IRWIN AND JOAN JACOBS
CENTER FOR COMMUNICATION AND INFORMATION TECHNOLOGIES

Tightened Exponential Bounds for Discrete Time, Conditionally Symmetric Martingales with Bounded Jumps

Igal Sason

CCIT Report #816
October 2012

Electronics
Computers
Communications

DEPARTMENT OF ELECTRICAL ENGINEERING
TECHNION - ISRAEL INSTITUTE OF TECHNOLOGY, HAIFA 32000, ISRAEL

Tightened Exponential Bounds for Discrete Time, Conditionally Symmetric Martingales with Bounded Jumps

Igal Sason
Department of Electrical Engineering
Technion - Israel Institute of Technology
Haifa 32000, Israel
E-mail: sason@ee.technion.ac.il

Abstract

This letter derives some new exponential bounds for discrete time, real valued, conditionally symmetric martingales with bounded jumps. The new bounds are extended to conditionally symmetric sub/ supermartingales, and are compared to some existing bounds.

Keywords: discrete-time (sub/ super) martingales, large deviations, concentration inequalities.

AMS 2000 subject classifications: 60F10, 60G40, 60G42.

I. Introduction and Main Results

Classes of exponential bounds for discrete-time real-valued martingales were extensively studied in the literature (see, e.g., [1], [2], [4], [5], [6]–[9], [12], and [15]–[18]). This letter further assumes conditional symmetry of these martingales, as it is defined in the following:

Definition 1: Let $\{X_k, \mathcal{F}_k\}_{k\in\mathbb{N}_0}$, where $\mathbb{N}_0 \triangleq \mathbb{N} \cup \{0\}$, be a discrete-time and real-valued martingale, and let $\xi_k \triangleq X_k - X_{k-1}$ for every $k \in \mathbb{N}$ designate the jumps of the martingale. Then $\{X_k, \mathcal{F}_k\}_{k\in\mathbb{N}_0}$ is called a *conditionally symmetric martingale* if, conditioned on $\mathcal{F}_{k-1}$, the random variable ξ_k is symmetrically distributed around zero.

Our goal in this letter is to demonstrate how the assumption of the conditional symmetry improves existing exponential inequalities for discrete-time real-valued martingales with bounded increments. Earlier results, serving as motivation, appear in [7, Section 4] and [15, Section 6]. The new exponential bounds are also extended to conditionally symmetric sub or supermartingales, where the construction of these objects is exemplified later in this section. The relation of some of the exponential bounds derived in this work with some existing bounds is discussed later in this letter. Additional results addressing weak-type inequalities, maximal inequalities and ratio inequalities for conditionally symmetric martingales were derived in [13], [14] and [19].

A. Main Results

Our main results for conditionally symmetric martingales with bounded jumps are introduced in Theorems 1, 3 and 4. Theorems 2 and 5 are existing bounds, for general martingales without the conditional symmetry assumption, that are introduced in connection to the new theorems. Corollaries 1 and 2 provide an extension of the new results to conditionally symmetric sub/ supermartingales with bounded jumps. Our first result is the following theorem:

Theorem 1: Let $\{X_k, \mathcal{F}_k\}_{k\in\mathbb{N}_0}$ be a discrete-time real-valued and conditionally symmetric martingale. Assume that, for some fixed numbers $d, \sigma > 0$, the following two requirements are satisfied a.s.

$$|X_k - X_{k-1}| \le d, \qquad \text{Var}(X_k|\mathcal{F}_{k-1}) = \mathbb{E}\big[(X_k - X_{k-1})^2 \,|\, \mathcal{F}_{k-1}\big] \le \sigma^2 \tag{1}$$

for every $k \in \mathbb{N}$. Then, for every $\alpha \ge 0$ and $n \in \mathbb{N}$,

$$\mathbb{P}\left(\max_{1\le k\le n} |X_k - X_0| \ge \alpha n\right) \le 2\exp\big(-nE(\gamma,\delta)\big) \tag{2}$$

where

$$\gamma \triangleq \frac{\sigma^2}{d^2}, \quad \delta \triangleq \frac{\alpha}{d} \tag{3}$$

and for $\gamma \in (0,1]$ and $\delta \in [0,1)$

$$E(\gamma, \delta) \triangleq \delta x - \ln\Big(1 + \gamma\big[\cosh(x) - 1\big]\Big) \tag{4}$$

$$x \triangleq \ln\left(\frac{\delta(1-\gamma) + \sqrt{\delta^2(1-\gamma)^2 + \gamma^2(1-\delta^2)}}{\gamma(1-\delta)}\right). \tag{5}$$

If $\delta > 1$, then the probability on the left-hand side of (2) is zero (so $E(\gamma, \delta) \triangleq +\infty$), and $E(\gamma, 1) = \ln\big(\frac{2}{\gamma}\big)$. Furthermore, the exponent $E(\gamma, \delta)$ is asymptotically optimal in the sense that there exists a conditionally symmetric martingale, satisfying the conditions in (1) a.s., that attains this exponent in the limit where $n \to \infty$.

Remark 1: From the above conditions, without any loss of generality, $\sigma^2 \leq d^2$ and therefore $\gamma \in (0, 1]$. This implies that Theorem 1 characterizes the exponent $E(\gamma, \delta)$ for all values of γ and δ.

Theorem 1 should be compared to the statement in [11, Theorem 6.1] (see also [6, Corollary 2.4.7]), which does not require the conditional symmetry property. It gives the following result:

Theorem 2: Let $\{X_k, \mathcal{F}_k\}_{k \in \mathbb{N}_0}$ be a discrete-time real-valued martingale with bounded jumps. Assume that the two conditions in (1) are satisfied a.s. for every $k \in \mathbb{N}$. Then, for every $\alpha \geq 0$ and $n \in \mathbb{N}$,

$$\mathbb{P}\left(\max_{1 \leq k \leq n} |X_k - X_0| \geq \alpha n\right) \leq 2 \exp\left(-n\, D\Big(\frac{\delta + \gamma}{1+\gamma} \Big|\Big| \frac{\gamma}{1+\gamma}\Big)\right) \tag{6}$$

where γ and δ are introduced in (3), and

$$D(p||q) \triangleq p \ln\Big(\frac{p}{q}\Big) + (1-p) \ln\Big(\frac{1-p}{1-q}\Big), \quad \forall\, p, q \in [0, 1] \tag{7}$$

is the divergence (a.k.a. relative entropy or Kullback-Leibler distance) to the natural base between the two probability distributions $(p, 1-p)$ and $(q, 1-q)$. If $\delta > 1$, then the probability on the left-hand side of (6) is zero. Furthermore, the exponent on the right-hand side of (6) is asymptotically optimal under the assumptions of this theorem.

Remark 2: The two exponents in Theorems 1 and 2 are both discontinuous at $\delta = 1$. This is consistent with the assumption of the bounded jumps that implies that $\mathbb{P}(|X_n - X_0| \geq nd\delta)$ is equal to zero if $\delta > 1$.

If $\delta \to 1^-$ then, from (4) and (5), for every $\gamma \in (0, 1]$,

$$\lim_{\delta \to 1^-} E(\gamma, \delta) = \lim_{x \to \infty} \big[x - \ln\big(1 + \gamma(\cosh(x) - 1)\big)\big] = \ln\left(\frac{2}{\gamma}\right). \tag{8}$$

On the other hand, the right limit at $\delta = 1$ is infinity since $E(\gamma, \delta) = +\infty$ for every $\delta > 1$. The same discontinuity also exists for the exponent in Theorem 2 where the right limit at $\delta = 1$ is infinity, and the left limit is equal to

$$\lim_{\delta \to 1^-} D\Big(\frac{\delta + \gamma}{1+\gamma} \Big|\Big| \frac{\gamma}{1+\gamma}\Big) = \ln\left(1 + \frac{1}{\gamma}\right) \tag{9}$$

where the last equality follows from (7). A comparison of the limits in (8) and (9) is consistent with the improvement that is obtained in Theorem 1 as compared to Theorem 2 due to the additional assumption of the conditional symmetry that is relevant if $\gamma \in (0, 1)$. It can be verified that the two exponents coincide if $\gamma = 1$ (which is equivalent to removing the constraint on the conditional variance), and their common value is equal to

$$f(\delta) = \begin{cases} \ln(2)\Big[1 - h_2\left(\frac{1-\delta}{2}\right)\Big], & 0 \leq \delta \leq 1 \\ +\infty, & \delta > 1 \end{cases} \tag{10}$$

where $h_2(x) \triangleq -x \log_2(x) - (1-x)\log_2(1-x)$ for $0 \leq x \leq 1$ denotes the binary entropy function to the base 2.

