Martingale Inequalities and Deterministic Counterparts

Mathias Beiglböck[*]   Marcel Nutz[†]

October 15, 2014

Abstract

We study martingale inequalities from an analytic point of view and show that a general martingale inequality can be reduced to a pair of deterministic inequalities in a small number of variables. More precisely, the optimal bound in the martingale inequality is determined by a fixed point of a simple nonlinear operator involving a concave envelope. Our results yield an explanation for certain inequalities that arise in mathematical finance in the context of robust hedging.

1   Introduction

Martingale inequalities are abundant in many areas of probability theory and analysis; see e.g. Burkholder's survey [11] for an extensive list of literature. We study general inequalities for discrete-time martingales from a bird's eye view and relate them to certain deterministic inequalities. Indeed, we shall see that every martingale inequality can be obtained as a consequence of

two deterministic ones, and in fact that martingale inequalities are not very probabilistic in nature.

A simple example of a martingale inequality is Doob's maximal quadratic inequality, stating that the running maximum $M_T^* := \sup_{0 \le t \le T} |M_t|$ of any martingale M satisfies

$$\|M_T^*\|_2 \le 2\|M_T\|_2,$$

where $\|\cdot\|_2$ is the L^2-norm. We may cast this in the form $E[f(M_T, M_T^*)] \le 0$ for a suitable function f; namely, $f(x, y) = y^2 - 4|x|^2$. The general form of the martingale inequality that we shall consider is

$$E[f(Z_T)] \le a, \tag{1.1}$$

where a is a constant and $Z = (Z_t)_{t \in \mathbb{N}}$ is a suitable state process defined as a function of M; in the preceding example, $Z = (M, M^*)$. More precisely, let $\mathbf{X}$ be a vector space (in which our martingales are taking values) and let $\mathbf{Z}$ be a set, to be used as the state space. Then the $\mathbf{Z}$-valued process Z is determined by a function $\phi : \mathbf{Z} \times \mathbf{X} \to \mathbf{Z}$ via $Z_{t+1} = \phi(Z_t, M_{t+1} - M_t)$ and some initial value z_0. Again in the example, $\phi(x, y, d) = (x+d, y \vee |x+d|)$ updates M by adding the next increment and increases the running maximum if necessary.

Given $f : \mathbf{Z} \to \overline{\mathbb{R}}$ and ϕ, we may ask if there exists a finite constant a such that (1.1) holds for all $T \in \mathbb{N}$ and all martingales with prescribed initial value, and what the optimal (minimal) value for a is. A possible answer runs as follows. Consider the operator A which acts on functions $g : \mathbf{Z} \to \overline{\mathbb{R}}$ by pre-composing with ϕ and taking the concave envelope at the origin in the variable corresponding to the martingale increment:

$$Ag(z) = g(\phi(z, \cdot))^\sharp(0), \quad z \in \mathbf{Z}.$$

If u is a fixed point of A dominating f; that is, $Au = u$ and $u \ge f$, then $a = u(z_0)$ is an admissible constant in (1.1). Under the natural condition $\phi(z, 0) = z$, a simple monotonicity argument shows that A has a *minimal* fixed point u dominating f. This fixed point can be obtained from f by iterating A and passing to the limit,

$$u = A^\infty f := \lim_{n \to \infty} A^n f,$$

and we shall see that $a = u(z_0)$ is the optimal constant in (1.1). In this sense, we may say that a martingale inequality can be reduced to the two deterministic inequalities

$$u \ge Au \quad \text{and} \quad u \ge f.$$

(Here $u \geq Au$ is actually equivalent to $Au = u$.) In fact, we may note that u defines a stronger martingale inequality altogether. Namely, as $u \geq f$,

$$E[u(Z_T)] \leq u(z_0)$$

is stronger than the original inequality $E[f(Z_T)] \leq u(z_0)$ with optimal constant, and we remark that the inequality $u \geq f$ is strict in most cases of interest. Returning to our example, we can check that the minimal fixed point is given by

$$u(x, y) = \begin{cases} y^2 - 4|x|^2 & \text{if } |x| < y/2, \\ 2y^2 - 4|x|y & \text{if } |x| \geq y/2, \end{cases}$$

and so the optimal constant corresponding to the initial value $z_0 = (x_0, |x_0|)$ is $a = -2|x_0|^2$, while the knowledge of u actually yields a further strengthening of Doob's maximal inequality (Corollary 4.3). There are of course very relevant martingale inequalities which hold only for some specific class of martingales; for instance, nonnegative martingales, martingales with increments bounded by one, etc. Many such inequalities can be fitted within our framework by choosing $\mathbf{Z}$ appropriately and assigning the value $-\infty$ to the function f on a suitable subset (see also Section 4.2).

All this has little to do with probability or measure theory; in fact, it seems that the latter is only needed to define the expectations. In order to clearly separate this aspect (and also to spare the reader some measurable selection arguments), we shall develop the theory for simple martingales (i.e., martingales taking finitely many values), so that all expectations are actually finite sums. In most cases of interest, f and ϕ (and then also the fixed point u) have some continuity properties and the passage to general martingales can be done a posteriori by approximation. However, we also provide an alternative argument which is more in the spirit of this paper and applies even to functions that are merely measurable, under the restriction that $\mathbf{X}$ be finite-dimensional. Namely, we devise a martingale version of Tchakaloff's theorem, stating that given a measurable (integrable) function $g : (\mathbb{R}^n)^T \to \mathbb{R}$ and an n-dimensional martingale M, we can find a *simple* martingale N such that

$$E[g(N_1, \ldots, N_T)] = E[g(M_1, \ldots, M_T)],$$

and moreover the (finite) support of the law of N lies in the support of the law of M. Note that we have here an actual equality; no approximation is necessary.

The theory outlined in this paper can be seen as a general formulation of a strategy of proof that was used in several works of D. L. Burkholder for martingale inequalities where Z consists of X, its running maximum and its square function. Namely, he used a class of functions u, corresponding roughly to what we call fixed points, to find admissible or sharp constants in various martingale inequalities, and in fact it seems that he was aware of at least part of the structure presented here; see in particular Theorem 2.1 in [13] but also [10, 11, 12], among others, as well as the recent monograph and review article of Osękowski [22, 23].

A different stream of literature about martingale inequalities has emerged in mathematical finance, starting with Hobson [17]. In this context, the process X takes values in $\mathbb{R}^n$ and represents the discounted prices of n tradable securities, while $f(Z_T)$ is seen as an option maturing at the fixed time horizon T. The problem is to find a minimal constant a and a predictable process H (i.e., H_t is a function of $X_0, \ldots, X_{t-1}$) such that

$$a + \sum_{t=1}^{T} \langle H_t, X_t - X_{t-1} \rangle \geq f(Z_T), \tag{1.2}$$

where the inner product $\langle H_t, X_t - X_{t-1} \rangle$ is interpreted as the gain or loss that occurs as the price X_{t-1} changes to X_t while H_t units of the security are held. Thus, if a is charged as the price of the option, the trading strategy H allows to hedge the risk of $f(Z_T)$ in a robust (model-free) way. We observe that by taking expectations on both sides, (1.2) implies the martingale inequality $E[f(Z_T)] \leq a$. Along these lines, "pathwise" proofs for several martingale inequalities have been obtained. For these and related results in robust finance, see [1, 2, 4, 5, 7, 14, 16, 21] among others; more references can be found in the surveys by Hobson [18] and Obłój [20]. In particular, a result of Bouchard and Nutz [6] implies that any martingale inequality in finite discrete time can be related to an inequality of the type (1.2). However, the machinery used there (to deal with a more general case) only yields a non-constructive existence result for H and little insight into the nature of the inequality. We shall see that, in essence, H is determined quite explicitly as the derivative of $u(\phi(z, \cdot))^\sharp$.

