SELF-NORMALIZED PROCESSES: EXPONENTIAL INEQUALITIES
MOMENTS AND LIMIT THEOREMS

by

Victor H. De La Pena
Michael J. Klass
Tze Leung Lai

Technical Report No. 2001-6
April 2001

Department of Statistics
STANFORD UNIVERSITY
Stanford, California
SELF-NORMALIZED PROCESSES: EXPONENTIAL INEQUALITIES,
MOMENTS AND LIMIT THEOREMS

by

Victor H. De La Pena
Columbia University

Michael J. Klass
University of California, Berkeley

Tze Leung Lai
Department of Statistics
Stanford University

Technical Report No. 2001-6
April 2001

This research was supported in part by National Science Foundation
grant DMS-00-72523, National Security Agency grant MDA
904-00-1-0018 and the Center for Advanced Study
in the Behavioral Sciences

Department of Statistics
Sequoia Hall
STANFORD UNIVERSITY
Stanford, California 94305-4065

http://www-stat.stanford.edu
SELF-NORMALIZED PROCESSES: EXPONENTIAL INEQUALITIES, 
MOMENTS AND LIMIT THEOREMS
BY VICTOR H. DE LA PEÑA¹, MICHAEL J. KLAS² AND TZE LEUNG LAI³

Columbia University, University of California at Berkeley, Stanford University

Abstract

In this paper we present several exponential and moment inequalities, as well as upper LILs, for self-normalized random variables including martingales. Tail probability bounds are also derived. Specifically for random variables $A_t, B_t > 0$, let $Y_t(\lambda) = \exp\{\lambda A_t - \frac{\lambda^2}{2} B_t^2\}$. We develop inequalities for the moments of $A_t/B_t$ and variants thereof, together with related LIL results, when $EY_t(\lambda) \leq 1$ or when $Y_t(\lambda)$ is a supermartingale, for all $\lambda$ in some subset of the real line. Our results are valid for a wide class of random processes including continuous martingales with $A_t = M_t$ and $B_t = \sqrt{\langle M \rangle_t}$, and sums of conditionally symmetric variables $d_i$ with $A_t = \sum_{i=1}^{t} d_i$ and $B_t = \sqrt{\sum_{i=1}^{t} d_i^2}$. We also prove a new type of almost sure and expectation LILs for discrete-time supermartingales $\Sigma_{i=1}^{t} d_i$ normalized by $\sqrt{\Sigma_{i=1}^{t} d_i^2}$ under a weak time-varying lower bound on $d_i$, which we show to be in some sense weakest possible. The key ingredient in this development is a new exponential supermartingale when the $d_i$’s are bounded from below.

Abbreviated Title: SELF-NORMALIZED PROCESSES

¹ Partially supported by NSF grant DMS-99-72237.
² Partially supported by NSF grant DMS-99-72417.
³ Research supported by NSF grant DMS-00-72523, NSA grant MDA 904-00-1-0018, and the Center for Advanced Study in the Behavioral Sciences.

AMS 1991 Subject Classifications: Primary 60E15, 60G42, 60G44; Secondary 60G40.

Key words and phrases: Martingales, self-normalized, exponential inequalities, laws of the iterated logarithm.
1. Introduction. In recent years, there has been increasing interest in limit theorems and moment bounds for self-normalized sums of i.i.d. zero-mean random variables $X_i$. In particular, Bentkus and Götze (1996) derive a Berry-Esseen bound for Student's t-statistic, and Giné, Götze and Mason (1997) prove that Student's t-statistic has a limiting standard normal distribution if and only if $X_1$ is in the domain of attraction of a normal law, by making use of exponential and $L_p$-bounds for the self-normalized sums $U_n = S_n/V_n$, where $S_n = \sum_{i=1}^n X_i$ and $V_n^2 = \sum_{i=1}^n X_i^2$. Shao (1997) provides large deviation results for $U_n$ without moment conditions and moderate deviation results when $X_1$ is in the domain of attraction of a normal or stable law. Egorov (1998) gives exponential inequalities for a centered variant of $U_n$. To see the connection between the t-statistic $T_n$ and the self-normalized sum $U_n$, observe that

$$T_n = \frac{S_n/V_n}{\sqrt{(n - (S_n/V_n)^2)/(n - 1)}}. \quad (1.1)$$

A recent paper of Caballero, Fernandez and Nualart (1998) contains moment inequalities for a continuous martingale over its quadratic variation and uses these results to show that if $\{M_t, t \geq 0\}$ is a continuous martingale null at zero, then for each $1 \leq p < q$, there exists a universal constant $C = C(p, q)$ such that

$$\|\frac{M_t}{\langle M \rangle_t^{1/2}}\|^p \leq C\|\frac{1}{\langle M \rangle_t^{1/2}}\|^q. \quad (1.2)$$

Related work in Revuz and Yor (1999, page 168) for continuous local martingales establishes for all $p > q > 0$ the existence of a constant $C_{pq}$ such that

$$\frac{E(\sup_{s \leq t} |M_s|)^p}{\langle M \rangle_t^{q/2}} \leq C_{pq} E(\sup_{s \leq t} |M_s|)^{p-q}. \quad (1.3)$$

It is important to point out that neither (1.2) nor (1.3) provide bounds in what is arguably the most important case of this type of inequalities, namely $p = q$. Bounds on $E\frac{|M_t|^p}{\langle M \rangle_t^{q/2}}$ are of particular interest because of their connection with the central limit theorem, as noted earlier in the case of self-normalized sums of i.i.d. random variables. For discrete-time martingales $\{\sum_{i=1}^n d_i, F_n, n \geq 1\}$, de la Peña (1999) provides exponential bounds for the tail probabilities of $\sum_{i=1}^n d_i/(\alpha + \beta V_n^2)$, where $V_n^2 = \sum_{i=1}^n E(d_i^2|F_{i-1})$ and $\beta > 0, \alpha \geq 0$. In view of the LIL (law of the iterated logarithm), it is of interest to consider $V_n$ (and $V_n\sqrt{2\log\log V_n}$) instead of $V_n^2$ in normalizing $\sum_{i=1}^n d_i$.

Motivated by these developments, we establish in this paper analogous exponential and $L_p$-bounds for a martingale divided by the square root of its quadratic variation or its conditional variance. We start more generally by considering random variables $A$ and $B$ with $B \geq 0$ such that

$$E \exp\{\lambda A - \frac{\lambda^2}{2} B^2\} \leq 1 \quad (1.4)$$
for all $\lambda \in \mathbb{R}$. Note that if we were allowed to maximize over $\lambda$ inside the expectation then the maximizing value $\lambda = \frac{A}{B^2}$ would give us $E\exp\left(\frac{A^2}{2B^2}\right) \leq 1$, which in turn would imply that $P(\Theta \geq x) \leq \exp(-x^2/2)$. Although we cannot interchange the order of $\max_\lambda$ and $E$, we can integrate over $\lambda$ with respect to a probability measure $F$ and interchange the order of integration with respect to $P$ and $F$. This approach gives not only tail probability bounds for $A/B$ but also bounds on the moments of $|A/B|$.

The remainder of the paper is organized as follows. We end this section with various lemmas identifying a large class of random variables for which (1.4) holds. In Section 2 we present a number of inequalities dealing with variations of the self-normalized theme, including $L_p$ and exponential bounds for $A/\sqrt{B^2 + (EB)^2}$. Section 3 considers the martingale case and develops maximal inequalities for self-normalized martingales. Section 4 provides iterated logarithm bounds for the moments of $A^+ / B$ and extends them to the case where (1.4) is replaced by $E\exp\{\lambda A - \Phi(\lambda B)\} \leq c$ for all $0 < \lambda < \lambda_0$, in which $\Phi$ is assumed to be any nonnegative, strictly convex function on $[0, \infty)$ such that $\Phi(0) = 0$, $\lim_{z \to \infty} \Phi(z) = \infty$ and $\limsup_{z \to \infty} \Phi''(z) < \infty$. Important examples are given in the context of discrete-time supermartingales, which are studied further in Section 5 where we prove an expectation form of the upper half of the LIL for discrete-time supermartingales whose increments satisfy certain one-sided boundedness assumptions and for continuous local martingales, thus providing an analogue of (1.3) in the case $q = p$. We also prove in Section 5 an almost sure upper LIL for discrete-time supermartingales $A_n = \sum_{i=1}^n d_i$ such that $d_i \geq -m_i$ a.s. for some non-negative $\mathcal{F}_{i-1}$-measurable random variable $m_i$. As in Griffin and Kuelbs (1991) and Shao (1997) who consider self-normalized LILs for sums of i.i.d. zero-mean variables, we self-normalize $A_n$ by $s_n = (\sum_{i=1}^n d_i^2)^{1/2}$. Whereas usual exponential upper bounds and LIL results require an upper bound on individual summands, for our self-normalized LIL a lower bound is required instead. Specifically, if $m_n = o(s_n^{-1} \log \log s_n^{-1} - \frac{1}{2})$ a.s. on $\{\lim_{n \to \infty} s_n = \infty\}$, then as a corollary of Theorem 5.3,

$$\limsup_{n \to \infty} A_n / \{2s_n (\log \log s_n)^{\frac{1}{2}}\} \leq 1 \text{ a.s. on } \{\lim_{n \to \infty} s_n = \infty\}.$$ 

The key tool used in developing this result is Lemma 5.1 which provides an exponential supermartingale when the $d_i$'s are bounded from below.

**Lemma 1.1.** Let $W_t$ be a standard Brownian motion. Assume that $T$ is a stopping time with $T < \infty$ a.s.. Then for all $\lambda \in \mathbb{R},$

$$E\exp\{\lambda W_T - \frac{\lambda^2 T}{2}\} \leq 1.$$  

3
LEMMA 1.2. Let \( M_t \) be a continuous, square-integrable martingale, with \( M_0 = 0 \). Then, for all \( \lambda \in \mathbb{R} \),

\[
E \exp\{\lambda M_t - \frac{\lambda^2 (M)_t}{2}\} \leq 1.
\]

If \( M_t \) is only assumed to be a continuous local martingale, then (1.6) is also valid (by application of Fatou’s lemma).

LEMMA 1.3. Let \( \{M_t : t \geq 0\} \) be a locally square-integrable martingale, with \( M_0 = 0 \). Let \( \{V_t\} \) be an increasing process, which is adapted, purely discontinuous and locally integrable; let \( V^{(p)} \) be its dual predictable projection. Set \( X_t = M_t + V_t \),

\[
C_t = \sum_{s \leq t} ((\Delta X_s)^+)^2, \quad D_t = \{\sum_{s \leq t} ((\Delta X_s)^-)^2\}_t^{(p)}, \quad H_t = (M^c)_t + C_t + D_t.
\]

Then \( \exp\{X_t - V_t^{(p)} - \frac{1}{2}H_t\} \) is a supermartingale and hence for all \( \lambda \in \mathbb{R} \),

\[
E \exp\{\lambda(X_t - V_t^{(p)}) - \frac{\lambda^2}{2} H_t\} \leq 1.
\]

Lemma 1.3 is taken from Proposition 4.2.1 of Barlow, Jacka and Yor (1986). A related bound can be found in Lemma 3.2 below which treats the case of discrete-time martingales with bounded increments. In this case, \( A \) in (1.4) is a martingale and \( B^2 \) is a multiple of its conditional variance; see also Kubilius and Mémin (1994) and Griffin and Kuelbs (1991). The following result holds without any integrability conditions on the variables involved.