Theorem 1 provides an improvement over the bound in Theorem 2 for conditionally symmetric martingales with bounded jumps. The bounds in Theorems 1 and 2 depend on the conditional variance of the martingale, but they do not take into consideration conditional moments of higher orders. The following bound generalizes the bound in Theorem 1, but it does not admit in general a closed-form expression.

Theorem 3: Let $\{X_k, \mathcal{F}_k\}_{k \in \mathbb{N}_0}$ be a discrete-time and real-valued conditionally symmetric martingale. Let $m \in \mathbb{N}$ be an even number, and assume that the following conditions hold a.s. for every $k \in \mathbb{N}$

$$|X_k - X_{k-1}| \leq d, \qquad \mathbb{E}\big[(X_k - X_{k-1})^l \,|\, \mathcal{F}_{k-1}\big] \leq \mu_l, \ \ \forall\, l \in \{2, 4, \ldots, m\}$$

for some $d > 0$ and non-negative numbers $\{\mu_2, \mu_4, \ldots, \mu_m\}$. Then, for every $\alpha \geq 0$ and $n \in \mathbb{N}$,

$$\mathbb{P}\left(\max_{1 \leq k \leq n} |X_k - X_0| \geq \alpha n\right) \leq 2 \left\{\min_{x \geq 0} e^{-\delta x} \left[1 + \sum_{l=1}^{\frac{m}{2}-1} \frac{\gamma_{2l}\, x^{2l}}{(2l)!} + \gamma_m\big(\cosh(x) - 1\big)\right]\right\}^n \tag{11}$$

where

$$\delta \triangleq \frac{\alpha}{d}, \quad \gamma_{2l} \triangleq \frac{\mu_{2l}}{d^{2l}}, \quad \forall\, l \in \Big\{1, \ldots, \frac{m}{2}\Big\}. \tag{12}$$

We consider in the following a different type of exponential inequalities for conditionally symmetric martingales with bounded jumps.

Theorem 4: Let $\{X_n, \mathcal{F}_n\}_{n \in \mathbb{N}_0}$ be a discrete-time real-valued and conditionally symmetric martingale. Assume that there exists a fixed number $d > 0$ such that $\xi_k \triangleq X_k - X_{k-1} \leq d$ a.s. for every $k \in \mathbb{N}$. Let

$$Q_n \triangleq \sum_{k=1}^{n} \mathbb{E}[\xi_k^2 \,|\, \mathcal{F}_{k-1}] \tag{13}$$

with $Q_0 \triangleq 0$, be the predictable quadratic variation of the martingale up to time n. Then, for every $z, r > 0$,

$$\mathbb{P}\left(\max_{1 \leq k \leq n} (X_k - X_0) \geq z,\, Q_n \leq r \;\text{ for some } n \in \mathbb{N}\right) \leq \exp\left(-\frac{z^2}{2r} \cdot C\left(\frac{zd}{r}\right)\right) \tag{14}$$

where

$$C(u) \triangleq \frac{2[u \sinh^{-1}(u) - \sqrt{1+u^2} + 1]}{u^2}, \quad \forall\, u > 0. \tag{15}$$

Theorem 4 should be compared to [8, Theorem 1.6] (see also [6, Exercise 2.4.21(b)]) that was stated without the requirement for the conditional symmetry of the martingale. It provides the following result:

Theorem 5: Let $\{X_n, \mathcal{F}_n\}_{n \in \mathbb{N}_0}$ be a discrete-time real-valued martingale. Assume that there exists a fixed number $d > 0$ such that $\xi_k \triangleq X_k - X_{k-1} \leq d$ a.s. for every $k \in \mathbb{N}$. Then, for every $z, r > 0$,

$$\mathbb{P}\left(\max_{1 \leq k \leq n} (X_k - X_0) \geq z,\, Q_n \leq r \;\text{ for some } n \in \mathbb{N}\right) \leq \exp\left(-\frac{z^2}{2r} \cdot B\left(\frac{zd}{r}\right)\right) \tag{16}$$

where

$$B(u) \triangleq \frac{2[(1+u)\ln(1+u) - u]}{u^2}, \quad \forall\, u > 0. \tag{17}$$

The proof of [8, Theorem 1.6] is modified by using Bennett's inequality for the derivation of the original bound in Theorem 5 (without the conditional symmetry requirement). Furthermore, this modified proof serves to derive the improved bound in Theorem 4 under the conditional symmetry assumption.

Extension of the inequalities to discrete-time, real-valued, and conditionally symmetric sub/ supermartingales:

Definition 2: Let $\{X_k, \mathcal{F}_k\}_{k \in \mathbb{N}_0}$ be a discrete-time real-valued sub or supermartingale, and $\eta_k \triangleq X_k - \mathbb{E}[X_k | \mathcal{F}_{k-1}]$ for every $k \in \mathbb{N}$. Then the martingale $\{X_k, \mathcal{F}_k\}_{k \in \mathbb{N}_0}$ is called, respectively, a conditionally symmetric sub or supermartingale if, conditioned on $\mathcal{F}_{k-1}$, the random variable η_k is symmetrically distributed around zero.

Remark 3: For martingales, $\eta_k = \xi_k$ for every $k \in \mathbb{N}$, so we obtain consistency with Definition 1.

An extension of Theorem 1 to conditionally symmetric sub and supermartingales is introduced in the following:

Corollary 1: Let $\{X_k, \mathcal{F}_k\}_{k \in \mathbb{N}_0}$ be a discrete-time, real-valued and conditionally symmetric supermartingale. Assume that, for some constants $d, \sigma > 0$, the following two requirements are satisfied a.s.

$$\eta_k \leq d, \qquad \text{Var}(X_k | \mathcal{F}_{k-1}) \triangleq \mathbb{E}\big[\eta_k^2 \,|\, \mathcal{F}_{k-1}\big] \leq \sigma^2 \tag{18}$$

for every $k \in \mathbb{N}$. Then, for every $\alpha \geq 0$ and $n \in \mathbb{N}$,

$$\mathbb{P}\Big(\max_{1 \leq k \leq n} (X_k - X_0) \geq \alpha n\Big) \leq \exp\big(-n\, E(\gamma, \delta)\big) \tag{19}$$

where γ and δ are defined in (3), and $E(\gamma, \delta)$ is introduced in (4). Alternatively, if $\{X_k, \mathcal{F}_k\}_{k \in \mathbb{N}_0}$ is a conditionally symmetric submartingale, the same bound holds for $\mathbb{P}\big(\min_{1 \le k \le n}(X_k - X_0) \le -\alpha n\big)$ provided that $\eta_k \ge -d$ and the second condition in (18) hold a.s. for every $k \in \mathbb{N}$. If $\delta > 1$, then these two probabilities are zero.

The following statement extends Theorem 4 to conditionally symmetric supermartingales.

