Time Averages of Stochastic Processes: a Martingale Approach

Bob Pepin *

October 26, 2018

Abstract

This work shows how exponential concentration inequalities for time averages of stochastic processes over a finite time interval can be obtained from a martingale representation formula. The approach relies on mixing properties of the underlying process, applies to a wide range of initial conditions and makes no assumptions on stationarity or time-homogeneity. A direct method is presented for diffusion processes and discrete-time Markov processes. For general square-integrable processes the constants in the concentration inequalities can be expressed in terms of the quadratic variation of a family of auxiliary martingale. For continuous-time Markov processes they admit a natural expression in terms of the squared field operator applied to the semigroup. The paper concludes with two examples: the squared Ornstein-Uhlenbeck process and the $M/M/\infty$ queue.

1 Introduction

1.1 Problem and Setting

Let $(X_t)_{t \geq 0}$ be a real-valued stochastic process and $S_t$ be an additive functional defined by

$$S_t = \int_0^t X_u \, du.$$ 

A typical example is $X_u = f(u, Y_u)$ for some Markov process $Y$ and $f$ in an appropriate class of functions. The main goal of the present work is to obtain for any $T > 0$ fixed exponential concentration inequalities for $S_T - E S_T$ with explicit constants. Our approach is entirely probabilistic, makes no assumptions about stationary measures and naturally deals with cases such as time-inhomogeneous Markov processes and periodic Markov chains.

*bobpepin@gmail.com. The present work was partially supported by the National Research Fund, Luxembourg
The essential ingredient is a martingale representation of \(S_T - \mathbb{E}S_T\) for fixed \(T\). By controlling the predictable quadratic variation and jumps of the martingale in this representation using a mixing condition on \(X\), we can then deduce concentration inequalities for \(S_T\) from standard concentration inequalities for martingales, yielding for example Hoeffding, Bennett and Bernstein-type inequalities. The martingale representation also opens the door to the use of inequalities on self-normalized martingales to show results of the form

\[
P\left( \frac{|S_T - \mathbb{E}S_T|}{\sqrt{\frac{2}{3}(V_T^2 + \mathbb{E}V_T^2)}} \geq R \right) \leq \min\{2^{1/3}, (2/3)^{2/3}R^{-2/3}\} \exp\left(-\frac{R^2}{2}\right)
\]

for an appropriate normalizing process \(V_T\). The prefactor \(R^{-2/3}\) improves upon known results for additive functionals. We will also see how the self-normalized form leads to a new probabilistic approach to obtaining Bernstein-type concentration inequalities for the classical examples of the squared Ornstein-Uhlenbeck process and the \(M/M/\infty\) queue.

Moreover, for a continuous Markov process \(Y\) with semigroup \(P_{t,u}\) and squared field operator \(\Gamma\), we also obtain the equality, for \(S_T f = \int_0^T f(u, Y_u)\,du\),

\[
\mathbb{E}\exp\left[\lambda (S_T f - \mathbb{E}S_T f) - \lambda^2 \int_0^T \int_t^T \int_t^T \Gamma(P_{t,u} f, P_{t,v} f)(t, Y_t) \,du \,dv \,dt\right] = 1, \quad \lambda \in \mathbb{R}
\]

which seems to be new.

**Organization of the paper.** The paper is structured as follows. In Section 2 we will treat the examples of an elliptic SDE and a discrete-time Markov process using a direct approach. Besides providing concise proofs for those particular cases, it also serves to highlight the main steps of our method and to motivate the objects introduced in the next section. Section 3 first introduces the general continuous-time setting, recalls some martingale inequalities and presents concentration inequalities for \(S_T - \mathbb{E}S_T\) involving the quadratic variations and jumps of an auxiliary martingale. We then proceed to show how to estimate these quantities in a number of concrete cases such as martingales and Markov processes. The Markov case involves in particular the squared field operator \(\Gamma\). Section 4 provides some concrete examples by applying the results from Section 3 to the squared Ornstein-Uhlenbeck process and the \(M/M/\infty\) queue.

**About the literature.** In the Markovian setting, our approach is closely related to the work of Joulin [Jou09], and we recover and extend the results from that work. The main innovation compared to that work is that we replace their tensorization argument by a martingale representation. In the discrete-time setting, the technique used to obtain the martingale representation using a so-called Doob martingale is well-known and commonly used to prove concentration inequalities.