LEMMA 1.4. Let \( \{d_i\} \) be a sequence of variables adapted to an increasing sequence of \( \sigma \)-fields \( \{\mathcal{F}_i\} \). Assume that the \( d_i \)'s are conditionally symmetric (that is \( \mathcal{L}(d_i|\mathcal{F}_{i-1}) = \mathcal{L}(-d_i|\mathcal{F}_{i-1}) \)). Then for all \( \lambda \in \mathbb{R} \),

\[
E \exp\{\lambda \sum_{i=1}^{n} d_i - \frac{\lambda^2}{2} \sum_{i=1}^{n} d_i^2\} \leq 1.
\]

Lemma 1.4 has a long history, including Wang (1989) and Hitczenko (1990). In Hitczenko (1990) it is proved for conditionally symmetric martingale difference sequences, and de la Peña (1999) points out that the same result still holds without the martingale difference assumption and hence without any integrability assumptions. Note that any sequence of real-valued random variables \( X_i \) can be “symmetrized” to produce an exponential supermartingale satisfying (1.8), by simply introducing random variables \( X'_i \) such that

\[
\mathcal{L}(X'_n|X_1, X'_1, \ldots, X_{n-1}, X'_n, X_{n-1}, X_n) = \mathcal{L}(X_n|X_1, \ldots, X_{n-1})
\]
and setting $d_n = X_n - X'_n$; see Section 6.1 of de la Peña and Giné (1999).

2. Some exponential moment bounds. In this section we present a simple method to derive exponential and $L_p$-bounds for $A/\sqrt{B + (EB)^2}$ under assumption (1.4).

THEOREM 2.1. Let $B \geq 0$ and $A$ be two random variables satisfying (1.4) for all $\lambda \in \mathbb{R}$. Then for all $y > 0$,

$$E \frac{y}{\sqrt{B^2 + y^2}} \exp\left\{ \frac{A^2}{2(B^2 + y^2)} \right\} \leq 1.$$  

(2.1)

Consequently, if $EB > 0$, then $E \exp(A^2/[4(B^2 + (EB)^2)]) \leq \sqrt{2}$ and

$$E \exp(xA/\sqrt{B^2 + (EB)^2}) \leq \sqrt{2} \exp(x^2) \quad \text{for all } x > 0.$$  

Moreover, for all $p > 0$,

$$E(|A|/\sqrt{B^2 + (EB)^2})^p \leq 2^{p+\frac{1}{2}} p \Gamma(a/2).$$  

(2.3)

PROOF. Multiplying both sides of (1.4) by $(2\pi)^{-1/2} y \exp(-\lambda^2 y^2/2)$ (with $y > 0$) and integrating over $\lambda$, we obtain by using Fubini’s theorem that

$$1 \geq \int_{-\infty}^{\infty} E \frac{y}{\sqrt{2\pi}} \exp \left( \lambda A - \frac{\lambda^2}{2} B^2 \right) \exp \left( -\frac{\lambda^2 y^2}{2} \right) d\lambda$$

$$= E \left[ \frac{y}{\sqrt{B^2 + y^2}} \exp\left\{ \frac{A^2}{2(B^2 + y^2)} \right\} \right] \times \int_{-\infty}^{\infty} \frac{\sqrt{B^2 + y^2}}{\sqrt{2\pi}} \exp\left\{ -\frac{B^2 + y^2}{2} \left( \lambda^2 - 2 \frac{A}{B^2 + y^2} \lambda + \frac{A^2}{(B^2 + y^2)^2} \right) \right\} d\lambda$$

$$= E \left[ \frac{y}{\sqrt{B^2 + y^2}} \exp\left( \frac{A^2}{2(B^2 + y^2)} \right) \right],$$

proving (2.1). By Schwarz’s inequality and (2.1),

$$E \exp\left\{ \frac{A^2}{4(B^2 + y^2)} \right\} \leq \left\{ \left( E \frac{y \exp\left( \frac{A^2}{2(B^2 + y^2)} \right)}{\sqrt{B^2 + y^2}} \right) \left( E \frac{\sqrt{B^2 + y^2}}{y^2} \right) \right\}^{1/2}$$

$$\leq \left( E \frac{\sqrt{B^2 + y^2}}{y^2} + 1 \right)^{1/2} \leq \left( E \left( \frac{B}{y} + 1 \right) \right)^{1/2} \leq \sqrt{2} \text{ for } y = EB.$$

To prove (2.2) and (2.3), we assume without loss of generality that $EB < \infty$. Using the inequality $|ab| \leq \frac{a^2 + b^2}{2}$, with $a = \sqrt{2cA}/\sqrt{B^2 + (EB)^2}$ and $b = x/\sqrt{2c}$, we get $xA/\sqrt{B^2 + (EB)^2} \leq$
\[
\frac{x^2 A^2}{B^2 + (EB)^2} + \frac{x^2}{4c}, \text{ which in the case } c = 1/4 \text{ yields } 
\]
\[
E \exp \left\{ \frac{x A}{\sqrt{B^2 + (EB)^2}} \right\} \leq E \exp \left\{ \frac{c A^2}{B^2 + (EB)^2} + \frac{x^2}{4c} \right\} \leq \sqrt{2} \exp(x^2),
\]
proving (2.2). Moreover, by Markov's inequality, \(P(|A|/\sqrt{B^2 + (EB)^2} \geq x) \leq \sqrt{2} \exp(-x^2/4)\) for all \(x > 0\). Combining this with the formula \(EU^p = \int_0^\infty px^{p-1}P(U > x)dx\) for any \(U \geq 0\), we obtain
\[
E(|A|/\sqrt{B^2 + (EB)^2})^p \leq \sqrt{2} \int_0^\infty px^{p-1} \exp(-x^2/4)dx = 2^{p+1/2} \Gamma(p/2). \tag{\ref*{eq:markov}}
\]

Another application of the basic inequality (2.1) is the following.

**COROLLARY 2.2.** Let \(B \geq 0\) and \(A\) be two random variables satisfying (1.4) for all \(\lambda \in \mathbb{R}\). Then for all \(x \geq \sqrt{2}, y > 0\) and \(p > 0\),

\[
P\left( \frac{|A|}{\sqrt{B^2 + y}} \left( 1 + \frac{1}{2} log \left( \frac{B^2}{y} + 1 \right) \right) \geq x \right) \leq \exp(-x^2/2), \tag{\ref*{eq:corollary2.2}}
\]

\[
E\left( \frac{|A|}{\sqrt{B^2 + y}} \left( 1 + \frac{1}{2} log \left( \frac{B^2}{y} + 1 \right) \right) \right)^p \leq 2^{p/2} + 2^{(p-2)/2} \Gamma(p/2). \tag{\ref*{eq:corollary2.5}}
\]

**PROOF.** Note that for \(x \geq \sqrt{2}\) and \(y > 0\),

\[
P\left\{ \frac{A^2}{2(B^2 + y)} \geq \frac{x^2}{2} \left( 1 + \frac{1}{2} log \left( \frac{B^2}{y} + 1 \right) \right) \right\} \leq P\left\{ \frac{A^2}{2(B^2 + y)} \geq \frac{x^2}{2} + \frac{1}{2} log \left( \frac{B^2}{y} + 1 \right) \right\}
\]
\[
\leq e^{-x^2/2} E\frac{\sqrt{y} \exp\left(\frac{1}{2} \frac{A^2}{B^2 + y}\right)}{\sqrt{B^2 + y}} \leq e^{-x^2/2},
\]
in which the last inequality follows from (2.1). The proof of (2.5) makes use of (2.4) and is similar to that of (2.3); the term \(2^{p/2}\) comes from \(p \int_0^{\sqrt{2}} x^{p-1}dx\). \(\square\)

3. Maximal inequalities for self-normalized martingales. We begin by considering a sequence of conditionally symmetric random variables \(\{d_i\}\). By Lemma 1.3, \(\exp\{\lambda \sum_{i=1}^n d_i - \lambda^2 \sum_{i=1}^n d_i^2/2\}\) is a supermartingale for all \(\lambda \in \mathbb{R}\), and hence the random variables

\[
M_n = \int_{-\infty}^{\infty} \exp\{\lambda \sum_{i=1}^n d_i - \lambda^2 \sum_{i=1}^n d_i^2/2\}dF(\lambda), n \geq 1,
\]

(3.1)
form a supermartingale, for any probability distribution $F$ on the real line. Therefore, if we take $F$ to be the normal distribution with mean 0 and variance $y^{-2}$ ($y > 0$), the same calculations as those in the proof of Theorem 2.1 show that the sequence

$$
\frac{y}{(y^2 + \sum_{i=1}^{n} d_i^2)^{1/2}} \exp \left\{ \frac{\left( \sum_{i=1}^{n} d_i \right)^2}{2 \left( y^2 + \sum_{i=1}^{n} d_i^2 \right)} \right\}, \quad n \geq 1, 
$$

forms a nonnegative supermartingale, yielding the following maximal inequality for self-normalized sums of conditionally symmetric random variables: For any $c > 1$ and $y > 0$,

$$
P\left\{ \frac{\left| \sum_{i=1}^{n} d_i \right|}{(y^2 + \sum_{i=1}^{n} d_i^2)^{1/2}} \geq \left[ \log \left( y^2 + \sum_{i=1}^{n} d_i^2 \right) + 2 \log \left( c/y \right) \right]^{1/2} \text{ for some } n \geq 1 \right\} \leq c^{-1}. \tag{3.3}
$$

When the $d_i$ are i.i.d. normal with mean 0 and variance 1, Robbins (1970) used the martingale $\{ \int_{-\infty}^{\infty} \exp(\lambda \sum_{i=1}^{n} d_i - \lambda^2 n/2) dF(\lambda), n \geq 1 \}$ to derive upper bounds for various boundary crossing probabilities for the martingale $\sum_{i=1}^{n} d_i$ by choosing different $F$'s. Using $\lambda \sum_{i=1}^{n} d_i - (\lambda^2/2) \sum_{i=1}^{n} d_i^2$ instead of $\lambda \sum_{i=1}^{n} d_i - \lambda^2 n/2$ in this “method of mixtures” leads to inequality (3.3), which considers the random (“self-”) normalization $(\sum_{i=1}^{n} d_i^2 + y^2)^{1/2}$ instead of the $\sqrt{n}$ normalization considered by Robbins (1970). Robbins and Siegmund (1970) also used this method of mixtures to analyze the probability that Brownian motion crosses a boundary of the order of magnitude $\sqrt{t}(\log t)^{1/2}$ as $t \to \infty$. Replacing $t$ by $\langle M \rangle_t$, we obtain the following generalization of their result to certain continuous local martingales.