Corollary 2: Let $\{X_n, \mathcal{F}_n\}_{n \in \mathbb{N}_0}$ be a discrete-time, real-valued supermartingale. Assume that there exists a fixed number $d > 0$ such that $\eta_k \le d$ a.s. for every $k \in \mathbb{N}$. Let $\{Q_n\}_{n \in \mathbb{N}_0}$ be the predictable quadratic variations of the supermartingale, i.e., $Q_n \triangleq \sum_{k=1}^n \mathbb{E}[\eta_k^2 \,|\, \mathcal{F}_{k-1}]$ for every $n \in \mathbb{N}$ with $Q_0 \triangleq 0$. Then, the result in (16) holds. Furthermore, if the supermartingale is conditionally symmetric, then the improved bound in (14) holds.

B. Construction of Discrete-Time, Real-Valued and Conditionally Symmetric Sub/ Supermartingales

Before proving the tightened inequalities for discrete-time conditionally symmetric sub/ supermartingales, it is in place to exemplify the construction of these objects.

Example 1: Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space, and let $\{U_k\}_{k \in \mathbb{N}} \subseteq L^1(\Omega, \mathcal{F}, \mathbb{P})$ be a sequence of independent random variables with zero mean. Let $\{\mathcal{F}_k\}_{k \ge 0}$ be the natural filtration of sub σ-algebras of $\mathcal{F}$, where

$$\mathcal{F}_0 = \{\emptyset, \Omega\}, \quad \mathcal{F}_k = \sigma(U_1, \ldots, U_k), \ \forall\, k \in \mathbb{N}.$$

Furthermore, for $k \in \mathbb{N}$, let $A_k \in L^\infty(\Omega, \mathcal{F}_{k-1}, \mathbb{P})$ be an $\mathcal{F}_{k-1}$-measurable random variable with a finite essential supremum. Define a new sequence of random variables in $L^1(\Omega, \mathcal{F}, \mathbb{P})$ where

$$X_n = \sum_{k=1}^n A_k U_k, \ \ \forall\, n \in \mathbb{N}$$

and $X_0 = 0$. Then, $\{X_n, \mathcal{F}_n\}_{n \in \mathbb{N}_0}$ is a martingale. Lets assume that the random variables $\{U_k\}_{k \in \mathbb{N}}$ are symmetrically distributed around zero. Note that $X_n = X_{n-1} + A_n U_n$ where A_n is $\mathcal{F}_{n-1}$-measurable and U_n is independent of the σ-algebra $\mathcal{F}_{n-1}$ (due to the independence of the random variables $U_1, \ldots, U_n$). It therefore follows that for every $n \in \mathbb{N}$, given $\mathcal{F}_{n-1}$, the random variable X_n is symmetrically distributed around its conditional expectation X_{n-1}. Hence, the martingale $\{X_n, \mathcal{F}_n\}_{n \in \mathbb{N}_0}$ is conditionally symmetric.

Example 2: In continuation to Example 1, let $\{X_n, \mathcal{F}_n\}_{n \in \mathbb{N}_0}$ be a martingale, and define $Y_0 = 0$ and

$$Y_n = \sum_{k=1}^n A_k (X_k - X_{k-1}), \quad \forall\, n \in \mathbb{N}.$$

The sequence $\{Y_n, \mathcal{F}_n\}_{n \in \mathbb{N}_0}$ is a martingale. If $\{X_n, \mathcal{F}_n\}_{n \in \mathbb{N}_0}$ is a conditionally symmetric martingale then also the martingale $\{Y_n, \mathcal{F}_n\}_{n \in \mathbb{N}_0}$ is conditionally symmetric (since $Y_n = Y_{n-1} + A_n(X_n - X_{n-1})$, and by assumption A_n is $\mathcal{F}_{n-1}$-measurable).

Example 3: In continuation to Example 1, let $\{U_k\}_{k \in \mathbb{N}}$ be independent random variables with a symmetric distribution around their expected value, and also assume that $\mathbb{E}(U_k) \le 0$ for every $k \in \mathbb{N}$. Furthermore, let $A_k \in L^\infty(\Omega, \mathcal{F}_{k-1}, \mathbb{P})$, and assume that a.s. $A_k \ge 0$ for every $k \in \mathbb{N}$. Let $\{X_n, \mathcal{F}_n\}_{n \in \mathbb{N}_0}$ be a martingale as defined in Example 1. Note that $X_n = X_{n-1} + A_n U_n$ where A_n is non-negative and $\mathcal{F}_{n-1}$-measurable, and U_n is independent of $\mathcal{F}_{n-1}$ and symmetrically distributed around its average. This implies that $\{X_n, \mathcal{F}_n\}_{n \in \mathbb{N}_0}$ is a conditionally symmetric supermartingale.

Example 4: In continuation to Examples 2 and 3, let $\{X_n, \mathcal{F}_n\}_{n \in \mathbb{N}_0}$ be a conditionally symmetric supermartingale. Define $\{Y_n\}_{n \in \mathbb{N}_0}$ as in Example 2 where A_k is non-negative a.s. and $\mathcal{F}_{k-1}$-measurable for every $k \in \mathbb{N}$. Then $\{Y_n, \mathcal{F}_n\}_{n \in \mathbb{N}_0}$ is a conditionally symmetric supermartingale.

Example 5: Consider a standard Brownian motion $(W_t)_{t \ge 0}$. Define, for some $T > 0$, the discrete-time process

$$X_n = W_{nT}, \quad \mathcal{F}_n = \sigma(\{W_t\}_{0 \le t \le nT}), \quad \forall\, n \in \mathbb{N}_0.$$

The increments of $(W_t)_{t \ge 0}$ over time intervals $[t_{k-1}, t_k]$ are statistically independent if these intervals do not overlap (except of their endpoints), and they are Gaussian distributed with a zero mean and variance $t_k - t_{k-1}$. The random variable $\xi_n \triangleq X_n - X_{n-1}$ is therefore statistically independent of $\mathcal{F}_{n-1}$, and it is Gaussian distributed with a zero mean and variance T. The martingale $\{X_n, \mathcal{F}_n\}_{n \in \mathbb{N}_0}$ is therefore conditionally symmetric.

II. Proofs

A. Combined Proof of Theorems 1 and 2

We rely here on the proof of the existing bound that is stated in Theorem 2, for discrete-time real-valued martingales with bounded jumps (see [11, Theorem 6.1] and [6, Corollary 2.4.7]), and then deviate from this proof at the point where the additional property of the conditional symmetry of the martingale is taken into consideration for the derivation of the improved exponential inequality in Theorem 1.

Write $X_n - X_0 = \sum_{k=1}^n \xi_k$ where $\xi_k \triangleq X_k - X_{k-1}$ for $k \in \mathbb{N}$. Since $\{X_k - X_0, \mathcal{F}_k\}_{k \in \mathbb{N}_0}$ is a martingale, $h(x) = \exp(tx)$ is a convex function on $\mathbb{R}$ for every $t \in \mathbb{R}$, and a composition of a convex function with a martingale gives a submartingale w.r.t. the same filtration, then $\bigl\{\exp(t(X_k - X_0)), \mathcal{F}_k\bigr\}_{k \in \mathbb{N}_0}$ is a sub-martingale for every $t \in \mathbb{R}$. By applying the maximal inequality for submartingales, then for every $\alpha \geq 0$ and $n \in \mathbb{N}$

$$\begin{aligned}
&\mathbb{P}\Bigl(\max_{1 \leq k \leq n} (X_k - X_0) \geq \alpha n\Bigr) \\
&= \mathbb{P}\Bigl(\max_{1 \leq k \leq n} \exp\left(t(X_k - X_0)\right) \geq \exp(\alpha n t)\Bigr) \qquad \forall\, t \geq 0 \\
&\leq \exp(-\alpha n t)\, \mathbb{E}\Bigl[\exp\bigl(t(X_n - X_0)\bigr)\Bigr] \\
&= \exp(-\alpha n t)\, \mathbb{E}\left[\exp\left(t \sum_{k=1}^n \xi_k\right)\right]
\end{aligned} \tag{20}$$