Most previous results on concentration inequalities for functionals of the form \(S_t\) have been obtained for time-homogeneous Markov processes using functional inequalities, see...
for example [WGD04; CG08; Gui+09; GGW14]. The results require the existence of a stationary measure and an initial distribution that has an integrable density with respect to the stationary measure. The same holds true for the combinatorial and perturbation arguments in the classic paper [Lez01]. Some concentration inequalities for homogeneous functionals have previously been established in [Gui01]. A different approach using renewal processes has been used in the work [LL12] to establish concentration inequalities for functionals with bounded integrands.

For Markov processes, the mixing conditions in this work are most naturally formulated in terms of bounds on either the Lipschitz seminorm, gradient or squared field (carré du champs) operator of the semigroup. Bounds on the Lipschitz seminorm are closely related to contractivity in the $L^1$ transportation distance and can for instance be found in [Ebe15; Ass+17] for elliptic diffusions, in [Wu09] for the Riemannian case, in [EGZ17] for Langevin dynamics or in [HMS11] for stochastic delay equations. See also [Oll09; JO10] for the discrete-time case and a large number of examples in both discrete and continuous time. Gradient estimates for semigroups can be obtained using Bismut-type formulas, see for example [EL94; Tha97; CO16], the work [CT18] for the non-autonomous case as well as the textbook [Wan14]. Finally, in terms of the squared field operator, our mixing conditions are a relaxation of the Bakry-Emery curvature-dimension condition [BGL14] since we allow for a prefactor strictly greater than 1.

**Notation.** For a right-continuous process $(X_t)_{t \geq 0}$ with left limits we write $X_{t-} = \lim_{\varepsilon \to 0^+} X_{t-\varepsilon}$ and $\Delta X_t = X_t - X_{t-}$. For a $\sigma$-field $\mathcal{F}$ and a random variable $X$, we denote $\mathbb{E}^\mathcal{F}X$ the conditional expectation of $X$ with respect to $\mathcal{F}$.

## 2 Direct Approach

### 2.1 Diffusion Processes on $\mathbb{R}^n$

First, we consider the case of an elliptic SDE on $\mathbb{R}^n$. Here, the martingale representation follows directly from the Itô formula and the mixing condition is naturally expressed as an assumption on the exponential decay of the gradient of the associated Markov semigroup. We will also see that the martingale representation naturally leads to the appearance of a function that solves a parabolic analogue of the Poisson equation, thereby establishing a link to spectral methods.

Consider a solution $X$ to the following elliptic SDE on $\mathbb{R}^n$:

$$dX_t = b(t, X_t)dt + \sigma(t, X_t)dB_t, \quad X_0 = x_0$$

for $x_0 \in \mathbb{R}^n$, $b: [0, \infty) \times \mathbb{R}^n \to \mathbb{R}^n$ locally bounded, once differentiable in its first argument and twice differentiable in the second with bounded first derivative, $\sigma: [0, \infty) \times \mathbb{R}^n \to \mathbb{R}^{n \times n}$ taking values in the real $n \times n$ matrices such that $\sigma(t, x)\sigma^T(t, x)$ is positive definite.
for every \((t, x)\) and a standard \(n\)-dimensional Brownian motion \(B\). Fix a bounded function \(f\) on \(\mathbb{R}^n\), twice continuously differentiable. For \(t > 0\) let

\[
S_t = \int_0^t f(u, X_u)du.
\]

Denote \(P_{t,u}f\) the two-parameter semigroup

\[
P_{t,u}f(x) = \mathbb{E}[f(u, X_u)|X_t = x].
\]

**Proposition 2.1.** If there exist constants \(\bar{\sigma}, \kappa > 0\) such that

\[
|\sigma^\top(t, x)\nabla P_{t,u}f(x)| \leq \bar{\sigma}e^{-\kappa(u-t)} \quad x \in \mathbb{R}^n, 0 \leq t \leq u
\]

then we have the following Gaussian concentration inequality:

\[
P_{x_0} \left( \frac{1}{T} (S_T - \mathbb{E}_{x_0}S_T) \geq R \right) \leq \exp \left( -\frac{\kappa^2 R^2 T}{2\bar{\sigma}^2} \right) \quad \text{for all } R > 0, T > 0.
\]