The remainder of this article is organized as follows. In Section 2 we consider martingale inequalities with a fixed time horizon T and relate the optimal constant to certain concave envelopes by dynamic programming. Section 3 focuses on martingale inequalities that do not depend explicitly on the time horizon T; this further condition of time-homogeneity leads to the fixed point considerations mentioned above. The connection to mathe-

4

matical finance is also discussed here. In Section 4, we illustrate the theory by two simple examples, Doob's maximal L^p-inequality and Burkholder's inequality for differentially subordinate martingales. Section 5 concludes with the martingale version of Tchakaloff's theorem.

2 Martingale Inequalities and Concave Envelopes

It will be convenient to work with functions taking values in the extended real line $\overline{\mathbb{R}} = [-\infty, \infty]$. The convention

$$\infty - \infty = -\infty \tag{2.1}$$

is used throughout; in particular, in the definitions of concavity and integrals. Let $\mathbf{X}$ be a real[1] vector space. Given a function $g : \mathbf{X} \to \overline{\mathbb{R}}$, we define its concave envelope $g^\sharp : \mathbf{X} \to \overline{\mathbb{R}}$ as the smallest concave function dominating g, or

$$g^\sharp(x) = \inf\{\psi(x) \,|\, \psi : \mathbf{X} \to \overline{\mathbb{R}} \text{ is concave and } \psi \geq g\}, \quad x \in \mathbf{X}.$$

We shall need to take consecutive envelopes over several variables. Given an integer $t \geq 0$ and $g : \mathbf{X}^{t+2} \to \overline{\mathbb{R}}$, we first introduce the function

$$g^{\sharp_t} : \mathbf{X}^{t+1} \to \overline{\mathbb{R}}, \quad g^{\sharp_t}(x_0, \ldots, x_t) := g(x_0, \ldots, x_t, \cdot)^\sharp(x_t);$$

in other words, we pass to the concave envelope in the ultimate variable and evaluate the resulting function at the penultimate variable. For an integer $T \geq 0$, we can then define the composition

$$\sharp(T) = \sharp_0 \circ \cdots \circ \sharp_{T-1}$$

which maps functions of $T + 1$ variables into functions of one variable.

Our first aim is to identify, for a fixed time horizon T, the optimal constant for a martingale inequality defined by $f : \mathbf{X}^{T+1} \to \overline{\mathbb{R}}$ in terms of the consecutive envelope $f^{\sharp(T)}$. Given $x_0 \in \mathbf{X}$, we shall denote by $\mathcal{M}^T(x_0)$ the set of all laws of $\mathbf{X}$-valued simple martingales $M_0, \ldots, M_T$ satisfying $M_0 = x_0$. Note that any expectation $E[f(M_0, \ldots, M_T)]$ on the original probability space of the martingale M can be expressed as the expectation $\mu[f] := E_\mu[f] := \int f \, d\mu$ of f under the law μ of M; the latter point of view will be more convenient in the sequel. We emphasize that the integral under $\mu \in \mathcal{M}^T(x_0)$ is a finite sum and therefore does not require any measurability conditions, and moreover that, according to (2.1), we have $\mu[f] = -\infty$ if $\mu[f^+] = \mu[f^-] = \infty$.

---

[1]The general case is no more difficult than $\mathbf{X} = \mathbb{R}$. Moreover, most of what follows applies to the complex case without change.

Proposition 2.1. *Let $f : \mathbf{X}^{T+1} \to \overline{\mathbb{R}}$. Then*

$$f^{\sharp(T)}(x_0) = \sup_{\mu \in \mathcal{M}^T(x_0)} \mu[f], \quad x_0 \in \mathbf{X}. \tag{2.2}$$

Or, to state the same in different words: proving that an inequality $E[f(M_0, \dots, M_T)] \leq a$ holds for all martingales M starting at x_0 boils down to checking that $f^{\sharp(T)}(x_0) \leq a$, and in fact $f^{\sharp(T)}(x_0)$ is the optimal constant.

As a first step towards the proof, we consider the case $T = 1$. Noting that $\mathcal{M}(x)$ is simply the set of all probability measures μ on $\mathbf{X}$ having finite support and barycenter $\mu[\mathrm{Id}_\mathbf{X}] = x$, the following identity is essentially classical (see Kemperman [19]); we state the details for the sake of completeness.

Lemma 2.2. *Let $g : \mathbf{X} \to \overline{\mathbb{R}}$. Then*

$$\sup_{\mu \in \mathcal{M}(x)} \mu[g] = g^{\sharp}(x), \quad x \in \mathbf{X}.$$

Proof. Let $x \in \mathbf{X}$ and $\mu \in \mathcal{M}(x)$; then μ is a convex combination of Dirac measures, $\mu = \sum_{i=1}^n \lambda_i \delta_{x_i}$, with $\sum \lambda_i x_i = x$. In particular,

$$\mu[g] \leq \mu[g^{\sharp}] = \sum \lambda_i g^{\sharp}(x_i) \leq g^{\sharp}(x)$$

as $g^{\sharp}$ is concave, showing that $\sup_{\mu \in \mathcal{M}(x)} \mu[g] \leq g^{\sharp}(x)$. To see the converse inequality, let $x_1, x_2 \in \mathbf{X}$ and $\lambda \in (0, 1)$. Given $\varepsilon > 0$, there are $\mu_i^{\varepsilon} \in \mathcal{M}(x_i)$ such that (with $a \wedge b := \min\{a, b\}$)

$$\mu_i^{\varepsilon}[g] \geq \varepsilon^{-1} \wedge \sup_{\mu \in \mathcal{M}(x_i)} \mu[g] - \varepsilon.$$

Using the fact that $\lambda \mu_1^{\varepsilon} + (1 - \lambda)\mu_2^{\varepsilon} \in \mathcal{M}(\lambda x_1 + (1 - \lambda)x_2)$, we then have

$$\lambda \sup_{\mu \in \mathcal{M}(x_1)} \mu[g] + (1 - \lambda) \sup_{\mu \in \mathcal{M}(x_2)} \mu[g] \leq \limsup_{\varepsilon \to 0} \lambda \mu_1^{\varepsilon}[g] + (1 - \lambda)\mu_2^{\varepsilon}[g]$$

$$\leq \sup_{\mu \in \mathcal{M}(\lambda x_1 + (1-\lambda)x_2)} \mu[g],$$

showing that $x \mapsto \sup_{\mu \in \mathcal{M}(x)} \mu[g]$ is concave. In view of $\sup_{\mu \in \mathcal{M}(x)} \mu[g] \geq \delta_x[g] = g(x)$, the definition of $g^{\sharp}(x)$ now yields $\sup_{\mu \in \mathcal{M}(x)} \mu[g] \geq g^{\sharp}(x)$. $\quad\square$

The extension to the case of a general horizon T can be understood as a dynamic programming argument where the martingale laws play the role

of the controls in a stochastic control problem. Given $g : \mathbf{X}^{t+2} \to \overline{\mathbb{R}}$, we therefore introduce the (value) function

$$\mathcal{E}_t(g) : \mathbf{X}^{t+1} \to \overline{\mathbb{R}}, \quad \mathcal{E}_t(g)(x_0, \ldots, x_t) := \sup_{\mu \in \mathcal{M}(x_t)} \mu[g(x_0, \ldots, x_t, \cdot)],$$

as well as the composition

$$\mathcal{E}^t := \mathcal{E}_t \circ \cdots \circ \mathcal{E}_{T-1}$$

which maps functions of $T + 1$ variables into functions of $t + 1$ variables.