Furthermore,

$$\begin{aligned}
&\mathbb{E}\left[\exp\left(t \sum_{k=1}^n \xi_k\right)\right] \\
&= \mathbb{E}\left[\mathbb{E}\left[\exp\left(t \sum_{k=1}^n \xi_k\right) \,|\, \mathcal{F}_{n-1}\right]\right] \\
&= \mathbb{E}\left[\exp\left(t \sum_{k=1}^{n-1} \xi_k\right) \mathbb{E}\bigl[\exp(t\xi_n) \,|\, \mathcal{F}_{n-1}\bigr]\right]
\end{aligned} \tag{21}$$

where the last transition holds since $\exp\bigl(t \sum_{k=1}^{n-1} \xi_k\bigr)$ is $\mathcal{F}_{n-1}$-measurable. The proof of Theorem 2 proceeds by applying Bennett's inequality [3] (see, e.g., [6, Lemma 2.4.1]) for the conditional law of ξ_k given the σ-algebra $\mathcal{F}_{k-1}$. Since $\mathbb{E}[\xi_k | \mathcal{F}_{k-1}] = 0$, $\text{Var}[\xi_k | \mathcal{F}_{k-1}] \leq \sigma^2$ and $\xi_k \leq d$ a.s. for $k \in \mathbb{N}$, then a.s. (see (3))

$$\mathbb{E}\left[\exp(t\xi_k) \,|\, \mathcal{F}_{k-1}\right] \leq \frac{\gamma \exp(td) + \exp(-\gamma t d)}{1+\gamma}\,, \quad \forall\, t \geq 0. \tag{22}$$

By induction, from (21) and (22), it follows that for every $t \geq 0$

$$\mathbb{E}\left[\exp\left(t \sum_{k=1}^n \xi_k\right)\right] \leq \left(\frac{\gamma \exp(td) + \exp(-\gamma t d)}{1+\gamma}\right)^n. \tag{23}$$

Combining (20) and (23), followed by an optimization of the bound on $\mathbb{P}\Bigl(\max_{1 \leq k \leq n} (X_k - X_0) \geq \alpha n\Bigr)$ over $t \geq 0$, gives that

$$\mathbb{P}\left(\max_{1 \leq k \leq n} (X_k - X_0) \geq \alpha n\right) \leq \exp\left(-n\, D\left(\frac{\delta+\gamma}{1+\gamma} \Big|\Big| \frac{\gamma}{1+\gamma}\right)\right). \tag{24}$$

Applying (24) to the martingale $\{-X_k, \mathcal{F}_k\}_{k \in \mathbb{N}_0}$ gives the same bound on $\mathbb{P}(\min_{1 \leq k \leq n}(X_k - X_0) \leq -\alpha n)$ for an arbitrary $\alpha \geq 0$. The union bound implies that

$$\mathbb{P}\left(\max_{1 \leq k \leq n} |X_k - X_0| \geq \alpha n\right) \leq \mathbb{P}\left(\max_{1 \leq k \leq n} (X_k - X_0) \geq \alpha n\right) + \mathbb{P}\left(\min_{1 \leq k \leq n} (X_k - X_0) \leq -\alpha n\right). \tag{25}$$

This doubles the bound on the right-hand side of (24), thus proving the exponential bound in Theorem 2.

In order to prove Theorem 1 for a discrete-time, real-valued and conditionally symmetric martingale with bounded jumps, we deviate from the proof of Theorem 2. This is done by a replacement of Bennett's inequality for the conditional expectation in (22) with a tightened bound under the conditional symmetry assumption. To this end, we need a lemma to proceed.

Lemma 1: Let X be a real-valued RV with a symmetric distribution around zero, a support $[-d, d]$, and assume that $\mathbb{E}[X^2] = \text{Var}(X) \leq \gamma d^2$ for some $d > 0$ and $\gamma \in [0, 1]$. Let h be a real-valued convex function, and assume that $h(d^2) \geq h(0)$. Then

$$\mathbb{E}[h(X^2)] \leq (1 - \gamma)h(0) + \gamma h(d^2) \tag{26}$$

where equality holds for the symmetric distribution

$$\mathbb{P}(X = d) = \mathbb{P}(X = -d) = \frac{\gamma}{2}, \quad \mathbb{P}(X = 0) = 1 - \gamma. \tag{27}$$

Proof: Since h is convex and $\text{supp}(X) = [-d, d]$, then a.s. $h(X^2) \leq h(0) + \left(\frac{X}{d}\right)^2 \big(h(d^2) - h(0)\big)$. Taking expectations on both sides gives (26), which holds with equality for the symmetric distribution in (27). ■

Corollary 3: If X is a random variable that satisfies the three requirements in Lemma 1 then, for every $\lambda \in \mathbb{R}$,

$$\mathbb{E}\big[\exp(\lambda X)\big] \leq 1 + \gamma\big[\cosh(\lambda d) - 1\big] \tag{28}$$

and (28) holds with equality for the symmetric distribution in Lemma 1, independently of the value of λ.

Proof: For every $\lambda \in \mathbb{R}$, due to the symmetric distribution of X, $\mathbb{E}\big[\exp(\lambda X)\big] = \mathbb{E}\big[\cosh(\lambda X)\big]$. The claim now follows from Lemma 1 since, for every $x \in \mathbb{R}$, $\cosh(\lambda x) = h(x^2)$ where $h(x) \triangleq \sum_{n=0}^{\infty} \frac{\lambda^{2n}|x|^n}{(2n)!}$ is a convex function (h is convex since it is a linear combination, with non-negative coefficients, of convex functions), and $h(d^2) = \cosh(\lambda d) \geq 1 = h(0)$. ■

We continue with the proof of Theorem 1. Under the assumption of this theorem, for every $k \in \mathbb{N}$, the random variable $\xi_k \triangleq X_k - X_{k-1}$ satisfies a.s. $\mathbb{E}[\xi_k \,|\, \mathcal{F}_{k-1}] = 0$ and $\mathbb{E}[(\xi_k)^2 \,|\, \mathcal{F}_{k-1}] \leq \sigma^2$. Applying Corollary 3 for the conditional law of ξ_k given $\mathcal{F}_{k-1}$, it follows that for every $k \in \mathbb{N}$ and $t \in \mathbb{R}$

$$\mathbb{E}\left[\exp(t\xi_k) \,|\, \mathcal{F}_{k-1}\right] \leq 1 + \gamma\big[\cosh(td) - 1\big] \tag{29}$$

holds a.s., and therefore it follows from (21) and (29) that for every $t \in \mathbb{R}$

$$\mathbb{E}\left[\exp\left(t\sum_{k=1}^{n} \xi_k\right)\right] \leq \Big(1 + \gamma\big[\cosh(td) - 1\big]\Big)^n. \tag{30}$$

Therefore, from (20), for every $t \geq 0$,

$$\mathbb{P}\Big(\max_{1 \leq k \leq n} (X_k - X_0) \geq \alpha n\Big) \leq \exp(-\alpha n t)\Big(1 + \gamma\big[\cosh(td) - 1\big]\Big)^n. \tag{31}$$

From (3) and a replacement of td with x, then for an arbitrary $\alpha \geq 0$ and $n \in \mathbb{N}$

$$\mathbb{P}\Big(\max_{1 \leq k \leq n} (X_k - X_0) \geq \alpha n\Big) \leq \inf_{x \geq 0} \left\{\exp\left(-n\Big[\delta x - \ln\big(1 + \gamma\big[\cosh(x) - 1\big]\big)\Big]\right)\right\}. \tag{32}$$

An optimization over the non-negative parameter x gives the solution for the optimized parameter in (5). Applying (32) to the martingale $\{-X_k, \mathcal{F}_k\}_{k \in \mathbb{N}_0}$ gives the same bound on $\mathbb{P}(\min_{1 \leq k \leq n}(X_k - X_0) \leq -\alpha n)$. Finally, using the union bound in (25) completes the proof of the exponential bound in Theorem 1.

