-known object in the optimal stopping theory. Let us briefly recall its definition, for details and properties we refer the reader to the monographs of Peskir and Shiryaev [22] and Shiryaev [23].

Definition 2.1 Let $(Z_\alpha)_{\alpha \in I}$ be a family of random variables. Then there is a countable subset J of I such that the random variable $\overline{Z} = \sup_{\alpha \in J} Z_\alpha$ satisfies the following two properties:

(i) $\mathbb{P}(Z_\alpha \leq \overline{Z}) = 1$ for each $\alpha \in I$,

(ii) if $\tilde{Z}$ is another random variable satisfying (i) in the place of $\overline{Z}$, then $\mathbb{P}(\overline{Z} \leq \tilde{Z}) = 1$.

The random variable $\overline{Z}$ is called the *essential supremum* of $(Z_\alpha)_{\alpha \in I}$.

Proof of Lemma 2.1 We will use some basic facts from the optimal stopping theory; for details, we refer the reader to Chapter I in Peskir and Shiryaev [22]. Suppose that $X = (X_1, X_2, \ldots, X_N, X_N, X_N, \ldots)$ is a finite sequence and let Y be the Snell envelope of X, i.e., the smallest adapted supermartingale majorizing the sequence $(X_n)_{n \geq 1}$. It is a well-known fact that for each n the variable Y_n is given by the formula

$$Y_n = ess\, sup\Big\{\mathbb{E}(X_\sigma | \mathcal{F}_n) \,:\, \sigma \in \mathcal{T}_n\Big\},$$

where $\mathcal{T}_n$ denotes the class of all finite adapted stopping times not smaller than n. Since $\sigma \equiv n$ belongs to $\mathcal{T}_n$, the inequality (2.1) is given for free. To show (2.2), observe that Y_1 is a constant random variable, since the σ-algebra $\mathcal{F}_1 = \sigma(X_1)$ is trivial. Thus the bound $Y_1 \geq V$ follows directly from the above formula for Y_1 and the definition of an essential supremum. On the other hand, for any finite stopping time σ we have $V \geq \mathbb{E}X_\sigma = \mathbb{E}(X_\sigma | \mathcal{F}_1)$ almost surely, which implies $\mathbb{P}(V \geq Y_1) = 1$, again from the definition of an essential supremum. This gives (2.2). Since the sequence X stabilizes after N steps, so does Y and therefore the second estimate in (2.3) holds true, directly from (2.2) and the supermartingale property. To prove the first bound in (2.3), recall that Y can be alternatively defined by the backward induction

$$Y_N = X_N, \quad Y_n = \max\big\{X_n, \mathbb{E}(Y_{n+1} | \mathcal{F}_n)\big\}, \quad n = 1, 2, \ldots, N-1.$$

This implies that $Y_n = \mathbb{E}(X_{\tau_n} | \mathcal{F}_n)$, where the stopping time τ_n is given by

$$\tau_n = \inf\big\{k \in \{n, n+1, \ldots, N\} : Y_k = X_k\big\}.$$

Thus, if we fix an arbitrary $\tau \in \mathcal{T}$, then

$$Y_\tau = Y_{\tau \wedge N} = \sum_{n=1}^{N} Y_n 1_{\{\tau \wedge N = n\}} = \sum_{n=1}^{N} 1_{\{\tau \wedge N = n\}} \mathbb{E}\big(X_{\tau_n} | \mathcal{F}_n\big),$$

which gives

$$\mathbb{E}Y_\tau^p = \sum_{n=1}^{N} \mathbb{E}\Big[\mathbb{E}\big(1_{\{\tau \wedge N = n\}} X_{\tau_n} | \mathcal{F}_n\big)^p\Big] \leq \sum_{n=1}^{N} \mathbb{E}\Big[1_{\{\tau \wedge N = n\}} X_{\tau_n}^p\Big] = \mathbb{E}X_\sigma^p,$$

where $\sigma = \sum_{n=1}^{N} 1_{\{\tau \wedge N = n\}} \tau_n$. We easily check that σ is a stopping time: for any $1 \leq k \leq N$, the event

$$\{\sigma = k\} = \bigcup_{n=1}^{N} (\{\tau \wedge N = n\} \cap \{\tau_n = k\}) = \bigcup_{n=1}^{k} (\{\tau \wedge N = n\} \cap \{\tau_n = k\})$$

belongs to $\mathcal{F}_k$. Therefore, the boundedness assumption (1.2) implies $\mathbb{E}Y_\tau^p \leq t$, as desired.