Lemma 2.3. *Let* $f : \mathbf{X}^{T+1} \to \overline{\mathbb{R}}$. *Then*

$$(\mathcal{E}_0 \circ \cdots \circ \mathcal{E}_{T-1})(f)(x_0) = \sup_{\mu \in \mathcal{M}^T(x_0)} \mu[f], \quad x_0 \in \mathbf{X}. \qquad (2.3)$$

Proof. We first suppose that f is bounded from above. To see the inequality "$\leq$", let $\varepsilon > 0$. For all $0 \leq t < T$ and $(x_0, \ldots, x_t) \in \mathbf{X}^{t+1}$, let $\mu_t(x_0, \ldots, x_t) \in \mathcal{M}(x_t)$ be such that

$$\mu_t(x_0, \ldots, x_t)[\mathcal{E}^{t+1}(f)(x_0, \ldots, x_t, \cdot)] \geq \sup_{\mu \in \mathcal{M}(x_t)} \mu[\mathcal{E}^{t+1}(f)(x_0, \ldots, x_t, \cdot)] - \varepsilon.$$

We may see μ_t as a stochastic kernel on $\mathbf{X}^{t+1}$ equipped with the discrete σ-field. Recalling that we are only using measures with finite support, we may form the product measure $\mu^\varepsilon := (\mu_0 \otimes \cdots \otimes \mu_{T-1})(x_0)$ which is an element of $\mathcal{M}^T(x_0)$ by Fubini's theorem. We then have

$$(\mathcal{E}_0 \circ \cdots \circ \mathcal{E}_{T-1})(f)(x_0) \leq \varepsilon T + \mu^\varepsilon[f] \leq \varepsilon T + \sup_{\mu \in \mathcal{M}^T(x_0)} \mu[f].$$

As $\varepsilon > 0$ was arbitrary, this yields the claimed inequality. To see the converse inequality "$\geq$", fix $x_0 \in \mathbf{X}$ and note that any $\mu \in \mathcal{M}^T(x_0)$ can be decomposed into the product $\mu = \mu_0 \otimes \mu_1 \otimes \cdots \otimes \mu_{T-1}$ of a measure $\mu_0 \in \mathcal{M}(x_0)$ and kernels μ_t on $\mathbf{X}^t$ such that $\mu_t(x_1, \ldots, x_t) \in \mathcal{M}(x_t)$ for all $x_1, \ldots, x_t \in \mathbf{X}$. By the definition of the operators $\mathcal{E}_t$, we then have

$$(\mathcal{E}_0 \circ \cdots \circ \mathcal{E}_{T-1})(f)(x_0) \geq (\mu_0 \otimes \mu_1 \otimes \cdots \otimes \mu_{T-1})[f] = \mu[f]$$

and the claim follows as $\mu \in \mathcal{M}^T(x_0)$ was arbitrary.

Finally, for the case of a general function f, we observe that both sides of (2.3) are continuous along increasing sequences (f_n) of $\overline{\mathbb{R}}$-valued functions having the property that $\{f_n = -\infty\} = \{f_{n+1} = -\infty\}$, $n \geq 1$. Thus, we may apply the above to $f \wedge n$ and pass to the limit as $n \to \infty$. $\qquad \square$

Proof of Proposition 2.1. Since Lemma 2.2 shows that $\sharp_t = \mathcal{E}_t$, Proposition 2.1 is a direct consequence of Lemma 2.3. $\qquad \square$

3 Time-Homogeneous Martingale Inequalities

Let $x_0 \in \mathbf{X}$, set $X_0 = x_0$ and let $(X_t)_{t=1,2,\ldots}$ be the coordinate-mapping process on $\mathbf{X} \times \mathbf{X} \times \cdots$. Moreover, let $\mathbf{Z}$ be a nonempty set and fix a function $\phi : \mathbf{Z} \times \mathbf{X} \to \mathbf{Z}$. Given $z_0 \in \mathbf{Z}$, we define the $\mathbf{Z}$-valued process $Z = (Z_t)_{t=0,1,\ldots}$ by

$$Z_0 = z_0, \quad Z_{t+1} = \phi(Z_t, X_{t+1} - X_t).$$

We write $\overline{\mathbb{R}}^{\mathbf{Z}}$ for the set of all functions $\mathbf{Z} \to \overline{\mathbb{R}}$, equipped with the pointwise partial order and convergence, and define the operator $A : \overline{\mathbb{R}}^{\mathbf{Z}} \to \overline{\mathbb{R}}^{\mathbf{Z}}$ via

$$Ag(z) := [g \circ \phi(z, \cdot)]^{\sharp}(0), \quad z \in \mathbf{Z}.$$

Moreover, we write A^T for the T-fold composition $A \circ \cdots \circ A$. Using this notation, Proposition 2.1 can be rephrased as follows.

Lemma 3.1. *Let $f : \mathbf{Z} \to \overline{\mathbb{R}}$ and let $(x_0, z_0) \in \mathbf{X} \times \mathbf{Z}$. Then*

$$A^T f(z_0) = \sup_{\mu \in \mathcal{M}^T(x_0)} \mu[f(Z_T)].$$

This lemma may look less general than Proposition 2.1, which allows for a general dependence on the path of X, but let us mention that with the choice $\mathbf{Z} = \mathbb{N} \times \mathbf{X}^{\mathbb{N}}$ we can arrange things so that $Z_t = (t, X_0, X_1, \ldots, X_t, 0, 0, \ldots)$.

From now on, we focus on martingale inequalities which hold for any time horizon T. The structural condition

$$\phi(z, 0) = z, \quad z \in \mathbf{Z} \tag{3.1}$$

seems to be natural in that setting and we make this a *standing assumption*. The operator A then has the following monotonicity properties.

Lemma 3.2. *Let $g, g' : \mathbf{Z} \to \overline{\mathbb{R}}$. Then*

(i) $Ag \geq g$;

(ii) $g \geq g'$ *implies* $Ag \geq Ag'$.

Proof. In view of (3.1), we have

$$Ag(z) = g(\phi(z, \cdot))^{\sharp}(0) \geq g(\phi(z, 0)) = g(z), \quad z \in \mathbf{Z}.$$

The second property follows from the monotonicity of $\sharp$. $\qquad\square$

Theorem 3.3. *Let $f : \mathbf{Z} \to \overline{\mathbb{R}}$. Then the limit*

$$A^\infty f(z) := \lim_{n \to \infty} A^n f(z), \quad z \in \mathbf{Z}$$

exists in $\overline{\mathbb{R}}$ and the function $A^\infty f \in \overline{\mathbb{R}}^{\mathbf{Z}}$ is characterized as the smallest fixed point of A which dominates f.

Remark 3.4. By Lemma 3.1, $A^\infty f(z_0)$ is the optimal *horizon-independent* constant for the martingale inequality determined by f, ϕ and z_0. In fact, Lemma 3.1 naturally extends to

$$A^\infty f(z_0) = \sup_{\mu \in \mathcal{M}^\infty(x_0)} \mu[f(Z_\infty)]$$

if we denote by $\mathcal{M}^\infty(x_0)$ the set of all laws of $\mathbf{X}$-valued simple[2] martingales $(M_t)_{t \in \mathbb{N}}$ satisfying $M_0 = x_0$. Note that any such martingale is eventually constant, so that $Z_\infty := \lim_n Z_n$ is well-defined μ-a.s. for all $\mu \in \mathcal{M}^\infty(x_0)$.