COROLLARY 3.1. Let $M_t$ be a continuous local martingale with $M_0 = 0$, $\lim_{t \to \infty} (M)_t = \infty$ a.s., and such that $E\exp(\lambda \langle M \rangle_t) < \infty$ for all $\lambda > 0$ and $t > 0$. Then for any $c > 1$ and $y > 0$,

$$
P\{ |M_t| \geq (y^2 + \langle M \rangle_t)^{1/2} \left[ \log(y^2 + \langle M \rangle_t) + 2 \log(c/y) \right]^{1/2} \text{ for some } t \geq 0 \} = c^{-1}. \tag{3.4}
$$

PROOF. By Novikov’s criterion (cf. Revuz and Yor (1999), page 332), $\{ \exp(\lambda M_t - \lambda^2 < (M)_t/2, t \geq 0 \}$ is a martingale. Therefore $\int_{-\infty}^{\infty} \exp\{\lambda M_t - \lambda^2 < (M)_t/2 \} dF(\lambda)$ is a nonnegative continuous martingale, and is equal to

$$
\frac{y}{(y^2 + \langle M \rangle_t)^{1/2}} \exp \left\{ \frac{M_t^2}{2(y^2 + \langle M \rangle_t)} \right\} \tag{3.5}
$$

when $F$ is chosen to be the normal distribution with mean 0 and variance $y^{-2}$. Therefore by Doob’s inequality, the probability in (3.4) is $\leq c^{-1}$. Equality actually holds in (3.4) by Lemma 1 of Robbins and Siegmund (1970) if it can be shown that (3.5) converges to 0 a.s. as $t \to \infty$. Since $\langle M \rangle_t \to \infty$
a.s., $M_t/(M)_t \to 0$ a.s. by the martingale strong law. Therefore $\exp\{\lambda M_t - \lambda^2(M)_t/2\} \to 0$ a.s. as $t \to \infty$, for every $\lambda \neq 0$. Hence $\int_{-\infty}^{\infty} \exp\{\lambda M_t - \lambda^2(M)_t/2\}dF(\lambda) \to 0$ a.s. by applying the dominated convergence theorem, noting that by Doob's inequality,

$$P\left\{ \int_{|\lambda| \geq a} \exp(\lambda M_t - \lambda^2(M)_t/2)dF(\lambda) \geq c \text{ for some } t \geq 0 \right\} \leq c^{-1} \int_{|\lambda| \geq a} dF(\lambda).$$

Note that the maximal inequalities (3.3) and (3.4) are similar to (2.4) which follows from (2.1) by the Markov inequality. The main difference is that (2.4) (with $e^{y^2/2} = c$) considers the event \(\{A_t^2/(B^2 + y) \geq (\log c)[2 + \log((B_t^2 + y)/y)]\}\) for a fixed $t$, while (3.3) and (3.4) (with $y^2$ replaced by $y$) consider an event of the form \(\{A_t^2/(B_t^2 + y) \geq \log((B_t^2 + y)/y) + 2 \log c \text{ for some } t\}\). Both (2.1) and (3.3) (or (3.4)) are derived by the method of mixtures with a normal mixing distribution $F$. In the remainder of this section we derive other results of the type (3.3) or (3.4) by using alternative mixing distributions $F$. In the next two sections we give several extensions of (2.1), (2.3) and (2.5) by choosing mixing distributions similar to (3.7) below.

Let $F$ be any finite measure on $(0, \infty)$ with $F(0, \infty) > 0$ and define the function

$$\psi(u, v) = \int_{0}^{\infty} \exp\{\lambda u - \lambda^2 v/2\}dF(\lambda).$$

Given any $c > 0$ and $v > 0$, the equation $\psi(u, v) = c$ has a unique solution $u = \beta_F(v, c)$. The function $v \to \beta_F(v, c)$ is called a Robbins-Siegmund boundary in Lai (1976), in which such boundaries are shown to have the following properties:

(a) $\beta_F(v, c)$ is a concave function of $v$.

(b) $\lim_{v \to \infty} \beta_F(v, c)/v = y_F/2$, where $y_F = \sup\{y > 0 : F(0, y) = 0\}$ (sup $\emptyset = 0$).

(c) If $dF(\lambda) = f(\lambda)d\lambda$ for $0 < \lambda < \lambda_0$ and $\inf_{0 < \lambda < \lambda_0} f(\lambda) > 0$ while $\sup_{0 < \lambda < \lambda_0} f(\lambda) < \infty$, then $\beta_F(v, c) \sim (v \log v)^{1/2}$ as $v \to \infty$.

(d) If $dF(\lambda) = f(\lambda)d\lambda$ for $0 < \lambda < e^{-2}$, and $= 0$ elsewhere, where

$$f(\lambda) = 1/(\lambda(\log \lambda^{-1})(\log \log \lambda^{-1})^{1+\delta}),$$

for some $\delta > 0$, then

$$\beta_F(v, c) = \{2v[(1 + \delta)(\log_2 v - \log 2) + \frac{1}{2} \log_3 v + \frac{1}{2} \log(1 + \delta) + \log(c/\sqrt{\pi}) + o(1)] \}^{1/2}$$

as $v \to \infty$, where we write $\log_k v = \log(\log_{k-1} v)$ for $k \geq 2$, $\log_1 v = \log v$.

Suppose $\{\exp\{\lambda A_t - \lambda^2 B_t^2/2\}, t \geq 0\}$ is a supermartingale for every $\lambda$ belonging to the support of $F$. Then in view of (3.6), $\{\psi(A_t, B_t^2), t \geq 0\}$ is also a supermartingale and

$$P\{A_t \geq \beta_F(B_t^2, c) \text{ for some } t \geq 0\} \leq F(0, \infty)/c,$$
for every $c > 0$. In particular by choosing $\delta$ in (3.7) arbitrarily small and $c$ in (3.9) arbitrarily large, it follows from (3.8) and (3.9) that

\begin{equation}
(3.10) \quad \limsup_{t \to \infty} \frac{A_t}{B_t(2\log \log B_t)^{1/2}} \leq 1 \quad \text{a.s. on} \quad \{ \lim_{t \to \infty} B_t = \infty \}.
\end{equation}

In view of Lemmas 1.1 and 1.2, (3.10) provides the upper half of the law of the iterated logarithm for continuous martingales and sums of conditionally symmetric random variables. In addition, the following result of Stout (1973) on bounded supermartingale difference sequences can also be used in conjunction with Robbins-Siegmund boundaries.

**Lemma 3.2.** Let $\{d_n\}$ be a sequence of random variables adapted to an increasing sequence of $\sigma$-fields $\{F_n\}$ such that $E(d_n|F_{n-1}) \leq 0$ and $d_n \leq M$ a.s. for all $n$ and some nonrandom positive constant $M$. Let $0 < \lambda_0 \leq M^{-1}$, $A_n = \sum_{i=1}^{n} d_i$, $B_n^2 = (1 + \frac{1}{2}\lambda_0 M) \sum_{i=1}^{n} E(d_i^2|F_{i-1})$. Then $\{\exp(\lambda A_n - \frac{1}{2}\lambda^2 B_n^2), F_n, n \geq 0\}$ is a supermartingale for every $0 \leq \lambda \leq \lambda_0$.

A natural extension of (3.3) to vector-valued $d_i$ is to self-normalize $\sum_{i=1}^{n} d_i$ with the matrix $(\sum_{i=1}^{n} d_i d_i^t + V)^{1/2}$, where $d_i^t$ denotes the transpose of the $m \times 1$ vector $d_i$ and $V$ is an $m \times m$ positive definite matrix (which reduces to $y^2 > 0$ in the case $m = 1$). A sequence of random vectors $\{d_i\}$ is called *conditionally symmetric* if $\{\lambda' d_i\}$ is a sequence of conditionally symmetric random variables for every $\lambda \in \mathbb{R}^m$. By Lemma 1.4, if $\{d_i\}$ is a sequence of conditionally symmetric random vectors, then for any probability distribution $F$ on $\mathbb{R}^m$, the sequence

\begin{equation}
(3.11) \quad \int_{\mathbb{R}^m} \exp\{\lambda' \sum_{i=1}^{n} d_i - \frac{1}{2}\lambda' \sum_{i=1}^{n} d_i d_i^t \lambda\} dF(\lambda), \quad n \geq 1,
\end{equation}

forms a nonnegative supermartingale, noting that $(\lambda' d_i)^2 = \lambda' d_i d_i^t \lambda$. In particular, if we choose $F$ to be the multivariate normal distribution with mean 0 and covariance matrix $V^{-1}$, then (3.11) reduces to

\begin{equation}
(3.12) \quad |V|^{1/2}|V + \sum_{i=1}^{n} d_i d_i^t|^{-1/2} \exp\{\left(\sum_{i=1}^{n} d_i\right)'(V + \sum_{i=1}^{n} d_i d_i^t)^{-1}(\sum_{i=1}^{n} d_i)/2\},
\end{equation}

where $|\cdot|$ denotes the determinant of a square matrix. Hence we can generalize (3.3) to the following maximal inequality for self-normalized sums of conditionally symmetric random vectors: For any $c > 0$ and any positive definite $m \times m$ matrix $V$,

\begin{equation}
(3.13) \quad P\left\{ \frac{(\sum_{i=1}^{n} d_i)'(V + \sum_{i=1}^{n} d_i d_i^t)^{-1}(\sum_{i=1}^{n} d_i)}{[\log |V + \sum_{i=1}^{n} d_i d_i^t| + 2\log(c/\sqrt{|V|})]^{1/2}} \geq 1 \quad \text{for some} \quad n \geq 1 \right\} \leq c^{-1}.
\end{equation}
Similarly we can generalize (3.4) to the multivariate case in the following result, in which \( \lambda_{\min}(\cdot) \) denotes the minimum eigenvalue of a nonnegative definite matrix.

