A weak-type inequality for the martingale square function

Adam Osękowski

Department of Mathematics, Informatics and Mechanics
University of Warsaw
Banacha 2, 02-097 Warsaw
Poland

Abstract

Let $f = (f_n)_{n\geq 0}$ be a real-valued martingale satisfying $f_0 \geq 0$ almost surely and let $S(f)$ denote the square function of f. The paper contains the proof of the weak-type bound

$$\lambda \mathbb{P}(S(f) \geq \lambda) \leq e\mathbb{E} \sup_{n\geq 0} f_n, \qquad \lambda > 0,$$

involving the one-sided maximal function on the right-hand side. The constant e is the best possible.

Keywords: Square function, martingale, weak-type inequality, best constant
2010 MSC: 60G42.

1. Introduction

Square function inequalities play a distinguished role in both classical and noncommutative probability theory, harmonic analysis, potential theory and many other areas of mathematics. The purpose of this note is to establish a sharp upper bound for the tail of the square function of a martingale, in terms of the first moment of the corresponding one-sided maximal function. This result is motivated by closely related classical works of Bollobás (1980), Burkholder (2002), Davis (1976), Novikov (1971), Pedersen & Peskir (2000), Shepp (1967), Wang (1991) and many others.

We need to introduce the necessary background and notation. Suppose that $(\Omega, \mathcal{F}, \mathbb{P})$ is a probability space, filtered by a nondecreasing family $(\mathcal{F}_n)_{n=0}^{\infty}$ of sub-σ-fields of $\mathcal{F}$. Let $f = (f_n)_{n\geq 0}$ be an adapted real-valued martingale and let $df = (df_n)_{n\geq 0}$ denote the associated difference sequence, given by

$$df_0 = f_0, \qquad df_n = f_n - f_{n-1}, \quad n = 1, 2, \ldots.$$

Then $S(f)$, the square function of f, is defined by

$$S(f) = \left(\sum_{n=0}^{\infty} |df_n|^2 \right)^{1/2}.$$

Email address: ados@mimuw.edu.pl (Adam Osękowski)
URL: www.mimuw.edu.pl/∼ados (Adam Osękowski)

We will also use the notation $S_n(f) = \left(\sum_{k=0}^n |df_k|^2\right)^{1/2}$ for the truncated square function. There is an interesting general question concerning the comparison of the sizes of f and $S(f)$, which is, for example, of fundamental importance to the theory of stochastic integration. The literature on this subject is extremely vast, the results are connected with many areas of mathematics and it is impossible to give even a brief review here. For some of the aspects of this subject, we refer the interested reader to the survey by Burkholder (1989), the book written by Osękowski (2012) or the monograph by Revuz & Yor (1999). To present our motivation, we start with the moment inequalities

$$c_p||S(f)||_p \le ||f||_p \le C_p||S(f)||_p, \qquad 1 \le p < \infty, \tag{1}$$

where $||f||_p = \sup_{n\ge 0} ||f_n||_p$ and c_p, C_p are absolute constants depending only on p. These estimates go back to the classical works of Khintchine (1923), Littlewood (1930), Marcinkiewicz (1937) and Paley (1932) (obviously, the concept of a martingale did not appear there; the results were stated in terms of partial sums of the Rademacher functions and the Haar system). In the 80's, Burkholder proved that if $1 < p < \infty$, then (1) holds with $c_p^{-1} = C_p = p^* - 1$, where $p^* = \max\{p, p/(p-1)\}$ (the survey Burkholder (1989) is a convenient reference). This choice of c_p is optimal for $1 < p \le 2$, and C_p is the best for $p \ge 2$. Furthermore, if $p = 1$, then the left inequality in (1) does not hold with any finite c_1, while the best choice for C_1 is 2 (cf. Osękowski (2005)). There is a substitute for the left inequality in the case $p = 1$. As shown by Cox (1982), we have the weak-type bound

$$\lambda \mathbb{P}(S(f) \ge \lambda) \le \sqrt{e}||f||_1, \qquad \lambda > 0$$

(see also the earlier work of Bollobás (1980) in this direction). Our objective is to establish the following related estimate involving the one-sided maximal function of f, given by $f^* = \sup_{n\ge 0} f_n$ (we will also use the truncated version $f_n^* = \sup_{0\le k\le n} f_k$).