**Proof.** Fix \(T > 0\) and define a martingale \(M^T\) by

\[
M^T_t = \mathbb{E}^\mathcal{F}_t S_T
\]

where \(\mathcal{F}_t = \sigma(\{X_s\}_{s \leq t})\) is the natural filtration of \(X\). Using the fact that \(f(u, X_u)\) is \(\mathcal{F}_t\)-measurable for all \(u \leq t\) and the Markov property we get

\[
M^T_t = \mathbb{E}^\mathcal{F}_t S_T = \int_0^t f(u, X_u)du + \int_t^T \mathbb{E}^\mathcal{F}_t f(u, X_u)du = \int_0^t f(u, X_u)du + R^T_t(X_t) \quad (2.1)
\]

with

\[
R^T_t(x) = \int_t^T P_{t,u}f(x)du.
\]

Denoting \(L_t\) the infinitesimal generator of \(X\), we have by the Kolmogorov backward equation \((\partial_t + L_t)P_{t,u}f = 0\) that

\[
(\partial_t + L_t)R^T_t(x) = -P_{t,t}f(x) + \int_t^T (\partial_t + L_t)P_{t,u}f(x)du = -f(t, x), \quad t > 0, x \in \mathbb{R}^n. \quad (2.2)
\]

Using the Itô formula we can now identify \(M^T\) from (2.1) and (2.2):

\[
dM^T_t = d \left( \int_0^t f(u, X_u)du \right) + dR^T_t(X_t)
\]

\[
= f(t, X_t)dt + (\partial_t + L_t)R^T_t f(X_t) dt + \nabla R^T_t f(X_t) \cdot \sigma(t, X_t) dB_t
\]

\[
= \nabla R^T_t f(X_t) \cdot \sigma(t, X_t) dB_t, \quad (2.3)
\]
From our assumption on $P_{t,u}f$ we can estimate for all $x \in \mathbb{R}^n$
\[
|\sigma(t,x)\nabla R_t^T f(x)| \leq \int_t^T |\sigma(t,x)\nabla P_{t,u}f(x)| du \leq \bar{\sigma} \int_t^T e^{-\kappa(t-u)} du \leq \frac{\bar{\sigma}}{\kappa}
\]
so that by (2.3)
\[
d(M^T)_t = |\sigma(t,x)\nabla R_t^T f(x)|^2 dt \leq \frac{\bar{\sigma}^2}{\kappa^2} dt.
\]
It is well known that for any continuous square-integrable martingale $M$ with $M_0 = 0$ and constants $x, y$ the following inequality holds:
\[
P(M_t \geq x, \langle M \rangle_t \leq y) \leq \exp \left( -\frac{x^2}{2y} \right).
\]
Since by the definition of $M^T$, $S_T - \mathbb{E}_{x_0} S_T = M_T^T - M_0^T$ and by our previous result
\[
\langle M^T \rangle_T \leq \frac{\bar{\sigma}^2}{\kappa^2} T
\]
we conclude that
\[
P_{x_0} \left( \frac{1}{T} (S_T - \mathbb{E}_{x_0} S_T) \geq R \right)
\]
\[
\leq \mathbb{P} \left( M_T^T - M_0^T \geq RT, \langle M^T \rangle_T \leq \frac{\bar{\sigma}^2}{\kappa^2} T \right)
\]
\[
\leq \exp \left( -\frac{\kappa^2 R^2 T^2}{2\bar{\sigma}^2} \right).
\]

2.2 Discrete Time Markov Process

For the second example, we consider the case of a discrete-time Markov chain in order to build some probabilistic intuition for our assumptions and to highlight the issues that appear in the presence of jumps.

Consider a discrete-time Markov Process $(X_t)_{t \in \mathbb{N}}$ taking values in $\mathbb{R}$ with $X_0 = x_0 \in \mathbb{R}$. Fix a measurable function $f : \mathbb{N} \times \mathbb{R} \to \mathbb{R}$ and define the associated two-parameter semigroup $P_{t,u}f$ by
\[
P_{t,u}f(x) = \mathbb{E}[f(u,X_u)|X_t = x].
\]
Let
\[
S_t = \sum_{u=1}^t f(u,X_u), \quad t \geq 1.
\]
Proposition 2.2. Assume that the jumps of $X$ are bounded by a constant $a$,
$$|X_t - X_{t-1}| \leq a, \quad t \geq 1,$$
and that there exist positive constants $\sigma$ and $\kappa$ such that
$$P_{t,u}f(x) - P_{t,u}f(y) \leq \sigma(1 - \kappa)^{u-t}|x - y|, \quad 0 \leq t \leq u, x, y \in \mathbb{R}.$$
Then for all $T > 0$ we have the following Gaussian concentration inequality:
$$P_{x_0}\left(\frac{1}{T}(S_T - \mathbb{E}_{x_0}S_T) > R\right) \leq \exp\left(-\frac{\kappa^2 R^2 T}{8a^2 \sigma^2}\right).$$