Proof of Theorem 3.3. It follows from Lemma 3.2 that

$$f \le Af \le \cdots \le A^n f, \quad n \ge 1.$$

In particular, the limit $A^\infty f(z) := \lim_{n \to \infty} A^n f(z) \in \overline{\mathbb{R}}$ exists for all $z \in \mathbf{Z}$. Next, let us observe that if $(g_n)_{n \ge 1} \subseteq \overline{\mathbb{R}}^{\mathbf{Z}}$ is a nondecreasing sequence, then

$$\lim_n g_n^\sharp = (\lim_n g_n)^\sharp.$$

Indeed, both limits are increasing and thus well-defined, and the monotonicity of $\sharp$ immediately implies that $\lim_n g_n^\sharp \le (\lim g_n)^\sharp$. Conversely, $\lim_n g_n^\sharp$ is concave as the pointwise limit of concave functions and dominates $\lim g_n$, so that $\lim_n g_n^\sharp \ge (\lim g_n)^\sharp$. Using this continuity property of $\sharp$, we see that

$$A^\infty f(z) = \lim_n A^{n+1} f(z) = \lim_n [A^n f \circ \phi(z, \cdot)]^\sharp(0) = [\lim_n A^n f \circ \phi(z, \cdot)]^\sharp(0)$$
$$= [A^\infty f \circ \phi(z, \cdot)]^\sharp(0) = A A^\infty f(z)$$

for all $z \in \mathbf{Z}$; that is, $A^\infty f$ is a fixed point. If $g \in \overline{\mathbb{R}}^{\mathbf{Z}}$ is another fixed point of A such that $g \ge f$, then the monotonicity of A from Lemma 3.2 yields that

$$g = A^n g \ge A^n f, \quad n \ge 0$$

and hence $g \ge A^\infty f$ by passing to the limit. $\square$

---

[2]"Simple" means that the support is a finite subset of $\mathbf{X}^{\mathbb{N}}$.

Remark 3.5. Let $u : \mathbf{Z} \to \overline{\mathbb{R}}$ be any function such that $f \leq u$ and $Au \leq u$ (hence $Au = u$; cf. Lemma 3.2). Then

$$\sup_{\mu \in \mathcal{M}^\infty(x_0)} \mu[f(Z_\infty)] = A^\infty f(z_0) \leq A^\infty u(z_0) = u(z_0);$$

that is, to prove that the martingale inequality holds with right-hand side a, it suffices to exhibit a fixed point u of A which dominates f and satisfies $u(z_0) \leq a$. As mentioned in the Introduction, this corresponds to a general formulation of the strategy of proof that has been used by Burkholder for several specific martingale inequalities. For the above conclusion, it is not necessary to establish that u is the minimal fixed point; however, this property guarantees that $u(z_0)$ is the optimal right-hand side.

To find an explicit formula for $A^\infty f$ (or any other fixed point), it is often useful to study properties of f that are preserved by A. We give a simple example to illustrate this point (see also Section 4.1).

Remark 3.6. Suppose that $\mathbf{Z}$ is a cone and that ϕ is positively homogeneous of degree one. If $f \in \overline{\mathbb{R}}^{\mathbf{Z}}$ is positively homogeneous of degree $p > 0$, then so are Af and $A^\infty f$. Indeed, let $\lambda \geq 0$; then

$$Af(\lambda z) = f(\phi(\lambda z, \cdot))^\sharp(0) = \inf\{\psi(0) \,|\, \psi(\lambda \cdot) \geq f(\phi(\lambda z, \lambda \cdot))\}$$
$$= \inf\{\psi(0) \,|\, \psi \geq \lambda^p f(\phi(z, \cdot))\} = \inf\{\lambda^p \psi(0) \,|\, \psi \geq f(\phi(z, \cdot))\} = \lambda^p Af(z),$$

where the infima are taken over all concave functions $\psi : \mathbf{X} \to \overline{\mathbb{R}}$. The homogeneity of $A^\infty f$ follows.

Next, we would like to explain a connection to certain inequalities which have arisen in mathematical finance—from our abstract point of view, we shall see that the latter are simply manifestations of the concavity that is imposed by A. For the purpose of the subsequent discussion, we assume that we are given a dual pair $\mathbf{X}, \mathbf{X}'$ with a separating pairing $\langle \cdot, \cdot \rangle$. Given a concave function $h : \mathbf{X} \to \overline{\mathbb{R}}$, the supergradient $\partial h(d_0)$ at $d_0 \in \mathbf{X}$ is defined as the set of all $\xi \in \mathbf{X}'$ such that $h(d_0) + \langle \xi, d - d_0 \rangle \geq h(d)$ for all $d \in \mathbf{X}$, and h is called superdifferentiable at d_0 if this set is nonempty.

Lemma 3.7. *Let $g : \mathbf{Z} \to \overline{\mathbb{R}}$. Each of the following conditions implies the subsequent one:*

(i) For all $z \in \mathbf{Z}$ there exists $\xi(z) \in \mathbf{X}'$ such that

$$g(\phi(z, d)) \leq g(z) + \langle \xi(z), d \rangle, \quad d \in \mathbf{X}. \tag{3.2}$$

(ii) $Ag = g$.

(iii) *For all* $z \in \mathbf{Z}$ *and all* $\xi(z) \in \partial g(\phi(z, \cdot))^\sharp(0)$,

$$g(\phi(z, d)) \leq g(z) + \langle \xi(z), d \rangle, \quad d \in \mathbf{X}.$$

If the concave function $g(\phi(z, \cdot))^\sharp$ *is superdifferentiable at* $d = 0$, *these conditions are equivalent. In particular, the conditions are equivalent if* $\mathbf{X}$ *is finite-dimensional and* $g(\phi(z, \cdot))^\sharp$ *is finite-valued.*

Proof. Let (i) hold. Taking concave envelopes on both sides of (3.2), we see that

$$Ag(z) = g(\phi(z, \cdot))^\sharp(0) \leq g(z) + \langle \xi(z), 0 \rangle = g(z),$$

which implies (ii) by Lemma 3.2. Let $\xi(z) \in \partial g(\phi(z, \cdot))^\sharp(0)$; that is,

$$g(\phi(z, \cdot))^\sharp(d) \leq g(\phi(z, \cdot))^\sharp(0) + \langle \xi(z), d \rangle \equiv Ag(z) + \langle \xi(z), d \rangle, \quad d \in \mathbf{X}.$$

Then (ii) and the fact that $g(\phi(z, d)) \leq g(\phi(z, \cdot))^\sharp(d)$ yield (iii). Finally, if $\partial g(\phi(z, \cdot))^\sharp(0) \neq \emptyset$ for all $z \in \mathbf{Z}$, it is evident that (iii) implies (i). $\square$

We mention that Lemma 3.7 can serve as a tool to verify that g is a fixed point: in examples, it is sometimes easier to verify a relation like (3.2) which does not involve the concave envelope (e.g. [5]).