Proof for the asymptotic optimality of the exponents in Theorems 1 and 2: In the following, we show that under the conditions of Theorem 1, the exponent $E(\gamma, \delta)$ in (4) and (5) is asymptotically optimal. To show this, let $d > 0$ and $\gamma \in (0, 1]$, and let $U_1, U_2, \ldots$ be i.i.d. random variables whose probability distribution is given by

$$\mathbb{P}(U_i = d) = \mathbb{P}(U_i = -d) = \frac{\gamma}{2}, \quad \mathbb{P}(U_i = 0) = 1 - \gamma, \quad \forall\, i \in \mathbb{N}. \tag{33}$$

Consider the particular case of the conditionally symmetric martingale $\{X_n, \mathcal{F}_n\}_{n \in \mathbb{N}_0}$ in Example 1 (see Section I-B) where $X_n \triangleq \sum_{i=1}^{n} U_i$ for $n \in \mathbb{N}$, and $X_0 \triangleq 0$. It follows that $|X_n - X_{n-1}| \leq d$ and $\text{Var}(X_n|\mathcal{F}_{n-1}) = \gamma d^2$ a.s. for every $n \in \mathbb{N}$. From Cramér's theorem in $\mathbb{R}$, for every $\alpha \geq \mathbb{E}[U_1] = 0$,

$$\begin{aligned} & \lim_{n \to \infty} \frac{1}{n} \ln \mathbb{P}(X_n - X_0 \geq \alpha n) \\ & = \lim_{n \to \infty} \frac{1}{n} \ln \mathbb{P}\biggl(\frac{1}{n} \sum_{i=1}^{n} U_i \geq \alpha\biggr) \\ & = -I(\alpha) \end{aligned} \tag{34}$$

where the rate function is given by

$$I(\alpha) = \sup_{t \geq 0} \{t\alpha - \ln \mathbb{E}[\exp(tU_1)]\} \tag{35}$$

(see, e.g., [6, Theorem 2.2.3] and [6, Lemma 2.2.5(b)] for the restriction of the supermum to the interval $[0, \infty)$). From (33) and (35), for every $\alpha \geq 0$,

$$I(\alpha) = \sup_{t \geq 0} \bigl\{t\alpha - \ln\bigl(1 + \gamma[\cosh(td) - 1]\bigr)\bigr\}$$

but it is equivalent to the optimized exponent on the right-hand side of (31), giving the exponent of the bound in Theorem 1. Hence, $I(\alpha) = E(\gamma, \delta)$ in (4) and (5). This proves that the exponent of the bound in Theorem 1 is indeed asymptotically optimal in the sense that there exists a discrete-time, real-valued and conditionally symmetric martingale, satisfying the conditions in (1) a.s., that attains this exponent in the limit where $n \to \infty$. The proof for the asymptotic optimality of the exponent in Theorem 2 (see the right-hand side of (6)) is similar to the proof for Theorem 1, except that the i.i.d. random variables $U_1, U_2, \ldots$ are now distributed as follows:

$$\mathbb{P}(U_i = d) = \frac{\gamma}{1+\gamma}, \quad \mathbb{P}(U_i = -\gamma d) = \frac{1}{1+\gamma}, \quad \forall\, i \in \mathbb{N}$$

and, as before, the martingale $\{X_n, \mathcal{F}_n\}_{n \in \mathbb{N}_0}$ is defined by $X_n = \sum_{i=1}^n U_i$ and $\mathcal{F}_n = \sigma(U_1, \ldots, U_n)$ for every $n \in \mathbb{N}$ with $X_0 = 0$ and $\mathcal{F}_0 = \{\emptyset, \Omega\}$ (in this case, it is not a conditionally symmetric martingale unless $\gamma = 1$).

B. Proof of Theorem 3

The starting point of the proof of Theorem 3 relies on (20) and (21). For every $k \in \mathbb{N}$ and $t \in \mathbb{R}$, since $\mathbb{E}\bigl[\xi_k^{2l-1} \,|\, \mathcal{F}_{k-1}\bigr] = 0$ for every $l \in \mathbb{N}$ (due to the conditionally symmetry property of the martingale),

$$\begin{aligned} & \mathbb{E}\bigl[\exp(t\xi_k) | \mathcal{F}_{k-1}\bigr] \\ & = 1 + \sum_{l=1}^{\frac{m}{2}-1} \frac{t^{2l}\, \mathbb{E}\bigl[\xi_k^{2l} \,|\, \mathcal{F}_{k-1}\bigr]}{(2l)!} + \sum_{l=\frac{m}{2}}^{\infty} \frac{t^{2l}\, \mathbb{E}\bigl[\xi_k^{2l} \,|\, \mathcal{F}_{k-1}\bigr]}{(2l)!} \\ & = 1 + \sum_{l=1}^{\frac{m}{2}-1} \frac{(td)^{2l}\, \mathbb{E}\bigl[\bigl(\frac{\xi_k}{d}\bigr)^{2l} \,|\, \mathcal{F}_{k-1}\bigr]}{(2l)!} + \sum_{l=\frac{m}{2}}^{\infty} \frac{(td)^{2l}\, \mathbb{E}\bigl[\bigl(\frac{\xi_k}{d}\bigr)^{2l} \,|\, \mathcal{F}_{k-1}\bigr]}{(2l)!} \\ & \leq 1 + \sum_{l=1}^{\frac{m}{2}-1} \frac{(td)^{2l}\, \gamma_{2l}}{(2l)!} + \sum_{l=\frac{m}{2}}^{\infty} \frac{(td)^{2l}\, \gamma_m}{(2l)!} \\ & = 1 + \sum_{l=1}^{\frac{m}{2}-1} \frac{(td)^{2l}\, \bigl(\gamma_{2l} - \gamma_m\bigr)}{(2l)!} + \gamma_m\bigl(\cosh(td) - 1\bigr) \end{aligned} \tag{36}$$

where the inequality above holds since $|\frac{\xi_k}{d}| \leq 1$ a.s., so that $0 \leq \ldots \leq \gamma_m \leq \ldots \leq \gamma_4 \leq \gamma_2 \leq 1$, and the last equality in (36) holds since $\cosh(x) = \sum_{n=0}^{\infty} \frac{x^{2n}}{(2n)!}$ for every $x \in \mathbb{R}$. Therefore, from (21),

$$\mathbb{E}\biggl[\exp\biggl(t \sum_{k=1}^{n} \xi_k\biggr)\biggr] \leq \left(1 + \sum_{l=1}^{\frac{m}{2}-1} \frac{(td)^{2l}\, \bigl(\gamma_{2l} - \gamma_m\bigr)}{(2l)!} + \gamma_m\bigl[\cosh(td) - 1\bigr]\right)^{n} \tag{37}$$

for an arbitrary $t \in \mathbb{R}$. The inequality then follows from (20). This completes the proof of Theorem 3.

C. A Combined Proof of Theorems 4 and 5

The proof of Theorem 4 relies on the proof of the known result in Theorem 5, where the latter dates back to Freedman's paper (see [8, Theorem 1.6], and also [6, Exercise 2.4.21(b)]). The original proof of Theorem 5 (see [8, Section 3]) is modified in a way that facilitates to realize how the bound can be improved for conditionally symmetric martingales with bounded jumps. This improvement is obtained via the refinement in (29) of Bennett's inequality for conditionally symmetric distributions. Furthermore, the following revisited proof of Theorem 5 simplifies the derivation of the new and improved bound in Theorem 4 for the considered subclass of martingales.

Without any loss of generality, lets assume that $d = 1$ (otherwise, $\{X_k\}$ and z are divided by d, and $\{Q_k\}$ and r are divided by d^2; this normalization extends the bound to the case of an arbitrary $d > 0$). Let $S_n \triangleq X_n - X_0$ for every $n \in \mathbb{N}_0$, then $\{S_n, \mathcal{F}_n\}_{n \in \mathbb{N}_0}$ is a martingale with $S_0 = 0$. The proof starts by introducing two lemmas.