Proof. Fix $T > 0$ and define a martingale $M^T$ by

$$M^T_t = \mathbb{E}^{F_t}S_T = \sum_{u=1}^{t} f(u, X_u) + \sum_{u=t+1}^{T} \mathbb{E}^{F_t}f(u, X_u).$$

By a direct calculation and the Markov property
$$M^T_t - M^T_{t-1} = \sum_{u=t}^{T} \mathbb{E}^{F_t}f(u, X_u) - \mathbb{E}^{F_{t-1}}f(u, X_u) = \sum_{u=t}^{T} P_{t,u}f(X_t) - P_{t-1,u}f(X_{t-1}).$$

We now use our assumptions on $P_{t,u}f$ and the jumps of $X_t$ to show that the terms in the sum on the right-hand side decay exponentially fast:
$$P_{t,u}f(X_t) - P_{t-1,u}f(X_{t-1}) \leq \sigma(1 - \kappa)^{u-t} \int |X_t - y|P_{t-1,u}(X_{t-1}, dy)$$
$$= \sigma(1 - \kappa)^{u-t} (\mathbb{E}_{F_t}f(X_{t-1}) + \delta - (X_{t-1} + \Delta X_t))|\delta = \Delta X_t$$
$$\leq 2a \sigma (1 - \kappa)^{u-t}.$$

This shows that the increments of the martingale $M^T_t$ are uniformly bounded by a constant independent of $t$ and $T$:
$$M^T_t - M^T_{t-1} \leq 2a \sigma \sum_{u=t}^{T} (1 - \kappa)^{u-t} \leq \frac{2a \sigma}{\kappa}.$$

Since $M^T_T - M^T_0 = S_T - \mathbb{E}^{F_0}S_T$ we get directly from the Azuma-Hoeffding inequality that
$$P_{x_0}\left(\frac{1}{T}(S_T - \mathbb{E}_{x_0}S_T) > R\right) = P_{x_0}(M^T_T - M^T_0 > RT) \leq \exp\left(-\frac{\kappa^2 R^2 T}{8a^2 \sigma^2}\right).$$

$\square$
3 Martingale and concentration inequalities

Consider a filtered probability space \((\Omega, \mathcal{F}, \mathbb{P}, (\mathcal{F}_t)_{t \geq 0})\) satisfying the usual conditions from the general theory of semimartingales, meaning that \(\mathcal{F}\) is \(\mathbb{P}\)-complete, \(\mathcal{F}_0\) contains all \(\mathbb{P}\)-null sets in \(\mathcal{F}\) and \(\mathcal{F}_t\) is right-continuous. In this section \((X_t)_{t \geq 0}\) will denote a real-valued stochastic process adapted to \(\mathcal{F}_t\), bounded in \(L^2(\Omega)\) in the sense that 
\[
\sup_{t \geq 0} \mathbb{E} |X_t|^2 < \infty
\]
and such that \(X_0\) is Lebesgue-measurable.

Define an adapted continuous finite-variation process \((S_t)_{t \geq 0}\) by
\[
S_t = \int_0^t X_u \, du.
\]
Fix \(T > 0\) and define a martingale \((M_T^t)_{t \geq 0}\) by
\[
M_T^t = \mathbb{E} \mathcal{F}_t S_T.
\]

By the boundedness and adaptedness assumptions on \(X\), \(M_T^t\) is a square integrable martingale (by Doob’s maximal inequality) which we can and will choose to be right-continuous with left limits so that the predictable quadratic variation \(\langle M_T^t \rangle\) and the jumps \((\Delta M_T^t)_{t \geq 0}\) are well-defined.

3.1 Martingale inequalities

Our key observation is that \(S_T - \mathbb{E}^{\mathcal{F}_0} S_T = M_T^T - M_0^T\). Concentration inequalities for \(S_T\) then follow from concentration inequalities for martingales. The goal of this section is to recall some well-known martingale inequalities together with conditions allowing us to pass from \(\mathbb{E}^{\mathcal{F}_0} S_T\) to \(\mathbb{E}[S_T]\).