Remark 3.8. In the context of mathematical finance, the $\mathbb{R}^n$-valued process X represents the discounted prices of n tradable securities, while $f(Z_T)$ is seen as an option maturing at time T. Inequality (3.2) with $g = u = A^\infty f$ expresses that the trading strategy $H_t := \xi(Z_{t-1})$ yields a superhedge for the seller of the option if $u(z_0)$ is charged as its price:

$$u(z_0) + \sum_{t=1}^{T} \langle H_t, X_t - X_{t-1} \rangle \geq u(Z_T) \geq f(Z_T), \tag{3.3}$$

where the left-hand side is the balance obtained from the amount $u(z_0)$ and the gains/losses from trading according to H. A similar observation applies if the time horizon T is seen as fixed (which is more natural in finance); namely, $H_t \in \partial[A^{T-t}(\phi(Z_{t-1}, \cdot))^\sharp](0)$ yields a process such that

$$A^T f(z_0) + \sum_{t=1}^{T} \langle H_t, X_t - X_{t-1} \rangle \geq f(Z_T). \tag{3.4}$$

In particular, this gives a simple and constructive proof for the result of [6] mentioned in the Introduction (note that an element of the supergradient can be chosen simply by taking a directional derivative).

By its definition, $A^T(z_0)$ is the minimal constant allowing for an inequality of the form (3.4) to hold almost-surely under all martingale laws and hence in all viable models, so that $A^T(z_0)$ is called the robust (or model-independent) superhedging price. To enlarge a bit further on the financial aspect, suppose that $\mathbf{Z} \subseteq \mathbf{X} \times \mathbf{Y}$ for some set $\mathbf{Y}$ and that $\phi(x, y, d) = \varphi(x+d, y)$ for some function $\varphi : \mathbf{X} \times \mathbf{Y} \to \mathbf{Z}$, where we now write (x, y) instead of z (see also Section 4.1 below). If $u = A^\infty f$, then $u(\cdot, y)$ is concave because $u(x, y)$ is the concave envelope of $u(\varphi(\cdot, y))$ evaluated at x, and moreover

$$\partial_x u(x, y) = \partial u(\phi(x, y, \cdot))^\sharp(0).$$

In other words, the hedging strategy is given by $\xi(x, y) = \partial_x u(x, y)$, which corresponds to the option's Delta in the language of finance.

Certain classical martingale inequalities hold also for submartingales. This can be related to the above as follows (the submartingale property is understood componentwise in the multivariate case).

Remark 3.9. Let $\mathbf{X} = \mathbb{R}^n$ and $g : \mathbf{X} \to \overline{\overline{\mathbb{R}}}$; then by Lemma 2.2, we have $\sup_{\mu \in \mathcal{M}(x)} \mu[g] = g^\sharp(x)$. Now let $\mathcal{M}^*(x)$ be the set of all probability measures on $\mathbf{X}$ having finite support and barycenter $x^* \geq x$. If the function g is (componentwise) nonincreasing, we also have

$$\sup_{\mu \in \mathcal{M}^*(x)} \mu[g] = g^\sharp(x), \quad x \in \mathbf{X}.$$

Indeed, for each $\mu^* \in \mathcal{M}^*(x)$ there is $\mu \in \mathcal{M}(x)$ such that $\mu^*[g] \leq \mu[g]$. As a consequence, the martingale inequality corresponding to f and ϕ extends to submartingales under the condition that

$$A^\infty f(\phi(z, \cdot)) \quad \text{is nonincreasing}.$$

Some martingale inequalities extend only to, e.g., nonnegative submartingales. Such a case can be covered by choosing a suitable state space $\mathbf{Z}$, as in Section 4.2 below.

We conclude this section with a brief remark about measurability questions (which we have avoided wherever possible).

Remark 3.10. Suppose that $\mathbf{X} = \mathbb{R}^n$ and $\mathbf{Z}$ is, say, a Polish space, and that ϕ is Borel-measurable. If f is Borel-measurable, one can check that Af and $A^\infty f$ are upper-semianalytic and in particular universally measurable; however, it can happen that Af is not Borel-measurable. As a consequence, the hedging strategy in Remark 3.8 can also be chosen to be universally measurable.

4   Examples

4.1   Doob's Maximal Inequality

The aim of this subsection is to illustrate the above abstract theory by a ramification of Doob's maximal L^p-inequality; in this case, all quantities of interest can be computed explicitly. In what follows, $\mathbf{X}$ is a vector space with norm $|\cdot|$.

Proposition 4.1. *Let* $1 < p < \infty$, $\mathbf{Z} = \{(x,y) \in \mathbf{X} \times \mathbb{R}_+ : |x| \leq y\}$ *and*

$$\phi(x,y,d) = (x+d, y \vee |x+d|), \quad f(x,y) = y^p - \left(\tfrac{p}{p-1}\right)^p |x|^p, \quad (x,y,d) \in \mathbf{Z} \times \mathbf{X}.$$

Then the minimal fixed point of A dominating f is given by

$$A^\infty f(x,y) = \begin{cases} f(x,y) & \text{if } |x| < \frac{p-1}{p} y, \\ \tilde{u}(x,y) & \text{if } |x| \geq \frac{p-1}{p} y, \end{cases} \tag{4.1}$$

where

$$\tilde{u}(x,y) := py^p - \tfrac{p^2}{p-1}|x|y^{p-1}, \quad (x,y) \in \mathbf{Z}.$$

Remark 4.2. The proof below also shows that the constant $\left(\frac{p}{p-1}\right)^p$ in the definition of f is optimal. Namely, if

$$f_c(x,y) = y^p - c|x|^p \tag{4.2}$$

for $c \geq 0$, we shall see that $A^\infty f_c \equiv \infty$ for $c < \left(\frac{p}{p-1}\right)^p$, whereas $A^\infty f_c$ is finite-valued for $c \geq \left(\frac{p}{p-1}\right)^p$.

Setting $|M|_T^* = \max_{0 \leq t \leq T} |M_t|$ and applying the results of the previous subsection, we immediately deduce the following ramification of Doob's maximal L^p-inequality.

Corollary 4.3. *For all $(x, y) \in \mathbf{Z}$, $T \geq 0$ and every (simple) $\mathbf{X}$-valued martingale M starting at $M_0 = x$, we have*

$$E\big[(|M|_T^*)^p \vee y^p - (\tfrac{p}{p-1})^p |M_T|^p\big] \leq \begin{cases} y^p - (\tfrac{p}{p-1})^p |x|^p & \text{if } |x| < \tfrac{p-1}{p} y, \\ py^p - \tfrac{p^2}{p-1}|x| y^{p-1} & \text{if } |x| \geq \tfrac{p-1}{p} y \end{cases}$$

and the right-hand side is optimal. In particular, for the case $y = |x|$, we have

$$E\big[(|M|_T^*)^p - (\tfrac{p}{p-1})^p |M_T|^p\big] \leq -\tfrac{p}{p-1}|x|^p \leq 0 \tag{4.3}$$

and thus $\||M|_T^*\|_p \leq \tfrac{p}{p-1}\|M_T\|_p$.

We mention that the function $\tilde{u}$ also appears in a proof of (4.3) in [11]. The optimality of the constant was not studied there; incidentally, we see that $\tilde{u}(x, y)$ actually yields the optimal constant for initial conditions with $y = |x|$. The function $\tilde{u}$ can also be extracted (with some additional work) from Cox [15], who considers the finite-horizon version of Doob's inequality in the case $\mathbf{X} = \mathbb{R}$.