Theorem 1.1. *For any martingale f satisfying $f_0 \ge 0$ almost surely, we have the bound*

$$\lambda \mathbb{P}(S(f) \ge \lambda) \le e\mathbb{E}f^* \tag{2}$$

and the constant e is the best possible.

Clearly, we cannot get rid of the assumption $\mathbb{P}(f_0 \ge 0) = 1$, even if we put $|\mathbb{E}f^*|$ on the right; for any $x < 0$ it is easy to construct a nonpositive martingale starting from x, satisfying $f^* = 0$ almost surely and $\mathbb{P}(S(f) \ge 1) > 0$.

The proof of the inequality (2) will rest on Burkholder's method: we will deduce the claim from the existence of a certain special function, possessing some majorization and concavity-type conditions. This is done in the next section. The final part of the paper is devoted to the optimality of the constant e: this is accomplished by providing the corresponding extremal examples.

2. Proof of (2)

As we have announced above, the proof of the weak-type inequality will rest on a certain special function. Consider $U : \{(x, y, z) \in \mathbb{R} \times [0,\infty)^2 : x \le z\} \to \mathbb{R}$ given by

$$U(x,y,z) = \begin{cases} 1 - ez & \text{if } (x-z)^2 + y^2 \ge 1, \\ 1 - \sqrt{1-y^2}\exp\left(-\dfrac{x-z}{\sqrt{1-y^2}}\right) - ex & \text{if } (x-z)^2 + y^2 < 1 \end{cases}$$

and extend it to the whole $\mathbb{R} \times [0, \infty)^2$ by the equality

$$U(x, y, z) = U(x, y, x \vee z). \tag{3}$$

We easily check that U is continuous. Its further properties are studied in lemmas below.

Lemma 2.1. *For any $x \in \mathbb{R}$ and any y, $z \geq 0$ we have the majorizations*

$$1_{\{y \geq 1\}} - e(x \vee z) \leq U(x, y, z) \leq 1 - e(x \vee z). \tag{4}$$

PROOF. Clearly, it suffices to show the claim under the additional requirement $x \leq z$ (replacing z by $x \vee z$ if necessary). We split the proof into two parts.

Left inequality. If $(x - z)^2 + y^2 \geq 1$, the majorization reduces to the trivial inequality $1 \geq 1_{\{y \geq 1\}}$. So, suppose that $(x - z)^2 + y^2 < 1$, or $y \in [0, \sqrt{1 - (x - z)^2}]$. Fix x and z, and take y from the interior of this interval. We easily check that

$$\frac{\partial}{\partial y} [U(x, y, z) - V(x, y, z)] = \frac{y}{\sqrt{1 - y^2}} \exp\left(-\frac{x - z}{\sqrt{1 - y^2}}\right)\left(1 + \frac{x - z}{\sqrt{1 - y^2}}\right) \geq 0,$$

so $U(x, y, z) - V(x, y, z) \geq U(x, 0, z) - V(x, 0, z)$. To handle the latter difference, keep z fixed and note that $x \in (z - 1, z]$. For x lying inside this interval, we have

$$\frac{\partial}{\partial x} [U(x, 0, z) - V(x, 0, z)] = e^{-x+z} - e \leq 0.$$

Therefore, $U(x, 0, z) - V(x, 0, z) \geq U(z, 0, z) - V(z, 0, z) = 0$. This completes the proof of the left estimate in (4).

Right inequality. The bound is clear when $(x - z)^2 + y^2 \geq 1$ (actually, both sides are equal). If, conversely, $(x - z)^2 + y^2 < 1$ (or $x > z - \sqrt{1 - y^2}$), then we have

$$U_x(x, y, z) = \exp\left(-\frac{x - z}{\sqrt{1 - y^2}}\right) - e \leq 0.$$

By the continuity of U, this implies $U(x, y, z) \leq U(z - \sqrt{1 - y^2}, y, z) = 1 - ez$. This finishes the proof of the lemma.