For a real-valued random variable \(Y\), denote \(\Psi_Y(\lambda)\) the logarithm of the moment-generating function of \(Y\) and \(\Psi_Y^*(x)\) its associated Cramér transform:
\[
\Psi_Y(\lambda) = \log \mathbb{E} e^{\lambda Y}, \quad \Psi_Y^*(x) = \sup_{\lambda \geq 0} (\lambda x - \Psi_Y(\lambda)).
\]

Denote \(\Lambda(\lambda)\) the logarithm of the moment-generating function of the centered random variable \(\mathbb{E}^{\mathcal{F}_0} S_T - \mathbb{E} S_T\) and \(I\) its domain:
\[
\Lambda(\lambda) = \Psi_{\mathbb{E}^{\mathcal{F}_0} S_T - \mathbb{E} S_T} = \log \mathbb{E} [\exp \lambda (\mathbb{E}[S_T|X_0] - \mathbb{E}[S_T])],
\]
\[
I = \{\lambda : \Lambda(\lambda) < \infty\}.
\]

In particular if \(X_0 = x \in \mathbb{R}\) is a deterministic constant then \(\mathbb{E}^{\mathcal{F}_0} S_T = \mathbb{E} S_T\) and \(\Lambda(\lambda) = 0\).

Define
\[
\varphi(x) = e^x - 1 - x, \quad \varphi_a(x) = \varphi(ax)/a^2, \quad a \geq 0
\]
and note that we have \( \varphi_0(x) = x^2/2 \).

Let

\[
H^a_t = \sum_{s \leq t} (\Delta M^T_s)^2 \mathbb{1}_{\{|\Delta M^T_s| > a\}} + \langle M^T \rangle_t.
\]

After recalling some martingale inequalities involving \( H^a \), the rest of the paper will be dedicated to estimating the terms \( \Delta M^T \) and \( \langle M^T \rangle \) in the expression of \( H^a \).

**Lemma 3.1.** For \( a \geq 0, \lambda \in I \)

\[
\mathbb{E} \exp \left( \lambda(S_T - \mathbb{E} S_T) - \varphi_a(|\lambda|) H^a_T - \Lambda(\lambda) \right) \leq 1.
\]

**Proof.** In [DZ01] Corollary 3.1 it is shown that for any square integrable martingale \( M \) and for all \( a \geq 0, \lambda \geq 0 \) the process

\[
\exp \left( \lambda M_t - \varphi_a(|\lambda|) H^a_t \right)
\]

is a supermartingale. Applying this to \( M^T \) and \( -M^T \) together with the supermartingale property yields that for \( a \geq 0, \lambda \in \mathbb{R} \)

\[
\mathbb{E}^{\mathcal{F}_0} \exp \left( \lambda(M^T_t - M^T_0) - \varphi_a(|\lambda|) H^a_t \right) \leq 1.
\]

By definition we have furthermore that for \( \lambda \in I \)

\[
\mathbb{E} \exp(\lambda(M^T_0 - \mathbb{E} M^T_0)) = \exp \Lambda(\lambda).
\]

Therefore for \( \lambda \in I \) and all \( t \in [0, T] \)

\[
\mathbb{E} \exp \left( \lambda(M^T_t - \mathbb{E} M^T_t) - \varphi_a(|\lambda|) H^a_t - \Lambda(\lambda) \right)
= \mathbb{E} \left\{ \mathbb{E}^{\mathcal{F}_0} \left[ \exp \left( \lambda(M^T_t - M^T_0) - \varphi_a(|\lambda|) H^a_t \right) \right] \exp \left( \lambda(M^T_0 - \mathbb{E} M^T_0) - \Lambda(\lambda) \right) \right\}
\leq \mathbb{E} \exp \left( \lambda(M^T_0 - \mathbb{E} M^T_0) - \Lambda(\lambda) \right) = 1.
\]

We conclude by taking the inequality at \( t = T \) and noting that \( M^T_T = S_T, \mathbb{E} M^T_T = \mathbb{E} S_T \).