Lemma 2: Under the assumptions of Theorem 5, let

$$U_n \triangleq \exp(\lambda S_n - \theta Q_n), \quad \forall\, n \in \{0, 1, \dots\} \tag{38}$$

where $\lambda \geq 0$ and $\theta \geq e^{\lambda} - \lambda - 1$ are arbitrary constants. Then, $\{U_n, \mathcal{F}_n\}_{n \in \mathbb{N}_0}$ is a supermartingale.

Proof: U_n in (38) is $\mathcal{F}_n$-measurable (since Q_n in (13) is $\mathcal{F}_{n-1}$-measurable, where $\mathcal{F}_{n-1} \subseteq \mathcal{F}_n$, and S_n is $\mathcal{F}_n$-measurable), Q_n and U_n are non-negative random variables, and $S_n = \sum_{k=1}^{n} \xi_k \leq n$ a.s. (since $\xi_k \leq 1$ and $S_0 = 0$). It therefore follows that $0 \leq U_n \leq e^{\lambda n}$ a.s. for $\lambda, \theta \geq 0$, so $U_n \in L^1(\Omega, \mathcal{F}_n, \mathbb{P})$. It is required to show that $\mathbb{E}[U_n | \mathcal{F}_{n-1}] \leq U_{n-1}$ holds a.s. for every $n \in \mathbb{N}$, under the above assumptions on the parameters λ and θ in (38).

$$\begin{aligned}
&\mathbb{E}[U_n|\mathcal{F}_{n-1}] \\
&\stackrel{\text{(a)}}{=} \exp(-\theta Q_n)\, \exp(\lambda S_{n-1})\, \mathbb{E}\big[\exp(\lambda \xi_n) \,|\, \mathcal{F}_{n-1}\big] \\
&\stackrel{\text{(b)}}{=} \exp(\lambda S_{n-1})\, \exp\big(-\theta(Q_{n-1} + \mathbb{E}[\xi_n^2|\mathcal{F}_{n-1}])\big)\, \mathbb{E}\big[\exp(\lambda \xi_n) \,|\, \mathcal{F}_{n-1}\big] \\
&\stackrel{\text{(c)}}{=} U_{n-1} \left(\frac{\mathbb{E}\big[\exp(\lambda \xi_n) \,|\, \mathcal{F}_{n-1}\big]}{\exp(\theta \mathbb{E}[\xi_n^2 \,|\, \mathcal{F}_{n-1}])} \right)
\end{aligned} \tag{39}$$

where (a) follows from (38) and because Q_n and S_{n-1} are $\mathcal{F}_{n-1}$-measurable and $S_n = S_{n-1} + \xi_n$, (b) follows from (13), and (c) follows from (38).

A modification of the original proof of Lemma 2 (see [8, Section 3]) is suggested in the following, which then enables to improve the bound in Theorem 5 for real-valued, discrete-time, conditionally symmetric martingales with bounded jumps. This leads to the improved bound in Theorem 4 for the considered subclass of martingales.

Since by assumption $\xi_n \leq 1$ and $\mathbb{E}[\xi_n \,|\, \mathcal{F}_{n-1}] = 0$ a.s., then applying Bennett's inequality in (22) to the conditional expectation of $e^{\lambda \xi_n}$ given $\mathcal{F}_{n-1}$ (recall that $\lambda \geq 0$) gives

$$\mathbb{E}\big[\exp\big(\lambda \xi_n\big) \,|\, \mathcal{F}_{n-1}\big] \leq \frac{\exp\big(-\lambda \mathbb{E}[\xi_n^2 \,|\, \mathcal{F}_{n-1}]\big) + \mathbb{E}[\xi_n^2 \,|\, \mathcal{F}_{n-1}]\, \exp(\lambda)}{1 + \mathbb{E}\big[\xi_n^2 \,|\, \mathcal{F}_{n-1}\big]}$$

which therefore implies from (39) and the last inequality that

$$\mathbb{E}[U_n|\mathcal{F}_{n-1}] \leq U_{n-1} \left(\frac{\exp\big(-(\lambda+\theta)\, \mathbb{E}[\xi_n^2 \,|\, \mathcal{F}_{n-1}]\big)}{1 + \mathbb{E}\big[\xi_n^2 \,|\, \mathcal{F}_{n-1}\big]} + \frac{\mathbb{E}[\xi_n^2 \,|\, \mathcal{F}_{n-1}]\, \exp\big(\lambda - \theta \mathbb{E}[\xi_n^2 \,|\, \mathcal{F}_{n-1}]\big)}{1 + \mathbb{E}[\xi_n^2 \,|\, \mathcal{F}_{n-1}]} \right). \tag{40}$$

In order to prove that $\mathbb{E}[U_n|\mathcal{F}_{n-1}] \leq U_{n-1}$ a.s., it is sufficient to prove that the second term on the right-hand side of (40) is a.s. less than or equal to 1. To this end, lets find the condition on $\lambda, \theta \geq 0$ such that for every $\alpha \geq 0$

$$\left(\frac{1}{1+\alpha}\right) \exp\big(-\alpha(\lambda+\theta)\big) + \left(\frac{\alpha}{1+\alpha}\right) \exp(\lambda - \alpha\theta) \leq 1 \tag{41}$$

which then assures that the second term on the right-hand side of (40) is less than or equal to 1 a.s. as required.

Lemma 3: If $\lambda \geq 0$ and $\theta \geq \exp(\lambda) - \lambda - 1$ then the condition in (41) is satisfied for every $\alpha \geq 0$.

Proof: This claim follows by calculus, showing that the function

$$g(\alpha) = (1+\alpha) \exp(\alpha\theta) - \alpha \exp(\lambda) - \exp(-\alpha\lambda), \quad \forall\, \alpha \geq 0$$

is non-negative on $\mathbb{R}_+$ if $\lambda \geq 0$ and $\theta \geq \exp(\lambda) - \lambda - 1$. ■

From (40) and Lemma 3, it follows that $\{U_n, \mathcal{F}_n\}_{n\in\mathbb{N}_0}$ is a supermartingale if $\lambda \geq 0$ and $\theta \geq \exp(\lambda) - \lambda - 1$. This completes the proof of Lemma 2. ■

At this point, we start to discuss in parallel the derivation of the tightened bound in Theorem 4 for conditionally symmetric martingales. As before, it is assumed without any loss of generality that $d = 1$.

Lemma 4: Under the additional assumption of the conditional symmetry in Theorem 4, then $\{U_n, \mathcal{F}_n\}_{n\in\mathbb{N}_0}$ in (38) is a supermartingale if $\lambda \geq 0$ and $\theta \geq \cosh(\lambda) - 1$ are arbitrary constants.

Proof: By assumption $\xi_n = S_n - S_{n-1} \leq 1$ a.s., and ξ_n is conditionally symmetric around zero, given $\mathcal{F}_{n-1}$, for every $n \in \mathbb{N}$. By applying Corollary 3 to the conditional expectation of $\exp(\lambda\xi_n)$ given $\mathcal{F}_{n-1}$, for every $\lambda \geq 0$,

$$\mathbb{E}\big[\exp(\lambda\xi_n) \,|\, \mathcal{F}_{n-1}\big] \leq 1 + \mathbb{E}[\xi_n^2 \,|\, \mathcal{F}_{n-1}] \, \big(\cosh(\lambda) - 1\big). \tag{42}$$

Hence, combining (39) and (42) gives

$$\mathbb{E}[U_n|\mathcal{F}_{n-1}] \leq U_{n-1} \left(\frac{1 + \mathbb{E}[\xi_n^2 \,|\, \mathcal{F}_{n-1}] \, \big(\cosh(\lambda) - 1\big)}{\exp\big(\theta \mathbb{E}[\xi_n^2|\mathcal{F}_{n-1}]\big)} \right). \tag{43}$$