Therefore, it suffices to establish the inequality (1.3) under the additional assumption that the process X is a finite supermartingale and the variable X_1 is constant almost surely. By some standard approximation arguments, we may further restrict ourselves to the class of *simple* supermartingales; recall that the sequence $X = (X_n)_{n\geq 1}$ is called simple, if for each n the random variable X_n takes only a finite number of values. We are ready to apply Burkholder's method, which is introduced in the next section.

3 Burkholder's method

Now we will describe the main tool which will be used to establish the inequality (1.3). Distinguish the set

$$D = \{(x, y, t) \in [0, \infty) \times [0, \infty) \times [0, \infty) : x^p \leq t, \ x \leq y\}$$

and, for each $(x, y, t) \in D$, let $\mathcal{S}(x, y, t)$ denote the class of all simple, finite and nonnegative supermartingales $X = (X_n)_{n\geq 1}$ satisfying $X_1 \equiv x$ and $\sup_\tau \mathbb{E}X_\tau^p \leq t$, where the supremum is taken over all stopping times adapted to the natural filtration of X. Suppose that we are interested in the explicit formula for the function

$$\mathbb{B}(x, y, t) = \sup \left\{ \mathbb{E}(X_n^* \vee y) \right\},$$

where the supremum is taken over all positive integers n and all $X \in \mathcal{S}(x, y, t)$. The key idea in the study of this problem is to introduce the class $\mathcal{C}$ which consists of all functions $B : D \to \mathbb{R}$ satisfying

$$B(x, y, t) \geq y \quad \text{for any } (x, y, t) \in D, \tag{3.1}$$

and the following concavity-type property: if $\alpha \in [0, 1]$, $(x, y, t) \in D$ and $(x_\pm, x_\pm \vee y, t_\pm) \in D$ satisfy $\alpha x_- + (1 - \alpha)x_+ \leq x$ and $\alpha t_- + (1 - \alpha)t_+ \leq t$, then

$$B(x, y, t) \geq \alpha B(x_-, x_- \vee y, t_-) + (1 - \alpha)B(x_+, x_+ \vee y, t_+). \tag{3.2}$$

We turn to the main result of this section. Recall that the probability space is the interval $[0, 1]$ with its Borel subsets and Lebesgue measure.

Theorem 3.1 *The function $\mathbb{B}$ is the least element of the class $\mathcal{C}$.*

Proof It is convenient to split the reasoning into two parts.

Step 1. First we will show that $\mathbb{B}$ belongs to the class $\mathcal{C}$. The majorization (3.1) is immediate, since $X_n^* \vee y \geq y$. The main difficulty lies in proving the concavity property (3.2). Fix the parameters $\alpha, x, t, x_\pm, t_\pm$ as in the statement and pick arbitrary supermartingales $X^- \in \mathcal{S}(x_-, y, t_-)$, $X^+ \in \mathcal{S}(x_+, y, t_+)$. We splice these two processes into one sequence $X = (X_n)_{n\geq 1}$ by setting $X_1 \equiv x$ and, for $n \geq 2$,

$$X_n(\omega) = \begin{cases} X_{n-1}^-(\omega/\alpha) & \text{if } \omega \in [0, \alpha], \\ X_{n-1}^+((\omega - \alpha)/(1 - \alpha)) & \text{if } \omega \in (\alpha, 1]. \end{cases}$$

Then X is a nonnegative supermartingale (with respect to its natural filtration), because the processes $X^\pm$ have this property and $\alpha x_- + (1 - \alpha)x_+ \leq x$. Furthermore, for any stopping time τ of X we have $\mathbb{E}X_\tau^p \leq t$. To see this, we consider two cases.

1. If $\mathbb{P}(\tau = 1) > 0$, then the set $\{\tau \leq 1\}$ is nonempty; combining this with the facts that X_1 is constant and τ is a stopping time of X, we see that $\{\tau \leq 1\} = \Omega$, or $\tau \equiv 1$. Then $\mathbb{E}X_\tau^p = x^p \leq t$ by the definition of D.

2. Suppose that $\{\tau = 1\} = \emptyset$, or $\tau \geq 2$ almost surely. Then we easily verify that the variables $\tau^\pm$, given by

$$\tau^-(\omega) = \tau(\alpha\omega) - 1, \qquad \tau^+ = \tau(\alpha + (1 - \alpha)\omega) - 1,$$

are stopping times of X^- and X^+. Therefore,

$$\mathbb{E}X_\tau^p = \alpha\mathbb{E}(X_{\tau^-}^-)^p + (1 - \alpha)\mathbb{E}(X_{\tau^+}^+)^p \leq \alpha t_- + (1 - \alpha)t_+ \leq t.$$

Hence $X \in \mathcal{S}(x, y, t)$. Since $x \leq y$, we have $X_n^* \vee y = \sup_{2 \leq k \leq n} X_k \vee y$ and thus

$$\mathbb{B}(x, y, t) \geq \mathbb{E}(X_n^* \vee y) = \alpha\mathbb{E}\big((X_{n-1}^-)^* \vee y\big) + (1 - \alpha)\mathbb{E}\big((X_{n-1}^+)^* \vee y\big).$$

Take the supremum over all n and $X^\pm$ as above to obtain the desired bound (3.2).