From Markov’s inequality applied to \( e^{\lambda Y} \) we immediately get Chernoff’s inequality

\[
P\{Y \geq x\} \leq \exp(-\Psi^*_Y(x)).
\]

By combining this with Lemma 3.1 and bounds on \( \Lambda \) we can immediately deduce the following Hoeffding, Bennett and Bernstein-type inequalities. The approach is classical and we follow [BLM13].
Corollary 3.2. If \( \Lambda(\lambda) \leq \frac{\lambda^2}{2} \rho^2 \) for some \( \rho \geq 0 \) then
\[
\mathbb{P}( S_T - \mathbb{E} S_T > R, H_T^0 \leq \sigma^2 ) \leq \exp \left( -\frac{R^2}{2(\rho^2 + \sigma^2)} \right).
\]

Proof. On the set \( \{ H_T^0 \leq \sigma^2 \} \), using \( \varphi_0(\lambda) = \frac{\lambda^2}{2} \), \( \Psi_{S_T-\mathbb{E}S_T} \) is upper bounded by the logarithmic MGF of a centered Gaussian random variable with variance \( \rho^2 + \sigma^2 \): \( \Psi_{S_T-\mathbb{E}S_T}(\lambda) \leq \frac{(\rho^2+\sigma^2)\lambda^2}{2} \). This implies that \( \Psi_{S_T-\mathbb{E}S_T}^*(\lambda) \) is lower bounded by the corresponding Cramér transform, \( \Psi_{S_T-\mathbb{E}S_T}^*(x) \geq \frac{\lambda^2}{2(\rho^2 + \sigma^2)} \), and the result follows immediately from Chernoff’s inequality. \( \square \)

Corollary 3.3. If \( \Lambda(\lambda) \leq \nu \varphi_\alpha(\lambda) \) for some \( \alpha, \nu \geq 0 \) then
\[
\mathbb{P}( S_T - \mathbb{E} S_T > R, H_T^0 \leq \mu ) \leq \exp \left( -\frac{\mu + \nu}{\alpha^2} h \left( \frac{aR}{\mu + \nu} \right) \right)
\]
with
\[
h(x) = (1 + x) \log(1 + x) - x, \quad x \geq -1.
\]

Proof. On the set \( \{ H_T^0 \leq \mu \} \), \( \Psi_{S_T-\mathbb{E}S_T}(\cdot/\alpha) \) is upper bounded by the logarithmic MGF of a centered Poisson random variable with parameter \( \frac{\mu + \nu}{\alpha} \): \( \Psi_{S_T-\mathbb{E}S_T}(\lambda/\alpha) \leq \frac{(\mu + \nu)\varphi_\alpha(\lambda)}{\alpha^2} \). This implies that \( \Psi_{S_T-\mathbb{E}S_T}^*(ax) = \sup_{\lambda \geq 0}(\lambda ax - \Psi_{S_T-\mathbb{E}S_T}(\lambda)) = \sup_{\lambda \geq 0}(\lambda x - \Psi_{S_T-\mathbb{E}S_T}(\lambda/\alpha)) \) is lower bounded by the corresponding Cramér transform, \( \Psi_{S_T-\mathbb{E}S_T}^*(ax) \geq \frac{\mu + \nu}{\alpha^2} h \left( \frac{a^2 x}{\mu + \nu} \right) \), and the result follows from Chernoff’s inequality after rescaling by \( \alpha \). \( \square \)

Corollary 3.4. If \( \Lambda(\lambda) \leq \frac{\lambda^2 \nu}{2(1-b\lambda)} \) for some \( b, \nu \geq 0 \) and all \( \lambda < 1/b \) then
\[
\mathbb{P}( S_T - \mathbb{E} S_T > R, H_T^0 \leq \mu ) \leq \exp \left( -\frac{\mu + \nu}{b^2} h_1 \left( \frac{bR}{\mu + \nu} \right) \right)
\]
with
\[
h_1(x) = 1 + x - \sqrt{1 + 2x}.
\]

Proof. On the set \( \{ H_T^0 \leq \mu \} \) using \( \varphi_0(\lambda) = \frac{\lambda^2}{2} \leq \frac{\lambda^2}{2(1-b\lambda)} \), \( \Psi_{S_T-\mathbb{E}S_T} \) is upper bounded by the (rescaled) logarithmic MGF of a sub-Gamma random variable (using the terminology of [BLM13]) with parameter \( (\mu + \nu, b) \): \( \Psi_{S_T-\mathbb{E}S_T}(\lambda) \leq \frac{(\mu + \nu)^2}{2(1-b\lambda)} \). This implies that \( \Psi_{S_T-\mathbb{E}S_T}^*(x) \) is lower bounded by the corresponding Cramér transform, \( \Psi_{S_T-\mathbb{E}S_T}^*(x) \geq \frac{\mu + \nu}{b^2} h_1 \left( \frac{b x}{\mu + \nu} \right) \), and the result follows as before from Chernoff’s inequality. \( \square \)