**COROLLARY 3.3.** Let \( M_t \) be a continuous local martingale taking values in \( \mathbb{R}^m \) such that \( M_0 = 0 \), \( \lim_{t \to \infty} \lambda_{\min}(\langle M \rangle_t) = \infty \) a.s., and such that \( \mathbb{E} \exp(\lambda' \langle M \rangle_t \lambda) < \infty \) for all \( \lambda \in \mathbb{R}^m \) and \( t > 0 \). Then for any \( c > 1 \) and any positive definite \( m \times m \) matrix \( V \),

\[
(3.14) \quad P \left\{ \frac{M_t' (V + \langle M \rangle_t)^{-1} M_t}{[\log |V + \langle M \rangle_t| + 2 \log (c/\sqrt{|V|})]^{1/2}} \geq 1 \text{ for some } t \geq 0 \right\} = c^{-1},
\]

**PROOF.** First note that

\[
\frac{M_t' (V + \langle M \rangle_t)^{-1} M_t}{[\log |V + \langle M \rangle_t| + 2 \log (c/\sqrt{|V|})]^{1/2}} = \int_{\mathbb{R}^m} \exp\{\lambda' M_t - \frac{1}{2} \lambda' \langle M \rangle_t \lambda\} dF(\lambda),
\]

where \( F \) is the \( m \)-variate normal distribution with mean 0 and covariance matrix \( V^{-1} \). Given any \( \lambda \in \mathbb{R}^m \) with \( \lambda \neq 0 \), \( \lambda' M_t \) is a univariate martingale and \( \langle \lambda' M \rangle_t = \lambda' \langle M \rangle_t \lambda \to \infty \) a.s. since \( \lambda_{\min}(\langle M \rangle_t) \to \infty \) a.s. Hence by the martingale strong law, \( \lambda' M_t / \lambda' \langle M \rangle_t \lambda \to 0 \) a.s. and therefore \( \exp\{\lambda' M_t - \lambda' \langle M \rangle_t \lambda/2\} \to 0 \) a.s. as \( t \to \infty \), for every \( \lambda \neq 0 \). The rest of the proof is similar to that of Corollary 3.1. \( \square \)

4. Iterated logarithm bounds for moments of self-normalized variables and their generalizations. In this section we prove bounds for \( E h(A^+ / B) \) in terms of \( E \{ H(B) \} \), where \( H \) is a function that depends on \( h \). The main results are Theorems 4.3 and 4.6. Applications of these results are given in Examples 4.4, 4.5 and 4.8, which relate in particular the \( p \)th absolute moment of \( A^+ / B \) to that of the iterated logarithm \( \sqrt{\log \log (B \lor B^{-1} \lor e^2)} \).

Let \( L : (0, \infty) \to (0, \infty) \) be a non-decreasing function such that

\[
(4.1) \quad L(cy) \leq 3cL(y) \text{ for all } c \geq 1 \text{ and } y > 0,
\]

\[
(4.2) \quad L(y^2) \leq 3L(y) \text{ for all } y \geq 1,
\]

\[
(4.3) \quad \int_1^\infty \frac{dx}{xL(x)} = \frac{1}{2}.
\]

An example satisfying (4.1)-(4.3) is the function

\[
(4.4) \quad L(y) = \beta \{ \log(y + \alpha) \} \{ \log \log(y + \alpha) \} \{ \log \log \log(y + \alpha) \}^{1+\delta},
\]

10
where $\delta > 0$, $\alpha$ is chosen sufficiently large to ensure (4.1)-(4.2) and $\beta$ is a normalizing constant so that (4.3) holds.

**LEMMA 4.1.** Let $\gamma \geq 1$. Then $yL(y/B \vee B/y) \leq 3\gamma\{L(\gamma) \vee L(B \vee B^{-1})\}$ for any $0 < y \leq \gamma$ and $B > 0$. Consequently, for any $A \geq B > 0$ and any $-\frac{A}{B} < x \leq 0$,

$$
(x + \frac{A}{B})L(\frac{x + \frac{A}{B}}{x + \frac{A}{B}} \vee \frac{B}{x + \frac{A}{B}}) \leq 3\frac{A}{B}\{L(\frac{A}{B}) \vee L(B \vee \frac{1}{B})\}.
$$

**PROOF.** First consider the case $y \leq 1$. From (4.1) and the fact that $L$ is non-decreasing, it follows that

$$
yL(\frac{y}{B} \vee \frac{B}{y}) \leq yL(\frac{1}{B} \vee B)) \leq 3L(B \vee \frac{1}{B}).
$$

For the remaining case $1 < y \leq \gamma$, since $L$ is non-decreasing, we have

$$
yL(\frac{y}{B} \vee \frac{B}{y}) \leq \gamma L(\frac{1}{B} \vee B)) 
\leq \gamma\{L(\gamma^2) \vee L\left((B \vee \frac{1}{B})^2\right)\} \leq 3\gamma\{L(\gamma) \vee L(B \vee \frac{1}{B})\},
$$

where the last inequality follows from (4.2). \(\Box\)

**LEMMA 4.2.** Let $B > 0$ and $A$ be random variables satisfying (1.4) for all $\lambda > 0$. Define

$$
g(x) = \frac{\exp\{x^2/2\}}{x}1(x \geq 1).
$$

Then

$$
E\frac{g(\frac{A}{B})}{L(\frac{A}{B}) \vee L(B \vee \frac{1}{B})} \leq \frac{3}{\int_0^\infty \exp(-x^2/2)dx}.
$$

**PROOF.** By a change of variables, $\int_0^1 (\lambda L(1/\lambda))^{-1}d\lambda = \int_1^\infty (\lambda L(\lambda))^{-1}d\lambda = \frac{1}{2}$. Let

$$
f(\lambda) = \frac{1}{\lambda L(\max(\lambda, \frac{1}{2}))}, \ \lambda > 0.
$$

Then $\int_0^\infty f(\lambda)d\lambda = \int_0^1 f(\lambda)d\lambda + \int_1^\infty f(\lambda)d\lambda = 1$, so $f$ is a density function on $(0, \infty)$. Therefore
integrating (1.4) with respect to this probability measure yields
\[ 1 \geq E \int_0^\infty \frac{\exp \{ A x - (B^2 x^2 / 2) \}}{x L(x \vee \frac{A}{B})} \, dx \]
\[ = E \int_0^\infty \frac{\exp \{ Ay - (y^2 / 2) \}}{y L(y \vee \frac{B}{y})} \, dy \text{ (letting } y = Bx) \]
\[ \geq E \{ \exp \left( \frac{A^2}{2B^2} \right) \} \int_{-\infty}^0 \frac{\exp \left( - \left( \frac{x^2}{2} \right) \right)}{x + \frac{A}{B} \text{L}(x + \frac{A}{B}) \vee \frac{B}{x + \frac{A}{B}}} \frac{1}{1} \, dx \text{ (letting } x = y - \frac{A}{B}) \]
\[ \geq E \{ \exp \left( \frac{A^2}{2B^2} \right) \} \int_0^1 \frac{\exp \left( - \left( \frac{x^2}{2} \right) \right)}{x + \frac{A}{B} \text{L}(x + \frac{A}{B}) \vee \frac{B}{x + \frac{A}{B}}} \frac{1}{1} \, dx \text{ (by (4.5))} \]
\[ = \left\{ \frac{1}{3} \int_0^1 e^{-z^2 / 2} \, dz \right\} E \frac{g(\frac{A}{B})}{L(\frac{A}{B}) \vee L(B \vee \frac{1}{B})}. \]

We next derive a bound on \( Eh(A^+/B) \) by making use of Lemma 4.2 for non-decreasing functions \( h \) that do not grow faster than \( g/L \).

**Theorem 4.3.** Let \( L : (0, \infty) \to (0, \infty) \) be a non-decreasing function satisfying (4.1)-(4.3). Define \( g \) by (4.7). Let \( h \) be a non-decreasing function on \([0, \infty)\) such that for some \( x_0 \geq 1 \) and \( c > 0 \),
\[ 0 < h(x) \leq cg(x)/L(x) \text{ for all } x \geq x_0. \]
(4.9)

Let \( q \) be a strictly increasing, continuous function on \([0, \infty)\) such that for some \( \bar{c} \geq c \),
\[ L(x) \leq q(x) \leq \frac{\bar{c}g(x)}{h(x)} \text{ for all } x \geq x_0. \]
(4.10)

Let \( B > 0 \) and \( A \) be random variables satisfying (1.4) for all \( \lambda > 0 \). Then
\[ Eh(A^+/B) \leq 4\bar{c} + h(x_0) + \text{Eh}(q^{-1}(L(B \vee B^{-1}))). \]
(4.11)

Consequently, \( Eh(A^+/B) < \infty \) if \( Eh(q^{-1}(L(B \vee B^{-1}))) < \infty \).

**Proof.** By Lemma 4.2,
\[ E \frac{g(A^+)}{L(\frac{A}{B}) \vee L(B \vee \frac{1}{B})} \leq 4. \]

Let \( Q = \{ L(B \vee \frac{1}{B}) \leq q(\frac{A}{B}) \}. \) Then, \( Eh(A^+/B) \) is majorized by
\[ h(x_0) + E \frac{h(A^+)}{g(\frac{A}{B})/L(\frac{A}{B}) \vee L(B \vee \frac{1}{B})} \cdot \frac{g(\frac{A}{B})}{L(\frac{A}{B}) \vee L(B \vee \frac{1}{B})} + \text{Eh}(A^+ \cdot 1(Q)1(\frac{A}{B} \geq x_0)) \]
\[ \leq h(x_0) + \text{sup}_{y \geq x_0} \frac{h(y)(L(y) \vee q(y))}{g(y)} \frac{g(\frac{A}{B})}{L(\frac{A}{B}) \vee L(B \vee \frac{1}{B})} + \text{Eh}(q^{-1}(L(B \vee \frac{1}{B}))) \]
\[ \leq h(x_0) + 4 \text{sup}_{y \geq x_0} \frac{h(y)g(y)}{g(y)} + \text{Eh}(q^{-1}(L(B \vee \frac{1}{B}))). \]
\[ \square \]
To apply Theorem 4.3, one can take \( L \) as given by (4.4) and choose \( q^{-1} \) that grows as slowly as possible (or equivalently, \( q \) that grows as rapidly as possible) subject to the constraint (4.10).

**EXAMPLE 4.4.** Define \( L \) by (4.4) and let \( h(x) = x^p \) for \( x \geq 0 \), with \( p > 0 \). Then (4.9) clearly holds with \( c = 1 \) and \( x_0 \) sufficiently large, for which (4.10) also holds with \( q(x) = g(x)/h(x) = \frac{x^{2/2}}{x^{p+1}} \). In this case,

\[
q^{-1}(y) = \{2 \log y + (p + 1 + o(1)) \log \log y\}^{1/2} \text{ as } y \to \infty .
\]

Since \( L(x) \sim \beta(\log x)(\log \log x) \) as \( x \to \infty \), Theorem 4.3 yields

\[
E(A^+ / B)^p < \infty \text{ if } E[\log[(\log(B \lor B^{-1}) \lor e)]^{p/2} < \infty,
\]

for random variables \( B > 0 \) and \( A \) satisfying (1.4) for all \( \lambda > 0 \).