Step 2. Now suppose that B is an arbitrary element of $\mathcal{C}$; we must prove that $\mathbb{B} \leq B$. To do this, rephrase the condition (3.2) as follows. Suppose that (X, T) is an arbitrary random variable with two-point distribution, such that $\mathbb{P}(X^p \leq T) = 1$. Then for any $(x, y, t) \in D$ such that $\mathbb{E}X \leq x$ and $\mathbb{E}T \leq t$, we have

$$B(x, y, t) \geq \mathbb{E}B(X, y, T). \tag{3.3}$$

Note that the set $\{(x, t) : x^p \leq t\}$ is convex. Therefore, by straightforward induction, the above inequality extends to the case when (X, T) is an arbitrary simple random variable satisfying $X^p \leq T$ with probability 1. Now, pick $X \in \mathcal{S}(x, y, t)$ and consider the sequence (X, Y, T), where $Y_n = X_n^* \vee y$ and $T_n = \text{ess sup}_{\tau \in \mathcal{T}_n} \mathbb{E}(X_\tau^p|\mathcal{F}_n)$ (here $\mathcal{T}_n$ denotes the class of all stopping times of X not smaller than n). Then the process $B(X, Y, T)$ is a supermartingale: to see this, fix $n \geq 1$ and apply (3.3) conditionally with respect to $\mathcal{F}_n$, with $\tilde{x} := X_n$, $\tilde{y} = Y_n$, $\tilde{t} = T_n$, $\tilde{X} = X_{n+1}$ and $\tilde{T} = T_{n+1}$. Let us verify the assumptions: the inequalities $\tilde{x}^p \leq \tilde{t}$, $\tilde{x} \leq \tilde{y}$ and $\tilde{X}^p \leq \tilde{T}$ are evident; the inequalities $\mathbb{E}(\tilde{X}|\mathcal{F}_n) \leq \tilde{x}$ and $\mathbb{E}(\tilde{T}|\mathcal{F}_n) \leq \tilde{t}$ follow from the supermartingale property of X and T (T is a supermartingale, since it is the Snell envelope of the sequence $(X_n^p)_{n \geq 1}$). Thus, (3.3) yields

$$\begin{aligned} B(X_n, Y_n, T_n) &\geq \mathbb{E}\Big[B\big(X_{n+1}, X_{n+1} \vee Y_n, T_{n+1}\big)|\mathcal{F}_n\Big] \\ &= \mathbb{E}\Big[B(X_{n+1}, Y_{n+1}, T_{n+1})|\mathcal{F}_n\Big]. \end{aligned}$$

Combining this with (3.1) yields

$$\mathbb{E}(X_n^* \vee y) \leq \mathbb{E}B(X_n, Y_n, T_n) \leq \mathbb{E}B(X_1, Y_1, T_1) = B(x, y, \sup_{\tau \in \mathcal{T}} \mathbb{E}X_\tau^p) \leq B(x, y, t).$$

Here in the last inequality we have used the fact that the function $t \mapsto B(x, y, t)$ is nondecreasing; this follows immediately from (3.2), applied to $x_+ = x_- = x$ and $t_+ = t_- < t$. Taking the supremum over all n and all $X \in \mathcal{S}(x, y, t)$, we obtain the bound $\mathbb{B}(x, y, t) \leq B(x, y, t)$. This finishes the proof. $\square$

4 Proof of Theorem 1.1

We will prove that the function $\mathbb{B}$ admits the following explicit formula:

$$\mathbb{B}(x, y, t) = \begin{cases} y + \frac{x}{p-1} \log\left(\frac{te}{xy^{p-1}}\right) & \text{if } y \leq (t/x)^{1/(p-1)}, \\ y + ty^{1-p}/(p-1) & \text{if } y > (t/x)^{1/(p-1)}. \end{cases}$$

This will clearly yield the assertion of Theorem 1.1, which is nothing else but the explicit formula for $\mathbb{B}(V, V, t)$. Denote expression on the right above by $B(x, y, t)$.