Going beyond the Chernoff inequality, we have for example the following result which follows directly from Lemma 5.1 and an inequality on self-normalized processes in [PP09] Theorem 2.1.
Corollary 3.5. If \( \Lambda(\lambda) \leq \frac{3}{2} \rho^2 \) for some \( \rho \geq 0 \) and all \( \lambda \in \mathbb{R} \) then

\[
\mathbb{P} \left( \frac{|S_T - \mathbb{E}S_T|}{\sqrt{\frac{3}{2}(H_T^0 + \mathbb{E}H_T^0 + 2\rho^2)}} \geq R \right) \leq (2/3)^{2/3} R^{-2/3} \exp \left( -\frac{R^2}{2} \right).
\]

3.2 General processes bounded in \( L^2(\Omega) \)

Recall that \( X \) is a right-continuous processes bounded in \( L^2(\Omega) \) adapted to a filtration \((\mathcal{F}_t)_{t \geq 0}\). Define a family of auxiliary martingales \((Z^u)_{u \in \mathbb{R}^+}\) by

\[
Z^u_t = \mathbb{E}^\mathcal{F}_t X_u
\]

which will be chosen right-continuous with left limits. Each \( Z^u \) is square integrable so that the predictable quadratic covariation \( \langle Z^u, Z^v \rangle \) is well-defined.

Formally the next result is just a consequence of the (bi)linearity of the integral and (predictable) quadratic variation. The proof shows that the formal calculation is justified under our assumption that \((X_t)_{t \geq 0}\) is bounded in \( L^2(\Omega) \). The main interest of the result lies in the fact that we can often find explicit expressions for \( \langle Z^u, Z^v \rangle \).

Theorem 3.6. For any \( T > 0 \) we have

\[
M^T_T = \int_0^T Z^u_t du, \tag{3.1}
\]

\[
\Delta M^T_T = \int_0^T \Delta Z^u_t du, \tag{3.2}
\]

\[
[M^T]_t = \int_0^T \int_0^T [Z^u, Z^v]_t du dv, \tag{3.3}
\]

\[
\langle M^T \rangle_t = \int_0^T \int_0^T \langle Z^u, Z^v \rangle_t du dv. \tag{3.4}
\]

Proof. We start by showing (3.1). For the sequence of stopping times \( \tau_n = \inf\{t \geq 0 : |X_t| \geq n\} \) we have

\[
\int_0^{T \wedge \tau_n} \mathbb{E}^{\mathcal{F}_t} |X_u| du < nT
\]

so that Fubini’s theorem applies and

\[
M^T_{T \wedge \tau_n} = \mathbb{E}^{\mathcal{F}_t} \int_0^{T \wedge \tau_n} X_u du = \int_0^{T \wedge \tau_n} \mathbb{E}^{\mathcal{F}_t} X_u = \int_0^{T \wedge \tau_n} Z^u_t du.
\]

Now we let \( n \to \infty \) so that \( \tau_n \to \infty \) by the càdlàg property and (3.1) follows by monotone convergence.
Next we show (3.2). Let \( L^1 \) be the space of real-valued \( \mathcal{F}_t \)-adapted stochastic processes equipped with the norm \( \| \xi \|_1 = \mathbb{E} \sup_{t \geq 0} |\xi_t| \). For each \( u \geq 0 \) we have by Doob’s maximal inequality and the observation that \( Z^u_t = X_u \) for \( t \geq u \) that
\[
\| Z^u \|_1^2 \leq \mathbb{E} \sup_{t} |Z^u_t|^2 \leq 4 \lim_{t \to \infty} \mathbb{E}|Z^u_t|^2 = 4 \mathbb{E}|X_u|^2 \leq 4 \sup \mathbb{E}|X_t|^2 < \infty.
\]
This shows that the family of martingales \( \{Z^u\}_{u \geq 0} \) is bounded in \( L^1 \). For \( \xi \in L^1 \) and \( h > 0 \) define \( \Delta_h \xi \in L^1 \) by
\[
(\Delta_h \xi)_t = 1_{\{t \geq h\}}(\xi_t - \xi_{t-h})
\]
and note that \( \| \Delta_h \xi \|_1 \leq 2 \| \xi \|_1 \).