**EXAMPLE 4.5.** Let \( 0 < \theta < 1 \) and \( h(x) = \exp(\theta x^2/2) \) for \( x \geq 0 \). Define \( L \) by (4.4). Then (4.9) holds with \( c = 1 \) and \( x_0 \) sufficiently large, for which (4.10) also holds with \( q(x) = g(x)/h(x) = \frac{x^{-1} \exp[(1 - \theta) x^2/2]}{x^{p+1}} \). In this case, \( h(q^{-1}(y)) = O((y(\log y)^{1/2})^{\theta/(1-\theta)}) \). Therefore if \( B > 0 \) and (1.4) holds for all \( \lambda > 0 \), then by Theorem 4.3,

\[
E \exp \left( \frac{\theta}{2} \left( \frac{A^+}{B} \right)^2 \right) < \infty \text{ if } E[\log(B) \log \log B]^{3/2} (\log \log \log B)^{1+\delta} \theta^{1/(1-\theta)} < \infty
\]

for some \( \delta > 0 \), where \( B = B \lor B^{-1} \lor e^3 \).

The following theorem modifies Theorem 4.3 by requiring (1.4) to hold only for the restricted range \( 0 < \lambda < \lambda_0 \). It also generalizes (1.4) by replacing the quadratic function \( \lambda^2 B^2/2 \) and the upper bound 1 in (1.4) by a convex function \( \Phi(\lambda B) \) and a finite positive constant \( c \).

**THEOREM 4.6.** Suppose that \( \Phi(\cdot) \) is a continuous function with \( \Phi'(x) \) strictly increasing, continuous and positive for \( x > 0 \), with \( \lim_{x \to \infty} \Phi(x) = \infty \) and \( \sup_{x > 0} \Phi''(x) < \infty \). Suppose \( B > 0 \) and \( A \) are random variables such that there exists \( c > 0 \) for which

\[
E \exp\{ \lambda A - \Phi(\lambda B) \} \leq c \text{ for all } 0 < \lambda < \lambda_0.
\]

For \( w > \Phi'(1) \), define \( y_w \) by the equation \( \Phi'(y_w) = w \), and let

\[
g(\Phi)(w) = y_w^{-1} \exp\{wy_w - \Phi(y_w)\}.
\]

Let \( \eta > \eta > 0 \). Let \( h: [0, \infty) \to (0, \infty) \) be a non-decreasing function. For \( b \geq \eta \), let \( g_b \) be a strictly increasing, continuous function on \( (0, \infty) \) such that for some \( \bar{c} > 0 \) and \( w_0 > \Phi'(2) \),

\[
g_b(w) \leq \bar{c} \{g(\Phi)(w)1(y_w \leq \lambda_0 b) + e^{\lambda_0 \eta w}1(y_w > \lambda_0 b)\}/h(w) \text{ for all } w \geq w_0.
\]
Let $L : (0, \infty) \rightarrow (0, \infty)$ be a non-decreasing function satisfying (4.1)-(4.3). Then there exists a constant $C$ depending only on $\lambda_0, \eta, \tilde{\eta}, c, \bar{c}$ and $\Phi$ such that

\begin{equation}
Eh(A^+/B) \leq C + h(w_0) + Eh(q^{-1}_{B \vee \eta}(L(B \vee \eta))).
\end{equation}

PROOF. Since $\Phi(x)$ is increasing in $x > 0$, (4.14) also holds with $B$ replaced by $B \vee \eta$. We shall therefore assume without loss of generality that $B \geq \eta$. Integrating (4.14) with respect to the probability measure defined by the density function (4.8) yields

\begin{equation}
c \geq E \int_0^{\lambda_0} \frac{\exp\{\lambda A - \Phi(\lambda B)\}}{\lambda L(\lambda \vee \lambda^{-1})} d\lambda = E \int_0^{\lambda_0} \frac{\exp\{x A/B - \Phi(x)\}}{x L(B/B \vee \eta)} dx.
\end{equation}

Let $\lambda_1 = \lambda_0 \vee \lambda_0^{-1} \vee \tilde{\eta}$. Since $\lambda_1 \geq \eta > \tilde{\eta}$, it follows from (4.18) and (4.1) that

\begin{equation}
c \geq E \int_{\lambda_0}^{\lambda_0 \eta} \frac{\exp\{\lambda_0 \tilde{\eta} A/B - \Phi(\lambda_0 \eta)\} dx}{L(\lambda_0 \vee B/\lambda_0 \eta)} \geq \frac{e^{-\Phi(\lambda_0 \eta)}}{3 \lambda_1/\tilde{\eta}} \log\left(\frac{\eta}{\tilde{\eta}}\right) E \frac{e^{\lambda_0 \eta A/B}}{L(B)}.
\end{equation}

Since $w_0 > \Phi'(2), y_{w_0} > 2$. Define

\begin{equation}
a_* = \sup\{a \leq 1 : a^2 \Phi''(x) \leq 1 \text{ for all } x > y_{w_0} - a\}.
\end{equation}

Note that $a_*>0$ and $y_{w_0} - a_* > 1$. Since $\Phi'(y_w) = w = 0$, a two-term Taylor expansion yields for $w \geq w_0$ and $x \in (y_w - a_*, y_w)$

\[wx - \Phi(x) = wy_w - \Phi(y_w) - \frac{(x - y_w)^2}{2} \Phi''(\xi^*) \geq wy_w - \Phi(y_w) - \frac{(x - y_w)^2}{2a_*^2},\]

in which $\xi^*$ lies between $x$ and $y_w$. The last inequality follows from (4.20), noting that $\xi^* > x > y_w - a_* \geq y_{w_0} - a_*$. It then follows from (4.18) that

\[c \geq E \left[1(y_{A/B} \leq \lambda_0 B, \frac{A}{B} \geq w_0) \int_{y_{A/B} - a_*}^{y_{A/B}} \frac{\exp\left\{\frac{A}{B} y_{A/B} - \Phi(y_{A/B}) - \frac{(x - y_{A/B})^2}{2a_*^2}\right\}}{x L(x/B \vee \eta)} dx \right] \]

\[\geq E \left[1(y_{A/B} \leq \lambda_0 B, \frac{A}{B} \geq w_0) \frac{\exp\{\frac{A}{B} y_{A/B} - \Phi(y_{A/B})\}}{3y_{A/B} L(y_{A/B}) \vee L(B \vee \frac{1}{B})} \int_{y_{A/B} - a_*}^{y_{A/B}} \exp\left\{-\frac{(x - y_{A/B})^2}{2a_*^2}\right\} dx \right].\]

The last inequality follows from Lemma 4.1. Hence
\[ (4.21) \]
\[
c \geq E \left[ 1(y_{A/B} \leq \lambda_0 B, \frac{A}{B} \geq w_0) \cdot \exp\left( \frac{A}{B} y_{A/B} - \Phi(y_{A/B}) \right) \frac{3y_{A/B}}{L(\lambda_0 B) \vee L(B \vee \frac{1}{B})} a^* \int_0^1 \exp(-\frac{z^2}{2}) \, dz \right] \]
\[
\geq E \left[ 1(y_{A/B} \leq \lambda_0 B, \frac{A}{B} \geq w_0) a^* \cdot \frac{\exp\left( \frac{A}{B} y_{A/B} - \Phi(y_{A/B}) \right)}{4y_{A/B}} \right] \]
\[
\geq \frac{a^*}{4} E \frac{g_{\Phi}(A/B)1(y_{A/B} \leq \lambda_0 B, \frac{A}{B} \geq w_0)}{L(\lambda_0 B) \vee L(B \vee \frac{1}{B})} \geq \frac{a^*}{12(\lambda_0 \vee 1 \vee \eta^{-2})} E \frac{g_{\Phi}(A/B)}{L(B)} 1(y_{A/B} \leq \lambda_0 B, \frac{A}{B} \geq w_0). \]

The last inequality in (4.21) follows from (4.1), noting that \( L(B \vee B^{-1}) \leq L(B/(\eta^2 \wedge 1)) \) since \( B \geq \eta \). Let \( Q = \{ L(B) \leq q_B(A/B) \} \). Then rewriting (4.16) as an upper bound for \( h \) and using the definition of \( Q \), we can majorize \( Eh(A^+/B) \) by

\[
h(w_0) + \tilde{c} E \left[ 1(Q) \left\{ \frac{g_{\Phi}(A/B)}{L(B)} 1\left( \frac{A}{B} \geq w_0, y_{A/B} \leq \lambda_0 B \right) \right. \right.
\[
+ \left. \left. \frac{e^{\lambda_0 w_{A/B}^B}}{L(B)} 1\left( \frac{A}{B} \geq w_0, y_{A/B} > \lambda_0 B \right) \right\} \right] + Eh(\frac{A}{B}) 1\left( Q^c \cap \left\{ \frac{A}{B} \geq w_0 \right\} \right)
\]
\[
\leq h(w_0) + C + Eh(q_{B \vee \eta}(L(B))),
\]
in which the inequality follows from (4.21) and (4.19). \( \square \)

**REMARK 4.7.** In the case \( \lambda_0 = \infty \) (as in Theorem 4.3 for which \( \Phi(x) = x^2/2 \)), the bounds (4.18) and (4.19) are not needed and the result for general \( \Phi \) is similar to (4.11) in Theorem 4.3. The main difference between (4.11) and (4.17) lies in \( q^{-1} \) in (4.11) versus the more elaborate \( q^{-1}_{B \vee \eta} \) in (4.17) to incorporate both (4.19) and (4.21).

**EXAMPLE 4.8.** Lemma 3.2 gives an example of \( (A, B) \) satisfying (1.4) only for \( 0 \leq \lambda \leq \lambda_0 \). Thus (4.14) holds with \( \Phi(x) = x^2/2 \) and \( g_{\Phi} \) reduces to the function \( g \) defined by (4.8) in this case, noting that \( y_w = w \). Define \( L \) by (4.4). First let \( h(x) = x^p \) for \( x \geq 0 \), with \( p > 0 \). For \( b \geq \eta > \tilde{\eta} > 0 \), let \( q_b \) be a strictly increasing function on \( (0, \infty) \) such that for all large \( b \),

\[
q_b(w) = e^{w^2/2}/w^{p+1} \quad \text{if} \quad w \leq (\lambda_0 \eta b)^{1/2},
\]
\[
\leq e^{w^2/2}/w^{p+1} \quad \text{if} \quad (\lambda_0 \eta b)^{1/2} < w \leq \lambda_0 b,
\]
\[
= e^{\lambda_0 w}/w^p \quad \text{if} \quad w > \lambda_0 b.
\]

Then (4.16) holds with \( \tilde{c} = 1 \). From (4.4) and (4.22), it follows that \( q_b^{-1}(L(b)) \sim (2 \log \log b)^{1/2} \) as \( b \to \infty \). Therefore (4.12) still holds with \( B \) replaced by \( B \vee \eta \) even though (1.4) holds only for \( 0 \leq \lambda \leq \lambda_0 \). Similarly, letting \( h(x) = e^{\zeta x} \) with \( 0 < \zeta < \lambda_0 \eta \), it follows from Theorem 4.6 that

\[
(4.23) \quad E \exp(\zeta A^+/B \vee \eta)) < \infty \quad \text{if} \quad E \exp\{\zeta(2(\log \log \tilde{B})(\log \log \log \tilde{B})^{1/4})^{1/2}\}
\]
for some $\delta > 0$, where $B = B \lor e^3$.