4.1 Proof of $\mathbb{B} \leq B$

In the light of Theorem 3.1, it suffices to verify that $B \in \mathcal{C}$. The condition (3.1) is obvious, and the main problem is to establish (3.2). First, we easily check that the functions $x \mapsto B(x, x \vee y, t)$ and $t \mapsto B(x, y, t)$ are nondecreasing. Consequently, in the proof of (3.2) we may restrict ourselves to the case $x = \alpha x_- + (1 - \alpha)x_+$ and $t = \alpha t_- + (1 - \alpha)t_+$. Since the region $\{(x, t) : x^p \leq t\}$ is convex, it is enough to prove the following. For any $h \in \mathbb{R}$ and any $(x, y, t) \in D$, the function

$$\varphi(s) = B(x + sh, (x + sh) \vee y, t + s)$$

(defined for those s, for which $(x + sh, (x + sh) \vee y, t + s) \in D)$ is concave. We start from observing that φ is of class C^1; this follows immediately from the fact that B is of class C^1 and $B_y(x, x, t) = 0$ (the latter condition guarantees that the one-sided derivatives $\varphi'(s-)$ and $\varphi'(s+)$ will match for $x + sh = y$). To deal with the concavity of φ on the set $\{s : x + sh \leq y\}$, we must prove that the matrix

$$\begin{bmatrix} B_{xx}(x + sh, y, t + s) & B_{xt}(x + sh, y, t + s) \\ B_{xt}(x + sh, y, t + s) & B_{tt}(x + sh, y, t + s) \end{bmatrix}$$

(defined for $y \neq (t + s)/(x + sh)$) is nonpositive-definite. Substituting $\tilde{x} = x + sh$ and $\tilde{t} = t + s$ if necessary, we may assume that $s = 0$. Now, if $y < (t/x)^{1/(p-1)}$, then the matrix equals

$$\begin{bmatrix} -\big((p-1)x\big)^{-1} & \big((p-1)t\big)^{-1} \\ \big((p-1)t\big)^{-1} & -x/\big((p-1)t^2\big) \end{bmatrix},$$

which clearly has the required property; if $y > (t/x)^{1/(p-1)}$, the situation is even simpler, since all the entries of the matrix are 0. Finally, it remains to show the concavity of φ on $\{s : x + sh > y\}$. Because $(x + sh)^p \le t + s$ (see the definition of D), we have $y < \big((t+s)/(x+sh)\big)^{1/(p-1)}$ and hence

$$\varphi(s) = x + sh + \frac{x+sh}{p-1} \log\left(\frac{(t+s)e}{(x+sh)^p}\right).$$

A direct differentiation yields

$$\varphi''(s) = -\frac{h^2}{x+sh} - \frac{(x-th)^2}{(p-1)(t+s)^2(x+sh)} \le 0$$

and the claim follows.

4.2 Proof of $\mathbb{B} \ge B$

Now we will use the second half of Theorem 3.1, which states that $\mathbb{B} \in \mathcal{C}$. We will also exploit the following additional homogeneity property of $\mathbb{B}$: for any $(x, y, t) \in D$ and $\lambda > 0$ we have

$$\mathbb{B}(\lambda x, \lambda y, \lambda^p t) = \lambda \mathbb{B}(x, y, t). \tag{4.1}$$

This condition follows at once from the very definition of $\mathbb{B}$ and the fact that $X \in \mathcal{S}(x, y, t)$ if and only if $\lambda X \in \mathcal{S}(\lambda x, \lambda y, \lambda^p t)$.

For the sake of clarity, we have split the reasoning into a few parts.

Step 1. Let δ be a small positive number. Using (3.2), we can write

$$\mathbb{B}(1, 1, 1) \ge \left(1 - \frac{1}{(1+\delta)^p}\right) \mathbb{B}(0, 1, 0) + \frac{1}{(1+\delta)^p} \mathbb{B}(1+\delta, 1+\delta, (1+\delta)^p). \tag{4.2}$$

Thus, using (3.1) and (4.1), the right-hand side is not smaller than

$$1 - \frac{1}{(1+\delta)^p} + \frac{1}{(1+\delta)^p} \mathbb{B}(1+\delta, 1+\delta, (1+\delta)^p) = 1 - \frac{1}{(1+\delta)^p} + \frac{\mathbb{B}(1, 1, 1)}{(1+\delta)^{p-1}}. \tag{4.3}$$

Combining the above facts, we get

$$\mathbb{B}(1, 1, 1) \ge \frac{(1+\delta)^p - 1}{(1+\delta)\big((1+\delta)^{p-1} - 1\big)},$$

so letting $\delta \to 0$ gives

$$\mathbb{B}(1, 1, 1) \ge \frac{p}{p-1}. \tag{4.4}$$

Step 2. Now we provide a lower bound for $\mathbb{B}(1, 1, t)$, where t is larger than 1. We argue as in the previous step, applying (3.2) and combining it with (3.1) and (4.1). Precisely, we fix a small positive δ and write

$$
\begin{aligned}
\mathbb{B}(1, 1, t) &\geq \frac{\delta}{1+\delta}\mathbb{B}(0, 1, 0) + \frac{1}{1+\delta}\mathbb{B}\big(1+\delta, 1+\delta, t(1+\delta)\big) \\
&\geq \frac{\delta}{1+\delta} + \mathbb{B}\left(1, 1, \frac{t}{(1+\delta)^{p-1}}\right).
\end{aligned}
\tag{4.5}
$$