Let $\lambda \geq 0$. Since $\mathbb{E}[\xi_n^2 \,|\, \mathcal{F}_{n-1}] \geq 0$ a.s. then in order to ensure that $\{U_n, \mathcal{F}_n\}_{n\in\mathbb{N}_0}$ forms a supermartingale, it is sufficient (based on (43)) that the following condition holds:

$$\frac{1 + \alpha\big(\cosh(\lambda) - 1\big)}{\exp(\theta\alpha)} \leq 1, \quad \forall\, \alpha \geq 0. \tag{44}$$

Calculus shows that, for $\lambda \geq 0$, the condition in (44) is satisfied if and only if

$$\theta \geq \cosh(\lambda) - 1 \triangleq \theta_{\min}(\lambda). \tag{45}$$

From (43), $\{U_n, \mathcal{F}_n\}_{n\in\mathbb{N}_0}$ is a supermartingale if $\lambda \geq 0$ and $\theta \geq \theta_{\min}(\lambda)$. This proves Lemma 4. ■

Hence, due to the assumption of the conditional symmetry of the martingale in Theorem 4, the set of parameters for which $\{U_n, \mathcal{F}_n\}$ is a supermartingale was extended. This follows from a comparison of Lemma 2 and 4 where indeed $\exp(\lambda) - 1 - \lambda \geq \theta_{\min}(\lambda) \geq 0$ for every $\lambda \geq 0$.

Let $z, r > 0$, $\lambda \geq 0$ and either $\theta \geq \cosh(\lambda) - 1$ or $\theta \geq \exp(\lambda) - \lambda - 1$ with or without assuming the conditional symmetry property, respectively (see Lemma 2 and 4). In the following, we rely on Doob's sampling theorem. To this end, let $M \in \mathbb{N}$, and define two stopping times adapted to $\{\mathcal{F}_n\}$. The first stopping time is $\alpha = 0$, and the second stopping time β is the minimal value of $n \in \{0, \ldots, M\}$ (if any) such that $S_n \geq z$ and $Q_n \leq r$ (note that S_n is $\mathcal{F}_n$-measurable and Q_n is $\mathcal{F}_{n-1}$-measurable, so the event $\{\beta \leq n\}$ is $\mathcal{F}_n$-measurable); if such a value of n does not exist, let $\beta \triangleq M$. Hence $\alpha \leq \beta$ are two bounded stopping times. From Lemma 2 or 4, $\{U_n, \mathcal{F}_n\}_{n\in\mathbb{N}_0}$ is a supermartingale for the corresponding set of parameters λ and θ, and from Doob's sampling theorem

$$\mathbb{E}[U_\beta] \leq \mathbb{E}[U_0] = 1 \tag{46}$$

($S_0 = Q_0 = 0$, so from (38), $U_0 = 1$ a.s.). Hence, it implies the following chain of inequalities:

$$\begin{aligned}
&\mathbb{P}(\exists\, n \leq M : \; S_n \geq z, Q_n \leq r) \\
&\stackrel{\text{(a)}}{=} \mathbb{P}(S_\beta \geq z, Q_\beta \leq r) \\
&\stackrel{\text{(b)}}{\leq} \mathbb{P}(\lambda S_\beta - \theta Q_\beta \geq \lambda z - \theta r) \\
&\stackrel{\text{(c)}}{\leq} \frac{\mathbb{E}[\exp(\lambda S_\beta - \theta Q_\beta)]}{\exp(\lambda z - \theta r)} \\
&\stackrel{\text{(d)}}{=} \frac{\mathbb{E}[U_\beta]}{\exp(\lambda z - \theta r)} \\
&\stackrel{\text{(e)}}{\leq} \exp\big(-(\lambda z - \theta r)\big)
\end{aligned} \tag{47}$$

where equality (a) follows from the definition of the stopping time $\beta \in \{0, \ldots, M\}$, (b) holds since $\lambda, \theta \geq 0$, (c) follows from Chernoff's bound, (d) follows from the definition in (38), and finally (e) follows from (46). Since (47) holds for every $M \in \mathbb{N}$, then from the continuity theorem for non-decreasing events and (47)

$$\begin{aligned} &\mathbb{P}(\exists\, n \in \mathbb{N}:\, S_n \geq z, Q_n \leq r) \\ &= \lim_{M \to \infty} \mathbb{P}(\exists\, n \leq M:\, S_n \geq z, Q_n \leq r) \\ &\leq \exp\bigl(-(\lambda z - \theta r)\bigr). \end{aligned} \tag{48}$$

The choice of the non-negative parameter θ as the minimal value for which (48) is valid provides the tightest bound within this form. Hence, without assuming the conditional symmetry property for the martingale $\{X_n, \mathcal{F}_n\}$, let (see Lemma 2) $\theta = \exp(\lambda) - \lambda - 1$. This gives that for every $z, r > 0$,

$$\mathbb{P}(\exists\, n \in \mathbb{N}:\, S_n \geq z, Q_n \leq r) \leq \exp\Bigl(-\bigl[\lambda z - \bigl(\exp(\lambda) - \lambda - 1\bigr) r\bigr]\Bigr), \quad \forall\, \lambda \geq 0.$$

The minimization w.r.t. λ gives that $\lambda = \ln\bigl(1 + \frac{z}{r}\bigr)$, and its substitution in the bound yields that

$$\mathbb{P}(\exists\, n \in \mathbb{N}:\, S_n \geq z, Q_n \leq r) \leq \exp\left(-\frac{z^2}{2r} \cdot B\left(\frac{z}{r}\right)\right) \tag{49}$$

where the function B is introduced in (17).

Furthermore, under the assumption that the martingale $\{X_n, \mathcal{F}_n\}_{n \in \mathbb{N}_0}$ is conditionally symmetric, let $\theta = \theta_{\min}(\lambda)$ (see Lemma 4) for obtaining the tightest bound in (48) for a fixed $\lambda \geq 0$. This gives the inequality

$$\mathbb{P}(\exists\, n \in \mathbb{N}:\, S_n \geq z, Q_n \leq r) \leq \exp\Bigl(-\bigl[\lambda z - r\, \theta_{\min}(\lambda)\bigr]\Bigr), \quad \forall\, \lambda \geq 0.$$

The optimized λ is equal to $\lambda = \sinh^{-1}\Bigl(\frac{z}{r}\Bigr)$. Its substitution in (45) gives that $\theta_{\min}(\lambda) = \sqrt{1 + \frac{z^2}{r^2}} - 1$, and

$$\mathbb{P}(\exists\, n \in \mathbb{N}:\, S_n \geq z, Q_n \leq r) \leq \exp\left(-\frac{z^2}{2r} \cdot C\Bigl(\frac{z}{r}\Bigr)\right) \tag{50}$$

where the function C is introduced in (15).

Finally, the proof of Theorems 4 and 5 is completed by showing that the following equality holds:

$$\begin{aligned} A &\triangleq \{\exists\, n \in \mathbb{N}:\, S_n \geq z, Q_n \leq r\} \\ &= \{\exists\, n \in \mathbb{N}:\, \max_{1 \leq k \leq n} S_k \geq z, Q_n \leq r\} \triangleq B. \end{aligned} \tag{51}$$

Clearly $A \subseteq B$, so one needs to show that $B \subseteq A$. To this end, assume that event B is satisfied. Then, there exists some $n \in \mathbb{N}$ and $k \in \{1, \ldots, n\}$ such that $S_k \geq z$ and $Q_n \leq r$. Since the predictable quadratic variation process $\{Q_n\}_{n \in \mathbb{N}_0}$ in (13) is monotonic non-decreasing, then it implies that $S_k \geq z$ and $Q_k \leq r$; therefore, event A is also satisfied and $B \subseteq A$. The combination of (50) and (51) completes the proof of Theorem 4, and respectively the combination of (49) and (51) completes the proof of Theorem 5.

We prove in the following Corollaries 1 and 2 that extend, respectively, Theorems 1 and 4 to real-valued discrete-time conditionally symmetric sub/ supermartingales.