Proof of Proposition 4.1 and Remark 4.2. Fix $c \geq 0$ and let $f := f_c$ be defined as in (4.2). By Remark 3.4, the function $u := A^\infty f$ has the representation

$$u(x, y) = \sup_{\mu \in \mathcal{M}^\infty(x)} \mu[f(Z_\infty)], \tag{4.4}$$

and in view of the form of f, this implies that $u(x, y)$ depends on x only through $|x|$. Moreover, we have the scaling property $u(\lambda x, \lambda y) = \lambda^p u(x, y)$ for $\lambda \geq 0$; cf. Remark 3.6. Thus, u is completely described by the function

$$\varrho : [0, 1] \to \overline{\mathbb{R}}, \quad \varrho(|x|) := u(x, 1);$$

namely, we have $u(0, 0) = 0$ and $u(x, y) = y^p \varrho(|x|/y)$ for all $(x, y) \in \mathbf{Z}$ with $y > 0$. On the other hand, we know that u is a fixed point of A,

$$u(x, y) = Au(x, y) = u(x + \cdot, y \vee |x + \cdot|)^\#(0) = u(\cdot, y \vee |\cdot|)^\#(x), \tag{4.5}$$

so that $x \mapsto u(x, y)$ is concave. In particular, using $u \geq f$ and the scaling property, we see that $u(x, y) = \infty$ at one point (x, y) if and only if $u \equiv \infty$ on $\mathbf{Z}$. For the time being, let us suppose that we are in the case where u is finite.

Under this condition, it follows from (4.5) and the scaling property, or also directly from (4.4), that $x \mapsto u(x, y)$ is continuous. Thus, ϱ is a continuous concave function on $[0, 1]$, and it follows from (4.5) that its (left)

tangent t at the boundary point $r = 1$ satisfies

$$t(r) \geq r^p \varrho(1), \quad r \in [1, \infty); \tag{4.6}$$

note that $r^p \varrho(1) = u(x_r, |x_r|)$ if $x_r \in \mathbf{X}$ is any point with $|x_r| = r$ (we may assume that $\mathbf{X} \neq \{0\}$). For later use, we remark that the converse is also true: a continuous concave function $\bar{\varrho}$ on $[0, 1]$ satisfying the analogue of (4.6) determines a fixed point $\bar{u}$ of A.

Let us establish that

$$\varrho(0) \geq 1 \quad \text{and} \quad \varrho(1) < 0. \tag{4.7}$$

Indeed, $\varrho(0) = u(0, 1) \geq f(0, 1) = 1$. Moreover, if $\varrho(1)$ were nonnegative, then $p > 1$ and (4.6) would imply that the tangent t has nonnegative slope, thus $\varrho(1) = t(1) \geq t(0) \geq \varrho(0) \geq 1$. But then (4.6) states that the affine function $t(r)$ dominates r^p on $[1, \infty)$, which is impossible.

As a result, $r \mapsto r^p \varrho(1)$ is concave and we see that the tangent condition (4.6) can be stated equivalently in differential terms. Namely, if $\varrho'(1)$ denotes the slope of t, (4.6) is equivalent to

$$0 > \varrho'(1) \geq p\varrho(1). \tag{4.8}$$

In view of (4.7), the tangent t has a unique zero r_1 in $[0, 1]$. Using the Intercept Theorem, (4.8) implies that

$$\frac{1 - r_1}{\varrho(1)} = \frac{1}{\varrho(1) - t(0)} = \frac{1}{\varrho'(1)} \leq \frac{1}{p\varrho(1)}$$

and hence

$$r_1 \leq 1 - 1/p. \tag{4.9}$$

Next, we construct another fixed point of A for comparison. Let $\bar{t}$ be the (uniquely determined) affine function which is parallel to t and touches

$$r \mapsto f(x_r, 1), \quad r \in [0, 1].$$

We denote by $(r_2, f(x_{r_2}, 1))$ the coordinates of this touching point. Set

$$\bar{\varrho}(r) := \begin{cases} f(x_r, 1) & \text{for } r \in [0, r_2], \\ \bar{t}(r) & \text{for } r \in (r_2, 1]. \end{cases} \tag{4.10}$$

15

By definition, $\bar{\varrho}$ is a continuous concave function satisfying (4.8). As remarked above, this implies that $\bar{\varrho}$ defines a fixed point $\bar{u}$ of A via $\bar{u}(0,0) := 0$ and $\bar{u}(x,y) := y^p \bar{\varrho}(|x|/y)$ for $(x,y) \in \mathbf{Z}$ with $y > 0$.

The fact that $f \leq u$ and the construction of $\bar{u}$ imply that $\bar{u} \leq u$. On the other hand, we have $\bar{u} \geq f$ and u is the minimal fixed point of A above f, so $u \leq \bar{u}$. As a result, $\bar{u} = u$, $\bar{\varrho} = \varrho$ and $\bar{t} = t$. In particular, this establishes that ϱ is of the specific form (4.10); it remains to determine the tangent t explicitly.

Consider $r_0 := (1/c)^{1/p}$, the zero of $r \mapsto 1 - cr^p = f(x_r, 1)$. By concavity, we must have $r_0 \leq r_1$; recall that r_1 is the zero of the tangent. In view of (4.9), we conclude that

$$r_0 \leq r_1 \leq 1 - 1/p; \tag{4.11}$$

hence, our assumption that u is finite is contradicted whenever $c < (\frac{p}{p-1})^p$.

Suppose that $c = (\frac{p}{p-1})^p$. Then $r_0 = 1 - 1/p$ and so (4.11) implies that $r_1 = r_0 = 1 - 1/p$. The slope of $r \mapsto f(x_r, 1)$ in this point is $-\frac{p^2}{p-1}$; therefore,

$$t(r) = -r\frac{p^2}{p-1} + p.$$

In view of (4.10), this corresponds to the claimed formula (4.1). Since we have seen that this form of u defines a fixed point dominating f, we are necessarily in the case where $A^\infty f$ is finite; moreover, as f is decreasing with respect to c, $A^\infty f$ is then also finite for all $c \geq (\frac{p}{p-1})^p$. $\qquad\square$

4.2 Differentially Subordinate Martingales

The main purpose of this subsection is to illustrate how one can accommodate a martingale inequality which holds only for a specific class of martingales. To this end, we shall treat an inequality for differentially subordinate martingales, first derived by Burkholder for real-valued processes in [9] and extended to the Hilbert-valued case in [10]. A martingale N is differentially subordinate to another martingale M if $|N_{t+1} - N_t| \leq |M_{t+1} - M_t|$ for all $t \geq 0$. In other words, this says that the increments of the bivariate martingale (M, N) take values in the cone $\{(d_1, d_2) : |d_2| \leq |d_1|\}$, and this is the condition defining the class of (bivariate) martingales for which the inequality will hold.

Let $\mathbf{H}$ be a Hilbert space. In what follows, our basic vector space is $\mathbf{X} := \mathbf{H} \times \mathbf{H}$ and our state space is $\mathbf{Z} = \mathbf{X} \cup \{\Delta\}$; the additional point Δ will be used as a cemetery state for paths that violate the subordination condition.