By induction, this implies

$$
\mathbb{B}(1, 1, t) \geq \frac{n\delta}{1+\delta} + \mathbb{B}\left(1, 1, \frac{t}{(1+\delta)^{n(p-1)}}\right),
$$

if only $(1+\delta)^{n(p-1)} \leq t$. Now we fix a large positive integer n, put $\delta = t^{1/(n(p-1))} - 1$ (so that $(1+\delta)^{n(p-1)} = t$) and let $n \to \infty$. Then the above bound gives

$$
\mathbb{B}(1, 1, t) \geq \mathbb{B}(1, 1, 1) + \frac{\log t}{p-1},
$$

which combined with (4.4) yields

$$
\mathbb{B}(1, 1, t) \geq \frac{p}{p-1} + \frac{\log t}{p-1}.
\tag{4.6}
$$

Step 3. The next move in our analysis is to prove the estimate $\mathbb{B}(x, y, t) \geq B(x, y, t)$ for $y \leq (t/x)^{1/(p-1)}$. We proceed as previously: first apply (3.2) to obtain

$$
\mathbb{B}(x, y, t) \geq \frac{y-x}{y}\mathbb{B}(0, y, 0) + \frac{x}{y}\mathbb{B}\left(y, y, \frac{ty}{x}\right)
$$

(here we use the assumption $y \leq (t/x)^{1/(p-1)}$; if this inequality does not hold, the point $(y, y, ty/x)$ does not belong to D and $\mathbb{B}(y, y, ty/x)$ does not make sense). Next, using (3.1), (4.1) and finally (4.6), we get

$$
\begin{aligned}
\mathbb{B}(x, y, t) &\geq \frac{y-x}{y} \cdot y + \frac{x}{y} \cdot y\mathbb{B}\left(1, 1, \frac{t}{xy^{p-1}}\right) \\
&\geq y - x + x\left(\frac{p}{p-1} + \frac{\log(t/(xy^{p-1}))}{p-1}\right) = B(x, y, t).
\end{aligned}
$$

Step 4. Now we will deal with $\mathbb{B}(1, y, 1)$ for $y > 1$. By (3.2), (3.1) and (4.1), we have

$$
\begin{aligned}
\mathbb{B}(1, y, 1) &\geq (1 - y^{-p})\mathbb{B}(0, y, 0) + y^{-p}\mathbb{B}(y, y, y^p) \\
&\geq (1 - y^{-p})y + y^{1-p}\mathbb{B}(1, 1, 1).
\end{aligned}
$$

Combining this with (4.4), we obtain

$$\mathbb{B}(1, y, 1) \geq y + \frac{y^{1-p}}{p-1}. \tag{4.7}$$

Step 5. This is the final part. Pick $(x, y, t) \in D$ such that $y > (t/x)^{1/(p-1)}$ and apply (3.2) and then (3.1) to get

$$\mathbb{B}(x, y, t) \geq \alpha \mathbb{B}(0, y, 0) + (1 - \alpha)\mathbb{B}\left(\frac{x}{1-\alpha}, y, \frac{t}{1-\alpha}\right)$$

$$\geq \alpha y + \mathbb{B}(x, y(1-\alpha), t(1-\alpha)^{p-1}),$$

where $\alpha = 1 - (x^p/t)^{1/(p-1)}$. For this choice of α, we have $x^p = t(1-\alpha)^{p-1}$ and hence, by (4.1) and (4.7),

$$\mathbb{B}(x, y, t) \geq \alpha y + x\mathbb{B}\left(1, \frac{y(1-\alpha)}{x}, 1\right)$$

$$\geq \left(1 - \left(\frac{x^p}{t}\right)^{1/(p-1)}\right)y + x\left[\frac{1}{p-1}\left(\frac{x}{y}\right)^{p-1}\frac{t}{x^p} + \frac{y}{x}\left(\frac{x^p}{t}\right)^{1/(p-1)}\right]$$

$$= B(x, y, t).$$

This completes the proof of the inequality $\mathbb{B} \geq B$ on the whole domain. Thus, Theorem 1.1 follows.

4.3 On the construction of extremal examples

The arguments presented in Steps 1–5 can be easily translated into the construction of extremal supermartingales X ("extremizers") corresponding to $\mathbb{B}(x, y, t)$, i.e., those for which the supremum defining $\mathbb{B}(x, y, t)$ is almost attained. The purpose of this section is to explain how to extract this construction from the above calculations. The reasoning will be a little informal, as our aim is to present the *idea* of the connection.