Another example to which Theorem 4.6 can be applied will be given by Lemma 5.1 (ii) in the next section, dealing with the more general case $\Phi(x) = x^r/r$ ($1 < r \leq 2$), for which

\begin{equation}
y_w = w^{1/(r-1)}, g_\Phi(w) = w^{-1/(r-1)} \exp\{(1 - r^{-1})w^r/(r-1)\}.
\end{equation}

In view of (4.24), it follows from Theorem 4.6 by arguments similar to Example 4.7 that under (4.14) with $\Phi(x) = x^r/r$, we have for any $p > 0$,

\begin{equation}
E(A^+/\{B \lor \eta\})^p < \infty \text{ if } E\{\log^+(\log(B \lor \eta))\}^{p(r-1)/r} < \infty.
\end{equation}

Moreover, (4.23) still holds if we replace 2 and 1/2 there by $r/(r-1)$ and its reciprocal, respectively.

5. A self-normalized LIL and its expectation extensions. While Lemma 3.2 is related to self-normalizing a martingale $A_n = \sum_{i=1}^n d_i$ by the square root of its conditional variance $\sum_{i=1}^n E(d_i^2 \mid F_{i-1})$, in this section we consider self-normalizing $A_n$ by its square function $\sum_{i=1}^n d_i^2$. We begin with the following lemma, which provides an analogue of Lemma 3.2. This new result gives an exponential supermartingale when the summands in $A_n$ are bounded from below rather than from above as in traditional versions of exponential supermartingales and their corresponding exponential bounds.

**Lemma 5.1.** Let $0 < \gamma < 1 < r \leq 2$. Define $c_{\gamma,r} = \max\{c_r, c_r^{(\gamma)}\}$, where

$c_r = \inf\{c > 0 : \exp(x - cx^r) \leq 1 + x \text{ for all } x \geq 0\}$,

$c_r^{(\gamma)} = \inf\{c > 0 : \exp(x - c|x|^r) \leq 1 + x \text{ for all } -\gamma \leq x \leq 0\}$.

(i) For all $x \geq -\gamma$, $\exp\{x - c_{\gamma,r} |x|^r\} \leq 1 + x$. Moreover, $c_r \leq (r - 1)^{r-1}(2 - r)^{2-r}/r$ and

$c_r^{(\gamma)} = -(\gamma + \log(1 - \gamma))/\gamma^r = \sum_{j=2}^\infty \gamma^{j-r}/j$.

(ii) Let $\{d_n\}$ be a sequence of random variables adapted to an increasing sequence of $\sigma$-fields $\{\mathcal{F}_n\}$ such that $E(d_n | \mathcal{F}_{n-1}) \leq 0$ and $d_n \geq -M$ a.s. for all $n$ and some nonrandom positive constant $M$. Let $A_n = \sum_{i=1}^n d_i$, $B_n = \sum_{i=1}^n |d_i|^r$. Then $\{\exp(\lambda A_n - (\lambda B_n)^r / r), \mathcal{F}_n, n \geq 0\}$ is a supermartingale for every $0 \leq \lambda \leq \gamma M^{-1}$.

**Proof.** The first assertion of (i) follows from the definition of $c_{\gamma,r}$. For $c > 0$, define $g_c(x) = \log(1 + x) - x + c|x|^r$ for $x > -1$. Let $c^* = -(\gamma + \log(1 - \gamma))/\gamma^r$. Then $g_{c^*}(-\gamma) = 0 = g_{c^*}(0)$.
Since $g_{c^*}(x) < 0$ for $-1 < x < 0$, it then follows that $g_{c^*}(x) > 0$ for all $-\gamma < x < 0$, and therefore $c^* \geq c_{r}(\gamma)$. If $c^* > c_{r}(\gamma)$, then $g_{c^*}(x) - \gamma < g_{c^*}(-\gamma) = 0$, contradicting the definition of $c_{r}(\gamma)$. Hence $c_{r}(\gamma) = c^*$. Take any $c \geq (r - 1)^{r-1}(2 - r)^{2-r}/r$. Then for all $x > 0$,

$$g_{c}(x) = \frac{1}{1 + cx - 1 + cx^{r-1}} \geq \frac{x}{1 + x} \left\{ -1 + c\inf_{y > 0} (y^{r-1} + y^{r-1}) \right\} = \frac{x}{1 + x} \left\{ -1 + \frac{c}{(r - 1)^{r-1}(2 - r)^{2-r}} \right\} \geq 0.$$

Since $g_{c}(0) = 0$, it then follows that $g_{c}(x) \geq 0$ for all $x \geq 0$. Hence $c_r \leq (r - 1)^{r-1}(2 - r)^{2-r}/r$.

To prove (ii), note that since $\lambda d_n \geq -\lambda M \geq -\gamma$ a.s. for $0 \leq \lambda \leq \gamma M^{-1}$, (i) yields

$$E[\exp\{\lambda d_n - c_{r}\lambda d_n\}]|\mathcal{F}_{n-1}] \leq E[1 + \lambda d_n]|\mathcal{F}_{n-1}] \leq 1 \text{ a.s.}. \quad \Box$$

The remainder of this section will be devoted to an upper half of the law of the iterated logarithm (LIL) for discrete-time supermartingales. First we prove, under the one-sided boundedness assumption of either Lemma 3.2 or Lemma 5.1, the following theorem giving an expectation form of the upper LIL, which in fact holds more generally when

$$\{\exp(\lambda A_t - \Phi_r(\lambda B_t)), t \in T\} \text{ is a supermartingale for } 0 \leq \lambda \leq \lambda_0,$$

where $T$ is either $\{0, 1, 2, \ldots\}$ (discrete-time case) or $[0, \infty)$ (continuous-time case) and $\Phi_r(x) = \frac{x^r}{r}$ for $1 < r \leq 2$. Applications of the theorem will be given in (5.9) and (5.10), and related results in the literature will also be discussed.

**THEOREM 5.2.** Let $T = \{0, 1, 2, \ldots\}$ or $T = [0, \infty)$, $1 < r \leq 2$, and $\Phi_r(x) = \frac{x^r}{r}$ for $x > 0$. Let $A_t, B_t$ be stochastic processes (on the same probability space) satisfying (5.1) and such that $B_t$ is positive and non-decreasing in $t > 0$, with $A_0 = 0$. In the case $T = [0, \infty)$, assume furthermore that $A_t$ and $B_t$ are right-continuous. Let $L : [1, \infty) \to (0, \infty)$ be a non-decreasing function satisfying (4.1)-(4.3). Let $\eta > 0$, $\lambda_0 \eta > \epsilon > 0$, and $h : (0, \infty) \to (0, \infty)$ be a non-decreasing function such that $h(x) \leq e^{\epsilon x}$ for all large $x$. Then there exists a constant $C$ depending only on $\lambda_0, \eta, \epsilon, h$ and $L$ such that

$$Eh(\sup_{t \geq 0} \{A_t(B_t \vee \eta)^{-r} \{1 \vee \log^{+} L(B_t \vee \eta)\}^{-1} \}^{(r-1)/r}) \leq C.$$

**PROOF.** It suffices to prove (5.2) with $\sup_{t \geq 0}$ replaced by $\sup_{s \geq t \geq 0}$ for every $s > 0$. Given any $s > 0$, there exists a sequence of nonnegative random times $\tau_n \leq s$ (in general not stopping times) such that

$$\lim_{n \to \infty} (B_{\tau_n} \vee \eta)^{1 \vee \log^{+} L(B_{\tau_n} \vee \eta)} \frac{A_n^{+}}{(B_{\tau_n} \vee \eta)^{1 \vee \log^{+} L(B_{\tau_n} \vee \eta)}} = \sup_{0 \leq t \leq s} \frac{A_t}{(B_t \vee \eta)^{1 \vee \log^{+} L(B_t \vee \eta)}}^{(r-1)/r},$$

$$17$$
since $A_0 = 0$. As in the proof of Theorem 4.6, we shall assume without loss of generality that $B_t \geq \eta$. Take any $q < 1$ such that $q\lambda_0 \eta > \varepsilon$.

It follows from Lemma 1 of Shao (2000) and Fatou’s lemma that for any nonnegative supermartingale $\{Y_t, t \in T\}$ (with right-continuous $Y_t$ in the case $T = [0, \infty)$), $E(\sup_{t \in T} Y_t)^q \leq (1-q)^{-1}(EY_0)^q$. Applying this result to (5.1) and noting that $A_0 = 0$, we obtain that for $0 \leq \lambda \leq \lambda_0$,

$$
(1-q)^{-1} \geq E\exp\{\sup_{t \in T} (\lambda A_t - \bar{\Phi}_r(\lambda B_t))\}^q \\
\geq E\exp\{q[\lambda A_{\tau_n} - \bar{\Phi}_r(\lambda B_{\tau_n})]\} = E\exp\{q\lambda A_{\tau_n} - \bar{\Psi}_r(q\lambda B_{\tau_n})\},
$$

where $\bar{\Psi}_r(x) = q^{1-r}x^r/r$.

Let $f_r(w) = \exp\{q(1-r)w r/(r-1)\}$ for $w > 0$. Note that in the notation of Theorem 4.6, $g_{y_w}(w) = y_w^{-1}f_r(w)$ with $y_w = qw^{1/(r-1)}$. Letting $A = A_{\tau_n}$ and $B = B_{\tau_n}$, it follows from (5.4) and (4.3) that

$$
(1-q)^{-1} \geq \int_0^{\lambda_0} E\exp\{q\lambda A - \bar{\Psi}_r(q\lambda B)\} \frac{d\lambda}{L(\lambda \lor \lambda_0^{-1})},
$$

which in turn yields the following analogues of (4.19) and (4.21), with $\eta > \tilde{\eta}$:

$$
(1-q)^{-1} \geq \frac{e^{-\bar{\Psi}_r(q\lambda_0 \eta)}}{3\lambda_1(\tilde{\eta}^{-1} \lor \eta)} \log \left( \frac{\eta}{\tilde{\eta}} \right) E\frac{e^{q\lambda_0 \bar{\eta} A/B}}{L(B \lor 1)},
$$

$$
(1-q)^{-1} \geq E \int_0^{q\lambda_0 B} \frac{\exp\{x A/B - \bar{\Psi}_r(x)\}}{xL(x^{1/(r-1)})} \frac{dx}{(A/B)^{1/(r-1)} L(B)}
\geq c(\lambda_0, q, \eta, r) E\frac{f_r(A/B)}{(A/B)^{1/(r-1)} L(B)}^{-1}((A/B)^{1/(r-1)} \leq \lambda_0 B),
$$

with $q\lambda_0 \tilde{\eta} > \varepsilon$, $\lambda_1 = \lambda_0 \lor \lambda_0^{-1}$ and the constant $c(\lambda_0, q, \eta, r)$ depending only on $\lambda_0, q, \eta$ and $r$. For (5.6), recall that $y_w = qw^{1/(r-1)}$ and $g_{y_w}(w) = y_w^{-1}f_r(w)$.