Proposition 4.4. *Let $1 < p < \infty$ and $p^* = \max\{p, p/(p-1)\}$. For $z \in \mathbf{Z}$ and $d = (d_1, d_2) \in \mathbf{X}$, define*

$$\phi(z, d) = \begin{cases} \Delta & \text{if } z = \Delta \text{ or } |d_2| > |d_1|, \\ z + d & \text{otherwise,} \end{cases}$$

$$f(z) = \begin{cases} -\infty & \text{if } z = \Delta, \\ |x_2|^p - (p^* - 1)^p |x_1|^p & \text{if } z = (x_1, x_2) \in \mathbf{X}, \end{cases}$$

$$\tilde{u}(z) = \begin{cases} -\infty & \text{if } z = \Delta, \\ p(1 - 1/p^*)^{p-1}(|x_2| - (p^* - 1)|x_1|)(|x_1| + |x_2|)^{p-1} & \text{if } z \in \mathbf{X}. \end{cases}$$

Then the minimal fixed point of A dominating f is given by $A^\infty f = u$, where u is defined for $1 < p \leq 2$ by

$$u(z) = \begin{cases} -\infty & \text{if } z = \Delta, \\ \tilde{u}(z) & \text{if } z = (x_1, x_2) \in \mathbf{X} \text{ and } |x_2| \leq (p^* - 1)|x_1|, \\ f(z) & \text{if } z = (x_1, x_2) \in \mathbf{X} \text{ and } |x_2| > (p^* - 1)|x_1| \end{cases}$$

and by the same identity with $\tilde{u}$ and f interchanged if $2 \leq p < \infty$.

Corollary 4.5. *Let $1 < p < \infty$ and $p^* = \max\{p, p/(p-1)\}$. Let M^1, M^2 be $\mathbf{H}$-valued (simple) martingales starting at $(M_0^1, M_0^2) = (x_1, x_2)$ and satisfying $|M_{t+1}^2 - M_t^2| \leq |M_{t+1}^1 - M_t^1|$ for all $t \geq 0$. Then*

$$E[|M_T^2|^p - (p^* - 1)^p |M_T^1|^p] \leq u(x_1, x_2)$$

and in particular $\|M_T^2\|_p \leq (p^* - 1)\|M_T^1\|_p$ if $x_1 = x_2$.

Proof of Proposition 4.4. All relevant properties are contained in [10]; we merely translate them into our setup. Indeed, we have $f(\Delta) = u(\Delta)$ by definition, and it is checked below Equation (1.10) in [10] that $f(z) \leq \tilde{u}(z)$ for $z \in \mathbf{X}$. Hence, $f \leq u$. Moreover, according to Remark 1.2 in [10], u is the smallest function which dominates f on $\mathbf{X}$ and has the property that $r \mapsto u(z + rd)$ is concave for all $z \in \mathbf{X}$ and all $d = (d_1, d_2) \in \mathbf{X}$ such that $|d_2| \leq |d_1|$. Using our notation and recalling that $u(\phi(z, \cdot)) = -\infty$ outside the set $\{|d_2| \leq |d_1|\}$, it follows that u is the smallest function dominating f on $\mathbf{Z}$ such that $u(\phi(z, \cdot))$ is concave on $\mathbf{X}$. The latter property implies that

$$Au(z) = u(\phi(z, \cdot))^\sharp(0) = u(\phi(z, 0)) = u(z),$$

so u is a fixed point of A. Conversely, if $g : \mathbf{Z} \to \overline{\mathbb{R}}$ is any fixed point of A, then $g(\phi(z, d)) = g(z + d + \cdot)^\sharp(0)$ and hence $g(\phi(z, \cdot))$ is concave. As a result, u is the smallest fixed point of A dominating u. $\square$

5 Tchakaloff's Theorem for Martingales

In the preceding sections, we have restricted our attention to simple martingales and we still have to argue that this entails no essential loss of generality. On the one hand, let us mention again that for nice functions f and ϕ, the extension from simple to general martingales can be done by direct approximation arguments; see, e.g., the proofs of Lemma 2.2 in [13] or Theorem 2.2 in [8]. On the other hand, we have developed the theory without regularity conditions and so we would like to see that the extension can be achieved under the natural requirement necessary to define the expectations; namely, the measurability alone. This will be achieved by a martingale version of Tchakaloff's theorem.

Following Bayer and Teichmann [3], a general version of Tchakaloff's classical theorem [24] about the existence of cubature formulas can be stated as follows: given an integrable function f on a probability space $(\Omega, \mathcal{F}, \mu)$, there exists a probability measure ν with finite support such that $\nu[f] = \mu[f]$, and moreover that support can be chosen to lie in the support of μ. The function f may be multivariate, which allows one to incorporate a finite number of linear constraints on ν; for instance, that ν should have the same first moment as μ. Our aim is to provide a version of the theorem where μ and ν are martingale laws. This extension is not immediate because the martingale property corresponds to an infinite number of constraints[3].

Theorem 5.1. *Let $k, n, T \in \mathbb{N}$ and $\mathbf{X} = \mathbb{R}^n$. Let $x_0 \in \mathbf{X}$ and let μ be the law of an $\mathbf{X}$-valued martingale $M_0, \ldots, M_T$ with $M_0 = x_0$, and let $A \subseteq \mathbf{X}^{T+1}$ be a (μ-measurable) set such that $\mu(A) = 1$. Moreover, let $f : \mathbf{X}^{T+1} \to \mathbb{R}^k$ be a μ-measurable function such that $\mu[|f|] < \infty$. There exists a martingale law ν, still starting at x_0, such that $\# \operatorname{supp} \nu \leq (n + k + 1)^T$, $\operatorname{supp} \nu \subseteq A$ and*

$$\nu[f] = \mu[f].$$

Proof. By changing f on a μ-nullset and replacing A with a smaller set of full μ-measure, we may assume that f and A are Borel. The case $T = 1$ is now a consequence of Tchakaloff's theorem in the form of [3, Corollary 2] applied to the function $\phi : \mathbf{X} \to \mathbb{R}^{n+k+1}$ given by $\phi(x) = (f(x), x, 1)$. Hence, we assume that the theorem holds for some $T \in \mathbb{N}$ and show how to pass to $T + 1$. So let μ be a martingale law on $\mathbf{X}^{T+1}$ and let $A \subseteq \mathbf{X}^{T+1}$ satisfy $\mu(A) = 1$. Let μ_0 be the marginal of μ on $\mathbf{X}^T$, given by $\mu_0(B) := \mu(B \times \mathbf{X})$ for $B \in \mathcal{B}(\mathbf{X}^T)$, and let μ_1 be a Borel-measurable stochastic kernel from $\mathbf{X}^T$

---

[3]We thank Josef Teichmann for the insightful discussions which led to this theorem.

to $(\mathbf{X}, \mathcal{B}(\mathbf{X}))$ such that

$$\mu = \mu_0 \otimes \mu_1. \tag{5.1}$$

It is easy to see that μ_0 is a martingale law on $\mathbf{X}^T$ and that $\mu_0(A_0) = 1$ if A_0 is the (universally measurable) canonical projection of A onto $\mathbf{X}^T$. On the other hand, it follows from (5.1) that there exists $N \in \mathcal{B}(\mathbf{X}^T)$ with $\mu_0(N) = 0$ such that for all $x \in \mathbf{X}^T \setminus N$, we have $\int |f(x, x')| \mu_1(x; dx') < \infty$ and

$$\mu_1(x) \text{ is a martingale law on } \mathbf{X} \text{ satisfying } \mu_1(x; A_x) = 1,$$

where $A_x \in \mathcal{B}(\mathbf{X})$ is the section $A_x = \{x' \in \mathbf{X} : (x, x') \in A\}$.