First we look at the value $\mathbb{B}(1, 1, 1)$, which was studied in Step 1. What about the extremal $X \in \mathcal{S}(1, 1, 1)$? For a fixed $\delta > 0$, consider a Markov process $((X_n, Y_n, T_n))_{n \geq 0}$ starting from $(1, 1, 1)$, satisfying the following two requirements.

(i) Any point of the form $(\lambda, \lambda, \lambda^p)$ (with some $\lambda > 0$) leads to $(0, \lambda, 0)$ or to $(\lambda(1+\delta), \lambda(1+\delta), \lambda^p(1+\delta)^p)$ with probabilities $1 - (1+\delta)^{-p}$ and $(1+\delta)^{-p}$, respectively.

(ii) The states of the form $(0, \lambda, 0)$ are absorbing.

Then one can check that the process X is a supermartingale. If we stop it after a large number N of steps, we obtain the desired extremizer: that is, if we take sufficiently small δ and sufficiently large N, then $\mathbb{E}X_N^*$ can be made arbitrarily close to $\mathbb{B}(1, 1, 1)$. One might wonder why we have introduced the more complicated three-dimensional process (X, Y, T) above. The reason is that the moves described in (i) and (ii) are

closely related to the inequality (4.2) and the bound (4.3) on which Step 1 rests. To explain this more precisely, note first that (4.2) *encodes* the Markov move from (i): to make this more apparent, combine (4.2) with (4.1) to get

$$\mathbb{B}(\lambda, \lambda, \lambda^p) \geq \left(1 - \frac{1}{(1+\delta)^p}\right) \mathbb{B}(0, \lambda, 0)$$
$$+ \frac{1}{(1+\delta)^p} \mathbb{B}(\lambda(1+\delta), \lambda(1+\delta), \lambda^p(1+\delta)^p).$$

Thus, a starting state appears on the left, while the destination states can be found on the right, with the appropriate transition probabilities constituting the appropriate weights. The condition (ii) is connected to the bound (4.3): generally speaking, all the states (x, y, t) at which we use the majorization $\mathbb{B}(x, y, t) \geq y$ in the above considerations, need to be assumed absorbing.

To gain more intuition about this interplay, let us look at Step 2, which concerns the value of $\mathbb{B}(1, 1, t)$ with $t > 1$. We will construct a supermartingale $X \in \mathcal{S}(1, 1, t)$ for which the supremum defining $\mathbb{B}(1, 1, t)$ is almost attained. As previously, we fix $\delta > 0$ and actually handle an appropriate three-dimensional process $((X_n, Y_n, T_n))_{n \geq 0}$ starting from $(1, 1, t)$. To do this, combine the first inequality in (4.5) with (4.1) to obtain

$$\mathbb{B}(\lambda, \lambda, \lambda^p t) \geq \frac{\delta}{1+\delta} \mathbb{B}(0, \lambda, 0) + \frac{1}{1+\delta} \mathbb{B}(\lambda(1+\delta), \lambda(1+\delta), \lambda^p t(1+\delta)).$$

This gives us some information about the dynamics of the triple (X, Y, T). Namely, we impose the condition

(i') Any point of the form $(\lambda, \lambda, \lambda^p t)$ (with some $\lambda > 0$ and $t > 1$) leads to $(0, \lambda, 0)$ or to $(\lambda(1+\delta), \lambda(1+\delta), \lambda^p t(1+\delta))$ with probabilities $\delta/(1+\delta)$ and $1/(1+\delta)$, respectively.

Furthermore, in the second inequality of (4.5), we use the bound $\mathbb{B}(0, \lambda, 0) \geq \lambda$; this suggests that the requirement (ii) above should be valid.

Now, suppose that δ is chosen appropriately as in Step 2: $\delta = t^{1/(n(p-1))} - 1$ for some large positive integer n. Then, after n steps, the process (X, Y, T) gets to the point $((1+\delta)^n, (1+\delta)^n, (1+\delta)^{np})$ with positive probability. Now the procedure described in (i') does not apply since the point is not of appropriate form. In Step 2 we encountered a similar phenomenon: the number $\mathbb{B}(1, 1, 1)$ came into play and the arguments of Step 2 did not apply. To overcome this difficulty, we exploited Step 1. Here we do the same and apply the procedure (i) to the point $((1+\delta)^n, (1+\delta)^n, (1+\delta)^{np})$. In other words, the Markov process (X, Y, T) is given by the starting position $(1, 1, t)$ and the conditions (i), (ii) and (i'). It remains to stop the process X after a large number of steps to obtain the extremizer corresponding to $\mathbb{B}(1, 1, t)$.