The key concavity-type property of U is studied in the next statement. We will need the auxiliary function $A : \mathbb{R} \times [0, \infty)^2 \to \mathbb{R}$, given by

$$A(x, y, z) = \begin{cases} 0 & \text{if } (x - z)^2 + y^2 \geq 1, \\ \exp\left(-\frac{x - z}{\sqrt{1 - y^2}}\right) - e & \text{if } (x - z)^2 + y^2 < 1 \end{cases}$$

(note that A coincides with U_x wherever the derivative exists). Clearly, A is bounded.

Lemma 2.2. *Suppose that $x \in \mathbb{R}$, $y \geq 0$ and $z \geq x$ are given numbers. Then for any $d \in \mathbb{R}$ we have*

$$U(x + d, \sqrt{y^2 + d^2}, z) \leq U(x, y, z) + A(x, y, z)d. \tag{5}$$

PROOF. If $(x-z)^2+y^2 \geq 1$, the inequality follows from the right inequality in (4). Indeed, using this bound and (3), we may write

$$U(x+d, \sqrt{y^2+d^2}, z) \leq 1 - e\big[(x+d) \vee z\big] \leq 1 - ez = U(x,y,z) + U_x(x,y,z)d.$$

Now, suppose that $(x-z)^2+y^2 < 1$. We consider two cases.

The case $d<0$. The function $d \mapsto U(x+d, \sqrt{y^2+d^2}, z)$ is continuous on $(-\infty, 0)$ and constant (equal to $1-ez$) on the halfline $\ell = \{d<0 : (x+d)^2+y^2+d^2 \geq 1\}$. Furthermore, we see that $A(x,y,z) \leq 0$. Putting all these facts together, we infer that it is enough to show the bound (5) on $(-\infty,0)\setminus \ell$, i.e., under the assumption $(x+d)^2+y^2+d^2 \leq 1$. Then the estimate can be rewritten in the equivalent form

$$\begin{aligned} &-\sqrt{1-y^2-d^2}\exp\left(-\frac{x+d-z}{\sqrt{1-y^2-d^2}}\right) \\ &\qquad \leq -\sqrt{1-y^2}\exp\left(-\frac{x-z}{\sqrt{1-y^2}}\right) + \exp\left(-\frac{x-z}{\sqrt{1-y^2}}\right)d. \end{aligned}$$

We may assume that $y=0$, dividing both sides by $\sqrt{1-y^2}$ and introducing the new variables $X = x/\sqrt{1-y^2}$, $Z = z/\sqrt{1-y^2}$ and $D = d/\sqrt{1-y^2}$ if this is not the case. Then, after some easy manipulations, the above inequality becomes

$$\sqrt{1-d^2}\exp\left(-\frac{x+d-z}{\sqrt{1-d^2}} + x - z\right) \geq 1-d. \tag{6}$$

The left-hand side, considered as a function of x, is nonincreasing (simply compute the derivative). So, it is enough to check the estimate (6) for $x=z$ which, after the substitution $s = -d/\sqrt{1-d^2} \geq 0$, becomes $F(s) := e^s - s - \sqrt{1+s^2} \geq 0$. However, we have $F(0)=0$ and

$$F'(s) = e^s - 1 - \frac{s}{\sqrt{1+s^2}} \geq e^s - 1 - s \geq 0,$$

so F is nonnegative and the inequality holds true.

The case $d>0$. If $x+d \leq z$, then, by the same argumentation as above, the estimate reduces to (6). Since the left-hand side is a nonincreasing function of x, we will be done if we prove this bound under the assumption $x+d=z$. Then (6) reads

$$\sqrt{1-d^2}e^{-d} \geq 1-d \tag{7}$$

which, after some simple manipulations, is equivalent to $F(d) := 1+d-(1-d)e^{2d} \geq 0$. However, we derive that $F(0) = F'(0) = 0$ and $F''(d) = 4de^{2d} \geq 0$, and hence the inequality is satisfied.