D. Proof of Corollary 1

The proof of Corollary 1 is similar to the proof of Theorem 1. The only difference is that for a supermartingale, $X_k - X_0 = \sum_{j=1}^k (X_j - X_{j-1}) \leq \sum_{j=1}^k \eta_j$ a.s., where $\eta_j \triangleq X_j - \mathbb{E}[X_j \,|\, \mathcal{F}_{j-1}]$ is $\mathcal{F}_j$-measurable. Hence

$$\mathbb{P}\Bigl(\max_{1 \leq k \leq n} X_k - X_0 \geq \alpha n\Bigr) \leq \mathbb{P}\Biggl(\max_{1 \leq k \leq n} \sum_{j=1}^k \eta_j \geq \alpha n\Biggr)$$

where a.s. $\eta_j \leq d$, $\mathbb{E}[\eta_j \,|\, \mathcal{F}_{j-1}] = 0$, and $\text{Var}(\eta_j \,|\, \mathcal{F}_{j-1}) \leq \sigma^2$. The continuation of the proof coincides with the proof of Theorem 1 for the martingale $\{\sum_{j=1}^k \eta_j, \mathcal{F}_k\}_{k \in \mathbb{N}_0}$ (starting from (20) and (30)). Since $\eta_k \leq d$ a.s., then $X_k - X_0 \leq nd$ a.s. for every $k \in \{1, \ldots, n\}$; hence, if $\delta > 1$ (i.e., $\alpha > d$), then the probability on the left-hand side of (19) is equal to zero. The other inequality for submartingales holds due to the fact that if $\{X_k, \mathcal{F}_k\}$ is a conditionally symmetric submartingale, then $\{-X_k, \mathcal{F}_k\}$ is a conditionally symmetric supermartingale.

E. Proof of Corollary 2

Since $\{X_n, \mathcal{F}_n\}_{n \in \mathbb{N}_0}$ is a supermartingale, then a.s., for every $k \in \mathbb{N}$, $X_k - X_0 \leq \sum_{j=1}^k \eta_j$ where, as before, $\eta_j \triangleq X_j - \mathbb{E}[X_j \,|\, \mathcal{F}_{j-1}]$. Consider the martingale $\{Y_n, \mathcal{F}_n\}_{n \in \mathbb{N}_0}$ where $Y_n \triangleq \sum_{j=1}^n \eta_j$ for every $n \in \mathbb{N}$, and $Y_0 \triangleq 0$ (it is a martingale since a.s. $\mathbb{E}[\eta_j \,|\, \mathcal{F}_{j-1}] = 0$). Since $Y_k - Y_{k-1} = \eta_k$ for every $k \in \mathbb{N}$, then the predictable quadratic variation process $\{Q_n\}_{n \in \mathbb{N}_0}$ which corresponds to the martingale $\{Y_n, \mathcal{F}_n\}_{n \in \mathbb{N}_0}$ is, from (13), the same process as the one which corresponds to the supermartingale $\{X_n, \mathcal{F}_n\}_{n \in \mathbb{N}_0}$. Furthermore, $X_k - X_0 \leq \sum_{j=1}^n \xi_j = Y_k - Y_0$ for every $k \in \mathbb{N}$. Hence, it follows that for every $z, r > 0$,

$$\mathbb{P}\Big(\exists\, n \in \mathbb{N} : \max_{1 \leq k \leq n} (X_k - X_0) \geq z,\, Q_n \leq r\Big) \leq \mathbb{P}\Big(\exists\, n \in \mathbb{N} : \max_{1 \leq k \leq n} (Y_k - Y_0) \geq z,\, Q_n \leq r\Big).$$

Theorem 5, applied to the martingale $\{Y_n, \mathcal{F}_n\}_{n \in \mathbb{N}_0}$, gives the bound in (16) and (17). If $\{X_n, \mathcal{F}_n\}_{n \in \mathbb{N}_0}$ is a conditionally symmetric supermartingale, then $\{Y_n, \mathcal{F}_n\}_{n \in \mathbb{N}_0}$ is a conditionally symmetric martingale. In this case, applying Theorem 4 to the martingale $\{Y_n, \mathcal{F}_n\}_{n \in \mathbb{N}_0}$ gives the improved bound in (14) and (15).

Acknowledgments. I thank Ofer Zeitouni for his help in improving the presentation of the results, and for a suggestion that led to a weakening of assumptions in Lemma 1.

References

[1] N. Alon and J. H. Spencer, *The Probabilistic Method*, Wiley Series in Discrete Mathematics and Optimization, Third Edition, 2008.
[2] K. Azuma, "Weighted sums of certain dependent random variables," *Tohoku Mathematical Journal*, vol. 19, pp. 357–367, 1967.
[3] G. Bennett, "Probability inequalities for the sum of independent random variables," *Journal of the American Statistical Association*, vol. 57, no. 297, pp. 33–45, March 1962.
[4] D. L. Burkholder, "Explorations in martingale theory and its applications," *Ecole d'Eté de Probabilités de Saint-Flour XIX–1989, Lecture Notes in Mathematics*, Springer-Verlag, vol. 1464, pp. 1–66, 1991.
[5] F. Chung and L. Lu, "Concentration inequalities and martingale inequalities: a survey," *Internet Mathematics*, vol. 3, no. 1, pp. 79–127, March 2006.
[6] A. Dembo and O. Zeitouni, *Large Deviations Techniques and Applications*, Springer, second edition, 1997.
[7] K. Dzhaparide and J. H. van Zanten, "On Bernstein-type inequalities for martingales," *Stochastic Processes and their Applications*, vol. 93, no. 1, pp. 109–117, May 2001.
[8] D. Freedman, "On tail probabilities for martingales," *Annals of Probability*, vol. 3, no. 1, pp. 100–118, January 1975.
[9] I. Grama and E. Haeusler, "Large deviations for martingales via Cramer's method," *Stochastic Processes and their Applications*, vol. 85, no. 2, pp. 279–293, February 2000.
[10] W. Hoeffding, "Probability inequalities for sums of bounded random variables," *Journal of the American Statistical Association*, vol. 58, no. 301, pp. 13–30, March 1963.
[11] C. McDiarmid, "On the method of bounded differences," *Surveys in Combinatorics*, vol. 141, pp. 148–188, Cambridge University Press, Cambridge, 1989.
[12] C. McDiarmid, "Concentration," *Probabilistic Methods for Algorithmic Discrete Mathematics*, pp. 195–248, Springer, 1998.
[13] A. Osekowski, "Weak type inequalities for conditionally symmetric martingales," *Statistics and Probability Letters*, vol. 80, no. 23-24, pp. 2009–2013, December 2010.
[14] A. Oskowski, "Sharp ratio inequalities for a conditionally symmetric martingale," *Bulletin of the Polish Academy of Sciences Mathematics*, vol. 58, no. 1, pp. 65–77, 2010.
[15] V. H. de la Pena, "A general class of exponential inequalities for martingales and ratios," *Annals of Probability*, vol. 27, no. 1, pp. 537–564, January 1999.
[16] V. H. de la Pena, M. J. Klass and T. L. Lai, "Self-normalized processes: exponential inequalities, moment bounds and iterated logarithm laws," *Annals of Probability*, vol. 32, no. 3, pp. 1902–1933, July 2004.
[17] I. Pinelis, "Optimum bounds for the distributions of martingales in Banach spaces," *Annals of Probability*, vol. 22, no. 4, pp. 1679–1706, October 1994.
[18] W. L. Steiger, "A best possible Kolmogoroff-type inequality for martingales and a characteristic property," *Annals of Mathematical Statistics*, vol. 40, no. 3, pp. 764–769, June 1969.
[19] G. Wang, "Sharp maximal inequalities for conditionally symmetric martingales and Brownian motion," *Proceedings of the American Mathematical Society*, vol. 112, no. 2, pp. 579–586, June 1991.