The remaining extremal processes, corresponding to the values of $\mathbb{B}$ at remaining points, are found similarly. We leave the details to the reader.

5 Lack of prophet inequalities for L^p bounded variables

In the final part of the paper we show that if the condition (1.2) is replaced by

$$X_1, \ X_2, \ \ldots \ \text{are nonnegative and} \ \sup_n \mathbb{E}X_n^p \leq 1, \tag{5.1}$$

then no non-trivial prophet inequality holds. To prove this, we will exploit the results of Sect. 4. Fix an arbitrary positive number K. We have $\mathbb{B}(1, 1, t) \to \infty$ as $t \to \infty$, and hence there is a positive integer L and a finite nonnegative supermartingale $(Y_n)_{n=1}^N$ on $([0, 1], \mathbb{B}([0, 1]), |\cdot|)$ which satisfies $Y_1 \equiv 1$,

$$\sup_{\tau \in \mathcal{T}} \mathbb{E}Y_\tau^p \leq L \quad \text{and} \quad \mathbb{E}Y^* \geq \mathbb{E}Y_1 + K. \tag{5.2}$$

Now we construct L "distinct" copies of Y which evolve on pairwise disjoint time intervals. Precisely, consider a sequence $X = (X_n)_{n=1}^{LN}$ defined as follows:

- If $\omega \in [0, 1/L)$, then $X_n(\omega) = Y_n(L\omega)$ for $n = 1, 2, \ldots, N$ and $X_n(\omega) = 0$ for other n.
- If $\omega \in [1/L, 2/L)$, then $X_n(\omega) = Y_{n-N}(L\omega - 1)$ for $n = N+1, N+2, \ldots, 2N$ and $X_n(\omega) = 0$ for other n.
- If $\omega \in [2/L, 3/L)$, then $X_n(\omega) = Y_{n-2N}(L\omega - 2)$ for $n = 2N+1, 2N+2, \ldots, 3N$ and $X_n(\omega) = 0$ for other n.
- …
- If $\omega \in [1 - 1/L, 1)$, then $X_n(\omega) = Y_{n-(L-1)N}(L\omega - (L-1))$ for $n = (L-1)N+1, (L-1)N+2, \ldots, LN$ and $X_n(\omega) = 0$ for other n.

The variables $X_1, X_2, \ldots, X_{LN}$ are nonnegative and enjoy the following properties. First, note that the conditional distribution of X_{mN+n} on $[m/L, (m+1)/L)$ coincides with the distribution of Y_n, so $\mathbb{E}X_{mN+n}^p = \mathbb{E}Y_n^p/L$ and thus

$$\sup_{1 \leq n \leq LN} \mathbb{E}X_n^p = \sup_{1 \leq n \leq N} \mathbb{E}Y_n^p/L \leq 1,$$

by the first inequality in (5.2). Next, if $\omega \in [m/L, (m+1)/L)$, we have $X^*(\omega) = Y^*(L\omega - m)$ and hence $\mathbb{E}X^* = \mathbb{E}Y^*$. Finally, we will prove that $\sup_{\tau \in \mathcal{T}} \mathbb{E}X_\tau = \mathbb{E}Y_1$. The inequality "$\geq$" follows by considering the stopping time given by $\tilde{\tau}(\omega) = mL+1$ for $\omega \in [m/L, (m+1)/L)$. To prove the reverse bound, note that while computing $\sup_{\tau \in \mathcal{T}} \mathbb{E}X_\tau$ we may restrict to stopping times which on each $[m/L, (m+1)/L)$ take values $\{mL+1, mL+2, \ldots, (m+1)L\}$. Indeed, if τ is an arbitrary stopping time, then $\bar{\tau} = (\tau \vee \tilde{\tau}) \wedge (\tilde{\tau} + L - 1)$ has this property and $\mathbb{E}X_\tau \leq \mathbb{E}X_{\bar{\tau}}$. Therefore, since Y is a supermartingale, we have

$$\mathbb{E}(X_\tau | [m/L, (m+1)/L)) \leq \mathbb{E}(X_{mL+1} | [m/L, (m+1)/L)) = \mathbb{E}Y_1$$

and hence $\mathbb{E}X_\tau \leq \mathbb{E}Y_1$.

Consequently, by the second estimate in (5.2), the sequence X satisfies

$$\mathbb{E}X^* \geq \sup_{\tau \in \mathcal{T}} \mathbb{E}X_\tau + K.$$

Since K was arbitrary, no universal prophet inequality holds under (5.1). This completes the proof.