Take any $\delta < 1$ such that $\tau(1 - \delta)/(r - 1) > 1$. Since $q\lambda_0 \tilde{\eta} > \varepsilon$, there exists $x_0 > \lambda_0^{-1} \lor 1$ such that

$$
h(x) \leq e^{q\lambda_0 \bar{\eta} x/L(x)} \leq f_r^{-\delta}(x)/x^{1/(r-1)}
$$

for all $x \geq x_0$, noting that $L(x) \leq 3x L(1)$ by (4.1). Let

$$
F = \{f_r^{1/\delta}(A/B) \leq L(B) \lor \varepsilon\} = \{A^+ / B \leq [(1 \lor \log L(B))/(\delta q(1 - r^{-1}))]^{(r-1)/r}\}.
$$

Let $k$ be the smallest integer such that $2^k (r - 1) \geq 1$. On $\{A/B \geq x_0 \lor (\lambda_0 B)^{r-1}\}$,

$$
L(A/B) \geq L(x_0 \lor (\lambda_0 B)^{r-1}) \geq \frac{1}{3}(\lambda_0 \land 1)^{r-1} L(1 \lor B^{r-1}) \geq 3^{-(k+1)}(\lambda_0 \land 1)^{r-1} L(1 \lor B),
$$
where the last two inequalities follow from (4.1) and (4.2), respectively. From (5.7), it then follows that

\[
E h \left( \frac{A^+}{B \{1 \lor \log^+ L(B)\}^{(r-1)/r}} \right) \leq h(x_0) + h(1/[\delta q(1 - r^{-1})]^{(r-1)/r})P(F)
\]

\[
+ E1 \left( F^c \cap \left\{ \frac{A}{B} \geq x_0 \right\} \right) \left\{ \frac{e^{\lambda_0 h(A/B)}}{3^{-(k+1)}(\lambda_0 \land 1)^{r-1} L(B \lor 1)} \right\}^{1 - \left( \frac{A}{B} \right)^{1/(r-1)} \leq \lambda_0 B}
\]

\[
+ \left( \sup_{x \geq x_0} \frac{h(x) x^{1/(r-1)}}{f_r^1(x)} \right) \frac{f_r(A/B)}{(A/B)^{1/(r-1)} L(B)} 1 \left( \left( \frac{A}{B} \right)^{1/(r-1)} < \lambda_0 B \right)\right\},
\]

noting that \( f_r^1(A/B) > L(B) \) and therefore \( f_r^{1-\delta}(A/B) < f_r(A/B)/L(B) \) on \( F^c \). The desired conclusion follows from (5.5), (5.6) and (5.8).

Consider the case of continuous local martingales \( A_t \). We can apply Theorem 5.2 with \( r = 2 \) and \( B_t = \sqrt{\langle A \rangle_t} \), in view of Lemma 1.2. Putting \( h(x) = x^p \) in (5.2) in this case yields the following extension of (1.3) to the case \( q = p \): There exists for \( p > 0 \) an absolute constant \( C_p \) such that

\[
E \left( \sup_{t \geq 0} \frac{|A_t|}{\langle A_t \log \log \langle A_t \rangle_t \lor e^2 \rangle^{1/2}} \right)^p \leq C_p.
\]

Since (5.1) holds for all \( \lambda_0 > 0 \) by Lemma 1.2, we can in fact set \( \lambda_0 = \infty \) in (5.8) with \( r = 2 \) to replace it by

\[
E h \left( \frac{A^+}{B \{1 \lor \log^+ L(B)\}^{1/2}} \right) \leq h(x_0) + h(2/\delta q)^{1/2})P(F)
\]

\[
+ \sup_{x \geq x_0} \exp \left\{ \frac{h(x) x^{1/2}}{2(1-\delta)x^2} \right\} E1 \left( F^c \cap \left\{ \frac{A}{B} \geq x_0 \right\} \right) \exp \left\{ \frac{2}{A/B^2} \right\}
\]

so we only require \( h(x) \leq e^{\epsilon x^2} \) for some \( \epsilon < \frac{1}{2} \) and all large \( x \) in this case. Putting \( h(x) = \exp(\alpha x^2) \) with \( 0 < \alpha < \frac{1}{2} \) in the preceding argument then yields an absolute constant \( C(\alpha) \) such that

\[
E \left[ \sup_{t \geq 0} \exp \left( \alpha A_t^2 / \langle A_t \rangle_t \log \log \langle A_t \rangle_t \lor e^2 \right) \right] \leq C(\alpha),
\]

which can be regarded as an extension to \( p = 0 \) of the following result of Kikuchi (1991): For every \( p > 0 \) and \( 0 < \alpha < \frac{1}{2} \), there exists an absolute constant \( C_{\alpha,p} \) such that

\[
E[A^*_\infty \exp(\alpha A^*_\infty^2 / \langle A \rangle_\infty)] \leq C_{\alpha,p} E(A^*_\infty),
\]

where \( A^*_\infty = \sup_{t \geq 0} |A_t| \).

Let \( \ell_2(x) = 2 \log(x \lor e^2), x \geq 0 \). Clearly the expectation form (5.10) of the LIL implies the almost sure version \( \lim \sup_{t \to \infty} A_t / \langle A_t \rangle_t \ell_2(\langle A_t \rangle_t)^{1/2} < \infty \) a.s. for continuous local martingales.
Replacing Lemma 1.2 by Lemma 5.1 or Lemma 3.2 in the preceding argument, we also have a similar LIL for discrete-time supermartingales \( A_n = \sum_{i=1}^{n} d_i \) under the boundedness assumption \( d_i \geq -M \) a.s. (Lemma 5.1) or \( d_i \leq M \) a.s. (Lemma 3.2). Let

\[
\sigma_n^2 = \sum_{i=1}^{n} d_i^2, \quad \sigma_n^2 = \sum_{i=1}^{n} E(d_i^2 | \mathcal{F}_{i-1}).
\]

Instead of imposing a constant lower (or upper) bound for \( d_n \), we can let these bounds vary with \( n \) in the a.s. LIL, as in Stout (1973) who proved the following result by making use of Lemma 3.2.

**COROLLARY 5.2.** Let \( \{A_n = \sum_{i=1}^{n} d_i, \mathcal{F}_n, n \geq 1\} \) be a supermartingale such that \( d_n \leq m_n \) a.s. for some \( \mathcal{F}_{n-1} \) measurable random variable \( m_n \geq 0 \). Suppose \( \sigma_n^2 < \infty \) a.s. for every \( n \), \( \lim_{n \to \infty} \sigma_n^2 = \infty \) a.s. and

\[
\lim \sup_{n \to \infty} m_n \sqrt{\ell_2(\sigma_n^2)} / \sigma_n \leq K \text{ a.s.}
\]

for some positive constant \( K \leq \frac{1}{2} \). Then there exists a function \( c(\cdot) \) such that \( \lim_{K \to 0} c(K) = 1 \) and

\[
\lim \sup_{n \to \infty} A_n / \{\sigma_n \sqrt{\ell_2(\sigma_n^2)}\} \leq c(K) \text{ a.s.}
\]

If we normalize by \( s_n \) instead of \( \sigma_n \), we no longer need to assume \( \sigma_n^2 < \infty \) a.s. to prove the following self-normalized LIL.

**THEOREM 5.3.** Let \( \{A_n = \sum_{i=1}^{n} d_i, \mathcal{F}_n, n \geq 1\} \) be a supermartingale such that \( d_n \geq -m_n \) a.s. for some \( \mathcal{F}_{n-1} \) measurable random variable \( m_n \geq 0 \) satisfying

\[
\lim \sup_{n \to \infty} m_n \sqrt{\ell_2(\sigma_n^2)} / s_n \leq K \text{ a.s. on } \{\lim_{n \to \infty} s_n = \infty\},
\]

for some positive constant \( K \). Then \( \sup_{n \geq 1} A_n < \infty \) a.s. on \( \{\lim_{n \to \infty} s_n < \infty\} \) and

\[
\lim \sup_{n \to \infty} A_n / \{s_n \sqrt{\ell_2(\sigma_n^2)}\} \leq g(K) \text{ a.s. on } \{\lim_{n \to \infty} s_n = \infty\},
\]

where the function \( h(K) := Kg(K) \) is given by the positive solution of the equation

\[
h = K^2 / 2 + \log(1 + h).
\]

As \( K \to 0 \), \( h \) given by (5.14) satisfies \( h \sim K \). Therefore \( \lim_{K \to 0} g(K) = 1 \). Unlike Corollary 5.2, Theorem 5.3 imposes a lower bound on the supermartingale difference \( d_n \) but no upper bound on \( d_n \). This is reminiscent of the convergence theory for supermartingales that imposes \( L_1 \)-boundedness.
on the negative parts but no assumption on the positive parts. The following example shows that we cannot dispense with the boundedness assumption (5.12) in Theorem 5.3.

EXAMPLE 5.4. Let \( d_1, d_2, d_3, \ldots \) be independent random variables such that

\[
P\{d_n = \frac{1}{\sqrt{n}}\} = \frac{1}{2} + \sqrt{\frac{\log n}{n}}, \quad P\{d_n = -\frac{1}{\sqrt{n}}\} = \frac{1}{2} - \sqrt{\frac{\log n}{n}} - \frac{1}{n(\log n)^2}, \quad P\{d_n = -m_n\} = \frac{1}{n(\log n)^2}
\]

for \( n \geq 2 \), where \( m_n \sim 2(\log n)^{5/2} \) is chosen so that \( Ed_n = 0 \). Let \( X_n = d_n 1(|d_n| \leq 1) \). Then \( P\{X_n \neq d_n \text{ i.o.}\} = P\{d_n = -m_n \text{ i.o.}\} = 0 \). Hence with probability 1, \( s_n^2 = \sum_{i=1}^{n} X_i^2 + O(1) = \sum_{i=1}^{n} i^{-1} + O(1) = \log n + O(1) \). Since \( \bar{X}_i := X_i - EX_i \) are independent bounded random variables with zero means and \( \text{Var}(\bar{X}_i) \sim i^{-1} \), Kolmogorov's LIL yields

\[
\limsup_{n \to \infty} \left( \frac{\sum_{i=1}^{n} \bar{X}_i}{2(\log n)(\log \log \log n)} \right)^{1/2} = 1 \text{ a.s.}
\]

Since \( \sum_{i=1}^{n} EX_i \sim 2\sum_{i=1}^{n} i^{-1}(\log i)^{1/2} \sim \frac{4}{3}(\log n)^{3/2} \), this implies that with probability 1

\[
\frac{\sum_{i=1}^{n} d_i}{s_n \sqrt{\ell_2(s_n^2)}} \sim \frac{\sum_{i=1}^{n} X_i}{((\log n)(\log \log \log n))^{1/2}} \sim \frac{4(\log n)^{3/2}}{3((\log n)(\log \log \log n))^{1/2}} \to \infty.
\]

Note that \( m_n \sqrt{\ell_2(s_n^2)}/s_n \to \infty \). This shows that without the boundedness condition (5.12), the LIL need not hold for martingales self-normalized by \( s_n \). On the other hand, \( d_n \) is clearly bounded above and therefore satisfies the boundedness condition of Corollary 5.2. Note that \( \text{Var}(d_i) \sim 4(\log i)^3/i \) and therefore \( \sigma_n^2 \sim (\log n)^4 \), yielding

\[
\frac{\sum_{i=1}^{n} d_i}{\sigma_n \sqrt{\ell_2(s_n^2)}} \sim \frac{4(\log n)^{3/2}}{3(\log n)^2(\log \log \log n)^{1/2}} \to 0 \text{ a.s.,}
\]

which is consistent with the conclusion of Corollary 5.2.