By the induction hypothesis, there exists a martingale law ν_0 on $\mathbf{X}^T$ such that

$$\# \operatorname{supp} \nu_0 \le (n+k+1)^T, \quad \operatorname{supp} \nu_0 \subseteq A_0 \setminus N \tag{5.2}$$

and

$$\nu_0[g] = \mu_0[g] \quad \text{for} \quad g(x) := \int f(x, x') \mu_1(x; dx').$$

Fix $x \in \mathbf{X}^T \setminus N$. By applying the case $T = 1$ to the function $f(x, \cdot)$ and the measure $\mu_1(x)$, we obtain a martingale law $\nu_1(x)$ on $\mathbf{X}$ such that $\# \operatorname{supp} \nu_1 \le n+k+1$, $\operatorname{supp} \nu_1 \subseteq A_x$ and

$$\int f(x, x') \nu_1(x; dx') = \int f(x, x') \mu_1(x; dx') \equiv g(x).$$

We may see $x \mapsto \nu_1(x)$ as a kernel and define

$$\nu = \nu_0 \otimes \nu_1;$$

this product is well defined as a consequence[4] of (5.2). By construction, we have $\# \operatorname{supp} \nu \le (n+k+1)^{T+1}$. Moreover, it follows from Fubini's theorem that

$$\mu[f] = \int \left[\int f(x, x') \mu_1(x; dx') \right] \mu_0(dx) = \mu_0[g] = \nu_0[g]$$
$$= \int \left[\int f(x, x') \nu_1(x; dx') \right] \nu_0(dx) = \nu[f],$$

and similarly that ν is a martingale law satisfying $\nu(A) = 1$. $\qquad \square$

---

[4]In particular, the finiteness of $\operatorname{supp} \nu_0$ implies that there are no measurability issues; ν is simply a finite weighted sum.

The preceding theorem entails that even for merely measurable functions f, simple martingales are sufficient to establish martingale inequalities; in particular, this yields an extension of the results from Section 3 to general martingales.

Corollary 5.2. *Let $\mathbf{X} = \mathbb{R}^n$ and let $f : \mathbf{X}^{T+1} \to \overline{\mathbb{R}}$ be universally measurable. Then*

$$\sup_{\mu \in \mathcal{M}^T(x_0)} \mu[f] = \sup_M E[f(M_0, \ldots, M_T)],$$

where the supremum on the right-hand side is taken over all n-dimensional martingales $M_0, \ldots, M_T$ with $M_0 = x_0$, each on its filtered probability space.

Proof. It suffices to show that $\sup_{\mu \in \mathcal{M}^T(x_0)} \mu[f(Z_T)] \geq E[f(M_0, \ldots M_T)]$ for any martingale M with $M_0 = x_0$. For this, we may assume without loss of generality that $E[f(M_0, \ldots M_T)] > -\infty$ and, by monotone convergence, that f is bounded from above. Hence, we may assume that f is real-valued.

Under these conditions, we have $\mu_M[|f|] < \infty$ for the law μ_M of M. Thus, Theorem 5.1 yields $\mu \in \mathcal{M}^T(x_0)$ such that $\mu[f] = \mu_M[f] = E[f(M_0, \ldots M_T)]$ and the claim follows. $\qquad\square$

References

[1] B. Acciaio, M. Beiglböck, F. Penkner, and W. Schachermayer. A model-free version of the fundamental theorem of asset pricing and the super-replication theorem. *To appear in Math. Finance*, 2013.

[2] B. Acciaio, M. Beiglböck, F. Penkner, W. Schachermayer, and J. Temme. A trajectorial interpretation of Doob's martingale inequalities. *Ann. Appl. Probab.*, 23(4):1494–1505, 2013.

[3] C. Bayer and J. Teichmann. The proof of Tchakaloff's theorem. *Proc. Amer. Math. Soc.*, 134(10):3035–3040 (electronic), 2006.

[4] M. Beiglböck, P. Henry-Labordère, and F. Penkner. Model-independent bounds for option prices: a mass transport approach. *Finance Stoch.*, 17(3):477–501, 2013.

[5] M. Beiglböck and P. Siorpaes. Pathwise versions of the Burkholder–Davis–Gundy inequality. *Preprint arXiv:1305.6188v1*, 2013.

[6] B. Bouchard and M. Nutz. Arbitrage and duality in nondominated discrete-time models. *To appear in Ann. Appl. Probab.*, 2013.

[7] H. Brown, D. Hobson, and L. C. G. Rogers. Robust hedging of barrier options. *Math. Finance*, 11(3):285–314, 2001.

[8] D. L. Burkholder. A geometrical characterization of Banach spaces in which martingale difference sequences are unconditional. *Ann. Probab.*, 9(6):997–1011, 1981.

[9] D. L. Burkholder. Boundary value problems and sharp inequalities for martingale transforms. *Ann. Probab.*, 12(3):647–702, 1984.

[10] D. L. Burkholder. Sharp inequalities for martingales and stochastic integrals. *Astérisque*, (157-158):75–94, 1988. Colloque Paul Lévy sur les Processus Stochastiques (Palaiseau, 1987).

[11] D. L. Burkholder. Explorations in martingale theory and its applications. In *École d'Été de Probabilités de Saint-Flour XIX—1989*, volume 1464 of *Lecture Notes in Math.*, pages 1–66. Springer, Berlin, 1991.

[12] D. L. Burkholder. Sharp norm comparison of martingale maximal functions and stochastic integrals. In *Proceedings of the Norbert Wiener Centenary Congress, 1994 (East Lansing, MI, 1994)*, volume 52 of *Proc. Sympos. Appl. Math.*, pages 343–358, Providence, RI, 1997. Amer. Math. Soc.

[13] D. L. Burkholder. The best constant in the Davis inequality for the expectation of the martingale square function. *Trans. Amer. Math. Soc.*, 354(1):91–105 (electronic), 2002.

[14] A. M. G. Cox and J. Obłój. Robust pricing and hedging of double no-touch options. *Finance Stoch.*, 15(3):573–605, 2011.

[15] D. C. Cox. Some sharp martingale inequalities related to Doob's inequality. In *Inequalities in statistics and probability (Lincoln, Neb., 1982)*, volume 5 of *IMS Lecture Notes Monogr. Ser.*, pages 78–83. Inst. Math. Statist., Hayward, CA, 1984.

[16] Y. Dolinsky and H. M. Soner. Martingale optimal transport and robust hedging in continuous time. *To appear in Probab. Theory Related Fields*, 2012.

[17] D. Hobson. Robust hedging of the lookback option. *Finance Stoch.*, 2(4):329–347, 1998.

[18] D. Hobson. The Skorokhod embedding problem and model-independent bounds for option prices. In *Paris-Princeton Lectures on Mathematical Finance 2010*, volume 2003 of *Lecture Notes in Math.*, pages 267–318. Springer, Berlin, 2011.

[19] J. H. B. Kemperman. The general moment problem, a geometric approach. *Ann. Math. Statist*, 39:93–122, 1968.

[20] J. Obłój. The Skorokhod embedding problem and its offspring. *Probab. Surv.*, 1:321–390, 2004.

[21] J. Obłój and M. Yor. On local martingale and its supremum: harmonic func-

tions and beyond. In *From stochastic calculus to mathematical finance*, pages 517–533. Springer, Berlin, 2006.

[22] A. Osękowski. *Sharp martingale and semimartingale inequalities*, volume 72 of *Instytut Matematyczny Polskiej Akademii Nauk. Monografie Matematyczne (New Series) [Mathematics Institute of the Polish Academy of Sciences. Mathematical Monographs (New Series)]*. Birkhäuser/Springer Basel AG, Basel, 2012.

[23] A. Osękowski. Survey article: Bellman function method and sharp inequalities for martingales. *Rocky Mountain J. Math.*, 43(6):1759–1823, 2013.

[24] V. Tchakaloff. Formules de cubatures mécaniques à coefficients non négatifs. *Bull. Sci. Math. (2)*, 81:123–134, 1957.