References

1. Allaart, P.C.: Prophet inequalities for I.I.D. random variables with random arrival times. Seq. Anal. **26**, 403–413 (2007)
2. Allaart, P.C., Monticino, M.G.: Pseudo-prophet inequalities in average-optimal stopping. Seq. Anal. **22**, 233–239 (2003)
3. Boshuizen, F.: Multivariate Prophet Inequalities for Negatively Dependent Random Vectors. Strategies for sequential search and selection in real time (Amherst, MA, 1990), Contemporary Mathematics, vol. 125, pp. 183–190. American Mathematical Society, Providence (1992)
4. Assaf, D., Goldstein, L., Samuel-Cahn, E.: Ratio prophet inequalities when the mortal has several choices. Ann. Appl. Probab. **12**(3), 972–984 (2002)
5. Assaf, D., Samuel-Cahn, E.: Prophet inequalities for optimal stopping rules with probabilistic recall. Bernoulli **8**(1), 39–52 (2002)
6. Boshuizen, F.: Prophet compared to gambler: additive inequalities for transforms of sequences of random variables. Statist. Probab. Lett. **29**, 23–32 (1996)
7. Burkholder, D.L.: Explorations in Martingale Theory and its Applications. Ecole d'Ete de Probabilités de Saint-Flour XIX–1989, Lecture Notes in mathematics, vol. 1464, pp. 1–66. Springer, Berlin (1991)
8. Hill, T.P.: Prophet inequalities and order selection in optimal stopping problems. Proc. Am. Math. Soc. **88**, 131–137 (1983)
9. Hill, T.P.: Prophet inequalities for averages of independent non-negative random variables. Math. Z. **192**, 427–436 (1986)
10. Hill, T.P., Kertz, R.P.: Additive comparisons of stop rule and supremum expectations of uniformly bounded independent random variables. Proc. Am. Math. Soc. **83**, 582–585 (1981)
11. Hill, T.P., Kertz, R.P.: Stop rule inequalities for uniformly bounded sequences of random variables. Trans. Am. Math. Soc. **278**, 197–207 (1983)
12. Hill, T.P., Kertz, R.P.: A Survey of Prophet Inequalities in Optimal Stopping Theory. Strategies for sequential search and selection in real time (Amherst, MA, 1990), Contemporary Mathematics, vol. 125, pp. 191–207. American Mathematical Society, Providence (1992)
13. Kennedy, R.P.: Optimal stopping of independent random variables and maximizing prophets. Ann. Probab. **13**, 566–571 (1985)
14. Kennedy, R.P.: Prophet-type inequalities for multi-choice optimal stopping. Stoch. Proc. Appl. **24**, 77–88 (1987)
15. Kennedy, D.P., Kertz, R.P.: A prophet inequality for independent random variables with finite variances. J. Appl. Probab. **34**(4), 945–958 (1997)
16. Kertz, R.P.: Comparison of optimal value and constrained maxima expectations for independent random variables. Adv. Appl. Prob. **18**, 311–340 (1986)
17. Krengel, U., Sucheston, L.: Semiamarts and finite values. Bull. Am. Math. Soc. **83**, 745–747 (1977)
18. Krengel, U., Sucheston, L.: On semiamarts, amarts, and processes with finite value. In: Kuelbs, J. (ed.) Probability on Banach Spaces. Marcel Dekker, New York (1978)

19. Krengel, U., Sucheston, L.: Prophet compared to the gambler: an inequality for transforms of processes. Ann. Probab. **15**, 1593–1599 (1987)
20. Nazarov, F., Treil, S.: The hunt for Bellman function: applications to estimates of singular integral operators and to other classical problems in harmonic analysis. Algebra i Analis **8**, 32–162 (1997)
21. Osękowski, A.: Sharp Martingale and Semimartingale Inequalities, vol. 72. Monografie Matematyczne, Birkhäuser (2012)
22. Peskir, G., Shiryaev, A.N.: Optimal Stopping and Free-boundary Problems. Lectures in Mathematics ETH Zürich. Birkhäuser Verlag, Basel (2006)
23. Shiryaev, A.N.: Optimal Stopping Rules. Translated from the 1976 Russian second edition by A.B. Aries. Reprint of the 1978 translation. Stochastic Modelling and Applied Probability, vol. 8. Springer-Verlag, Berlin (2008)
24. Tanaka, T.: Prophet inequalities for two-parameter optimal stopping problems. J. Inf. Optim. Sci. **28**, 159–172 (2007)
25. Tanaka, T.: Lower semicontinuity property of multiparameter optimal stopping value and its application to multiparameter prophet inequalities. J. Math. Anal. Appl. **359**, 240–251 (2009)
26. Vasyunin, V., Volberg, A.: Monge-Ampére Equation and Bellman Optimization of Carleson Embedding Theorems. Linear and complex analysis. American Mathematical Society Translations: Series 2, vol. 226, pp. 195–238. American Mathematical Society, Providence (2009)