It remains to verify (5) in the case $x+d \geq z$. If $y^2+d^2 \leq 1$, the inequality becomes

$$-\sqrt{1-y^2-d^2} \leq -\sqrt{1-y^2}\exp\left(-\frac{x-z}{\sqrt{1-y^2}}\right) + \exp\left(-\frac{x-z}{\sqrt{1-y^2}}\right)d.$$

As previously, we may assume that $y = 0$; then the inequality can be transformed into $\sqrt{1-d^2}e^{x-z} \geq 1-d$. Now, the left-hand side is a nondecreasing function of x. Hence we will be done if we check it for smallest x, i.e., for $x = z-d$ (remember that we work under the assumption $x+d \geq z$). But then the bound reduces to (7), which has been already proved above. The final case corresponds to the inequalities $x+d \geq z$ and $y^2+d^2 > 1$. Then, after some straightforward manipulations, (5) is equivalent to $-\sqrt{1-y^2}+d \geq 0$, or $y^2+d^2 \geq 1$, which follows from the assumptions above.

This completes the proof of the concavity-type property of U.

Proof of (2). Pick an arbitrary martingale f satisfying $f_0 \geq 0$ almost surely and $\mathbb{E}f^* < \infty$. The key observation is that the sequence $(U(f_n, S_n(f), f_n^*))_{n\geq 0}$ forms a supermartingale. Indeed, for any $n \geq 0$ we have

$$\begin{aligned}\mathbb{E}U(f_{n+1}, S_{n+1}(f), f_{n+1}^*) &= \mathbb{E}U(f_{n+1}, S_{n+1}(f), f_n^*)\\ &= \mathbb{E}\Big[\mathbb{E}\big[U\big(f_n+df_n, \sqrt{S_n^2(f)+df_n^2}, f_n^*\big)|\mathcal{F}_n\big]\Big],\end{aligned}$$

where in the first passage we have exploited (3). Now, apply the inequality (5) with $x = f_n$, $y = S_n(f)$, $z = f_n^*$ and $d = df_n$ to get

$$U\big(f_n+df_n, \sqrt{S_n^2(f)+df_n^2}, f_n^*\big) \leq U(f_n, S_n(f), f_n^*) + A(f_n, S_n(f), f_n^*)df_n.$$

Both sides are integrable: combine the assumption $\mathbb{E}f^* < \infty$ with the right inequality in (4) and the boundedness of A. Hence, applying the conditional expectation with respect to $\mathcal{F}_n$ gives the supermartingale property. Thus, using the left inequality in (4), we get

$$\mathbb{P}(S_n(f) \geq 1) - e\mathbb{E}f_n^* = \mathbb{E}\Big[1_{\{S_n(f)\geq 1\}} - ef_n^*\Big] \leq \mathbb{E}U(f_n, S_n(f), f_n^*) \leq \mathbb{E}U(f_0, S_0(f), f_0^*).$$

But $f_0 = S_0(f) = f_0^*$ almost surely; hence, using (3) and applying (5) with $x = y = z = 0$ and $d = f_0$, we get

$$U(f_0, S_0(f), f_0^*) = U(f_0, S_0(f), 0) \leq U(0,0,0) = 0.$$

This gives the inequality $\mathbb{P}(S_n(f) \geq 1) \leq e\mathbb{E}f_n^*$, and the estimate (2) follows from a straightforward limiting argument.