PROOF OF THEOREM 5.3. Let \( u_n = \sqrt{\ell_2(s_n^2)} \). Since \( |d_n| \leq s_n \),

\[
(5.15) \quad d_n/(s_n u_n) \to 0 \text{ a.s. on } \{|\lim_{n \to \infty} s_n = \infty\}.
\]

Fix any \( \epsilon > 0 \). Let \( b = 1 + \frac{\epsilon}{2} \). For \( j = 1, 2, \ldots \), define the stopping times

\[
\tau_j = \inf\{n \geq 1 : s_n \geq b\} \quad (\inf \emptyset = \infty).
\]

On \( \{|\lim_{n \to \infty} s_n = \infty\} \), \( \tau_j < \infty \) for all \( j \). Define

\[
\Omega_j = \{\tau_j < \infty \text{ for all } j\} \cap \{|d_n| < \epsilon g(K)s_n u_n \text{ for all } n \geq \tau_j\}
\]

\[
\cap \{m_n \leq (1 + \epsilon^2)Ks_n/\sqrt{\ell_2(s_n)} \text{ for all } n \geq \tau_j\}.
\]

21
In view of (5.12) and (5.15), it suffices for the proof of (5.13) that for all small $\epsilon > 0$,

$$\lim_{J \to \infty} \sum_{j=J}^{\infty} P(\cup_{\tau_j < n \leq \tau_{j+1}} \{A_n \geq (1 + 3\epsilon)g(K)s_n u_n\} \cap \Omega_J) = 0.$$ 

Note that $\tau_j$ may equal $\tau_{j+1}$, in which case the union $\cup_{\tau_j < n \leq \tau_{j+1}}$ yields the empty set. Since on $\Omega_J \cap \{A_{\tau_{j+1} - 1} < (1 + 2\epsilon)g(K)s_{\tau_{j+1} - 1} u_{\tau_{j+1} - 1}\}$ we have $A_{\tau_{j+1}} = d_{\tau_{j+1}} + A_{\tau_{j+1} - 1} < (1 + 3\epsilon)g(K)s_{\tau_{j+1}} u_{\tau_{j+1}}$, for $j \geq J$, we need only show that

$$\lim_{J \to \infty} \sum_{j=J}^{\infty} P(\cup_{\tau_j < n \leq \tau_{j+1}} \{A_n \geq (1 + 2\epsilon)g(K)s_n u_n\} \cap \Omega_J) = 0.$$ 

For $\tau_j < n < \tau_{j+1}$ with $j \geq J$, we have $s_n < b^{j+1}$ and therefore $m_n \leq (1 + \epsilon^2)Kb^{j+1}/\sqrt{\ell_2(b^{2j})}$ on $\Omega_J$, implying that $d_n = \tilde{d}_n$ on $\Omega_J$, where we define

$$\alpha_j = g(K)\ell_2(b^{2j})/(bK), \quad \beta_j = Kb^{j+1}/\sqrt{\ell_2(b^{2j})},$$

$$\tilde{d}_n = d_n1(\cup_{j=1}^{\infty}\{m_n \leq (1 + \epsilon^2)\beta_j, \tau_j < n \leq \tau_{j+1}\}).$$

Since $m_n$ is $\mathcal{F}_{n-1}$-measurable,

$$E(\tilde{d}_n|\mathcal{F}_{n-1}) = E(d_n|\mathcal{F}_{n-1})1(\cup_{j=1}^{\infty}\{m_n \leq (1 + \epsilon^2)\beta_j\} \cap \{n \leq n - 1\} \cap \{\tau_{j+1} \leq n - 1\}) \leq 0,$$

so $\tilde{A}_n := \Sigma_{i=1}^{\infty}\tilde{d}_i$ is still a supermartingale. Let $\tilde{s}_n^2 = \Sigma_{i=1}^{\infty}\tilde{d}_i^2$. On $\Omega_J$, we have $\tilde{A}_n = A_n$ and $b^{j} \leq \tilde{s}_n = s_n < b^{j+1}$ for $\tau_j < n < \tau_{j+1}$ with $j \geq J$. Take any $0 < \gamma < 1$. Then for $j \geq J$,

$$P(\cup_{\tau_j < n \leq \tau_{j+1}} \{A_n \geq (1 + 2\epsilon)g(K)s_n u_n\} \cap \Omega_J)$$

$$\leq P(\cup_{\tau_j < n \leq \tau_{j+1}} \{\tilde{A}_n \geq (1 + 2\epsilon)\alpha_j\beta_j\} \cap \Omega_J)$$

$$\leq P\left(\cup_{\tau_j < n \leq \tau_{j+1}} \left\{ \frac{\gamma\tilde{A}_n}{(1 + \epsilon^2)\beta_j} - \frac{\gamma^2 c_7 \tilde{s}_n^2}{(1 + \epsilon^2)2\beta_j^2} \geq \frac{\gamma(1 + 2\epsilon)\alpha_j}{1 + \epsilon^2} - \frac{\gamma^2 c_7 b\alpha_j}{(1 + \epsilon^2)^2 K g(K)} \right\} \right)$$

$$\leq P\left(\sup_{\tau_j < n \leq \tau_{j+1}} \exp\left(\frac{\gamma\tilde{A}_n}{(1 + \epsilon^2)\beta_j} - \frac{\gamma^2 c_7 \tilde{s}_n^2}{(1 + \epsilon^2)2\beta_j^2} \right) \geq \exp\left((1 + \epsilon)\alpha_j\left(\gamma - \frac{\gamma^2 c_7}{K g(K)}\right)\right) \right)$$

for sufficiently small $\epsilon$, noting that $b^{2(j+1)}/\beta_j^2 = b\alpha_j/K g(K)$ and $b = 1 + \frac{\epsilon}{2}$. We can apply Lemma 5.1 (ii) and Doob's inequality for nonnegative supermartingales to conclude that the above probability is bounded by

$$\exp\left\{-(1 + \epsilon)\alpha_j\left(\gamma - \frac{\gamma^2 c_7}{K g(K)}\right)\right\} = \exp\left\{-\frac{2(1 + \epsilon)}{(1 + \epsilon/2)K} \left(\log \log b^{2j}\right) \left(\gamma - \frac{\gamma^2 c_7}{K g(K)}\right)\right\},$$

22
for $j$ sufficiently large. Therefore (5.16) holds if $g(K)$ is chosen so that

\[(5.19) \quad \sup_{0<\gamma<1} 2K^{-1} g(K) \{\gamma + (Kg(K))^{-1}(\gamma + \log(1 - \gamma))\} = 1,\]

recalling that $c_\gamma = -\{\gamma + \log(1 - \gamma)\}/\gamma^2$. The supremum in (5.19) is attained at $\gamma = Kg(K)/(1 + Kg(K))$, and therefore simple algebra reduces (5.19) to the equation (5.14) for $h(K) = Kg(K)$.

Note that $\{\lim_{n\to\infty} s_n < \infty\} = \cup_{i=1}^\infty \{\sum_{i=1}^\infty d_i^2 \leq J\}$ and that $m_n \leq \sqrt{J}$ for all $n$ on $\{\sum_{i=1}^\infty d_i^2 \leq J\}$. For fixed $J \geq 1$, let $\hat{d}_n = d_n 1(m_n \leq J), \hat{A}_n = \sum_{i=1}^n \hat{d}_i, \hat{s}_n^2 = \sum_{i=1}^n d_i^2$. Application of Lemma 5.1(ii) and Theorem 5.2 (with $r = 2$) to the supermartingale $\hat{A}_n$ shows that $\sup_{n\geq 1} \hat{A}_n/(\hat{s}_n \vee \eta)\sqrt{\ell_2(\hat{s}_n^2)} < \infty$ a.s., which implies that $\sup_{n\geq 1} A_n < \infty$ a.s. on $\{\sum_{i=1}^\infty d_i^2 \leq J\}$. $\Box$

Condition (5.12) in Theorem 5.3 can be restated in the following form, which is more convenient for applications since the lower bound $-m_n$ on $d_n$ is assumed to be $\mathcal{F}_{n-1}$-measurable:

\[(5.20) \quad \lim_{n\to\infty} \sup_{n\to\infty} m_n \sqrt{\ell_2(s_{n-1}^2)} / s_{n-1} \leq K \quad \text{a.s. on} \quad \{\lim_{n\to\infty} s_n = \infty\}.\]

Clearly (5.20) implies (5.12). Conversely, (5.12) implies that $m_n/s_n \to 0$ a.s. on $\{\lim_{n\to\infty} s_n = \infty\}$. On $\{d_n \geq 0\}$, the lower bound $d_n \geq -m_n$ is automatically satisfied for any choice of $m_n \geq 0$. On $\{-m_n \leq d_n \leq 0\} \cap \{\lim_{n\to\infty} s_n = \infty\}$, $d_n^2 \leq m_n^2$ and therefore $s_n^2 \leq m_n^2 + s_{n-1}^2 = s_{n-1}^2 + o(s_n^2)$ a.s., implying that $s_n \sim s_{n-1}$ a.s.

EXAMPLE 5.5. Let $X_1, X_2, \ldots$ be independent random variables such that $EX_n = 0$ and $X_n \geq -M$ a.s. for all $n \geq 1$, where $M$ is a nonnegative constant. Let $\mathcal{F}_n$ be the $\sigma$-field generated by $X_i, i \leq n$. Let $u_n \geq 0$ be $\mathcal{F}_{n-1}$-measurable and let $A_n = \sum_{i=1}^n u_i X_i, s_n^2 = \sum_{i=1}^n u_i^2 X_i^2$. Suppose (5.20) holds with $m_n = Mu_n$. Then the self-normalized LIL (5.13) holds for $A_n$. 

23
REFERENCES