3. Sharpness

Now we will construct a family of examples which will show that the constant e cannot be replaced in (2) by a smaller number. Fix a large positive integer N and consider i.i.d. random variables $\xi_1, \xi_2, \ldots, \xi_N$, with the distribution given by

$$\mathbb{P}(\xi_i = -N^{-1}) = 1 - \mathbb{P}(\xi_i = 1) = N/(N+1).$$

In addition, take a Rademacher variable ξ_{N+1}, independent of $\xi_1, \xi_2, \ldots, \xi_N$. Consider the stopping time $\tau = \inf\{n : |\xi_n| = 1\}$ and the sequence $f = (f_n)_{n=0}^{N+1}$ given by $f_0 \equiv 0$ and the equality

$$f_n = \xi_1 + \xi_2 + \ldots + \xi_{\tau\wedge n}, \qquad n = 1, 2, \ldots.$$

Since all ξ_j's are mean zero, the sequence f is a martingale, by virtue of Doob's optional sampling theorem. Furthermore, with probability 1 the last jump of f is of size 1, so we have $S(f) \geq 1$ almost surely. In addition, we see that the variable f^* takes values 1, $1 - N^{-1}$, $1 - 2N^{-1}$, ..., N^{-1} and 0 with probabilities $\frac{1}{N+1}$, $\frac{N}{(N+1)^2}$, $\frac{N^2}{(N+1)^3}$, ..., $\frac{N^{N-1}}{(N+1)^N}$ and $\frac{N^N}{(N+1)^N}$, respectively. Consequently,

$$\mathbb{E}f^* = \sum_{k=0}^{N-1} \left(1 - \frac{k}{N}\right) \frac{N^k}{(N+1)^{k+1}} = \left(\frac{N}{N+1}\right)^N,$$

which can be made arbitrarily close to e^{-1} by taking N sufficiently large. Thus the ratio $\mathbb{P}(S(f) \geq 1)/\mathbb{E}f^*$ can be made as close to e as we wish; this shows that the inequality (2) is indeed sharp.

Acknowledgments

The author would like to express his gratitude to an anonymous referee for the careful reading of the first version of the paper. The research was partially supported by the NCN grant DEC-2012/05/B/ST1/00412.

References

Bollobás, B., 1980. Martingale inequalities. Math. Proc. Cambridge Phil. Soc. 87, 377–382.

Burkholder, D. L., 1989. Explorations in martingale theory and its applications. École d'Ete de Probabilités de Saint-Flour XIX—1989, 1–66, Lecture Notes in Math., 1464, Springer, Berlin.

Burkholder, D. L., 2002. The best constant in the Davis' inequality for the expectation of the martingale square function. Trans. Amer. Soc. 354, 91–105.

Cox, D. C., 1982. The best constant in Burkholder's weak-L^1 inequality for the martingale square function. Proc. Amer. Math. Soc. 85, 427–433.

Davis, B., 1976. On the L^p norms of stochastic integrals and other martingales. Duke Math. J. 43, 697–704.

Khintchine, A., 1923. Über dyadische Brüche. Math. Z. 18, 109–116.

Littlewood, J. E., 1930. On bounded bilinear forms in an infinite number of variables. Quart. J. Math. Oxford 1, 164–174.

Marcinkiewicz, J., 1937. Quelques théoremes sur les séries orthogonales, Ann. Soc. Polon. Math. 16, 84–96.

Novikov, A. A. 1971. On moment inequalities for stochastic integrals. Theory Probab. Appl. 16, 538–541.

Osękowski, A., 2005. Two inequalities for the first moments of a martingale, its square function and its maximal function. Bull. Pol. Acad. Sci. Math. 53, 441–449.

Osękowski, A., 2012. Sharp martingale and semimartingale inequalities. Monografie Matematyczne 72, Birkhäuser.

Paley, R. E. A. C., 1932. A remarkable series of orthogonal functions I. Proc. London Math. Soc. 34, 241–264.

Pedersen, J. L., and Peskir, G., 2000. Solving non-linear optimal stopping problems by the method of time-change. Stochastic Anal. Appl., 18, 811–835.

Revuz D. and Yor, M., 1999. Continuous martingales and Brownian motion. Third edition, Fundamental Principles of Mathematical Sciences, 293. Springer-Verlag, Berlin.

Shepp, L. A., 1967. A first passage time for the Wiener process. Ann. Math. Statist. 38, 1912–1914.

Wang, G., 1991. Sharp Square-Function Inequalities for Conditionally Symmetric Martingales. Trans. Amer. Math. Soc., 328, 393–419.