Since \( M^T \) and \( Z^u \) have left limits we have \( \Delta_h M^T \to \Delta M^T \) and \( \Delta_h Z_u \to \Delta Z^u \) a.s. as \( h \to 0^+ \).

By dominated convergence
\[
\Delta_h M^T = \int_0^T \Delta_h Z^u du \to \int_0^T \Delta Z^u du \quad \text{in } L^1
\]
as \( h \to 0^+ \). Since we already saw that the left-hand side converges almost surely to \( \Delta M^T \) this proves (3.2).

Finally we show (3.3) and (3.4). Since the argument is identical in both cases we present it only for (3.3). We are going to use the characterisation of \( [M^T] \) as the unique process such that \( (M^T)^2 \) is a martingale and \( \Delta [M^T] = (\Delta M^T)^2 \). First, by stopping and Fubini as above, we see that for \( 0 \leq s \leq t \)
\[
\mathbb{E}^F_s (M^T) = \int_0^T \int_0^T \mathbb{E}^F_s (Z^u_t Z^v_t) du dv.
\]
Since \( Z^u Z^v - [Z^u, Z^v] \) is a martingale
\[
\mathbb{E}^F_s \left( (M^T)^2 - \int_0^T \int_0^T [Z^u_t Z^v_t]_s du dv \right) = \int_0^T \int_0^T \mathbb{E}^F_s [Z^u_t Z^v_t - [Z^u, Z^v]_t]_s du dv
\]
\[
= (M^T)^2 - \int_0^T \int_0^T [Z^u_t Z^v_t]_s du dv.
\]
We have shown the martingale part of the characterisation. To show the jump part, we will proceed to show that \( [Z^u, Z^v] \) is bounded in \( L^1 \). Since \( [Z^u, Z^v]_t = [Z^u, Z^v]_{t \wedge \wedge \wedge v} \) we have
\[
\|[Z^u, Z^v]\|_1 = \mathbb{E} \sup_t |Z^u_t Z^v_t| \leq \mathbb{E} |Z^u_t Z^v_t|_{t \wedge \wedge \wedge v} \leq \mathbb{E} X^2_t \mathbb{E} X^2_t^{1/2} \leq \sup_t \mathbb{E}|X_t|^2 < \infty.
\]
Since \( [Z^u, Z^v] \) has left limits this shows as above that \( \Delta \) commutes with all the time integrals involved and so
\[
\Delta \int_0^T \int_0^T [Z^u, Z^v]_t du dv = \int_0^T \int_0^T \Delta Z^u Z^v du dv
\]
\[
= \left( \int_0^T \Delta Z^u du \right)^2 = (\Delta M^T)^2.
\]
This concludes the proof of (3.3).

\[\square\]

**Proposition 3.7.** If there exist real-valued processes \(\sigma_t, J_t\) and a constant \(\kappa \geq 0\) such that for all \(0 \leq t \leq u \leq T\)

\[
d\langle Z^u \rangle_t \leq \sigma^2_t e^{-2\kappa(u-t)} dt,
\]

\[|\Delta Z^u_t| \leq |\Delta J_t| e^{-\kappa(u-t)}\]

then

\[
\langle M^T \rangle_t \leq \int_0^t \frac{\sigma^2_s}{\kappa^2} \left(1 - e^{-\kappa(T-s)}\right)^2 ds,
\]

\[|\Delta M^T_t| \leq \frac{|\Delta J_t|}{\kappa} \left(1 - e^{-\kappa(T-t)}\right)\]

where the case \(\kappa = 0\) is to be understood in the sense of the limit as \(\kappa \to 0\).

**Proof.** The second inequality is immediate from Lemma 3.6 and the observation that \(\Delta Z^u_t = 0\) for \(t > u\). We now proceed to prove the first one. For \(t \leq u \wedge v\) we have

\[
\langle Z^u, Z^v \rangle_t \leq e^{-2\kappa u} \int_0^t \sigma^2_s e^{2\kappa s} ds e^{-2\kappa v} \sigma^2_t e^{2\kappa t} = \frac{1}{2} d \left(\int_0^t \sigma^2_s e^{-\kappa(u-s)} e^{-\kappa(v-s)} ds\right)^2
\]

which is symmetric in \(u\) and \(v\). Using Cauchy-Schwarz for the predictable quadratic variation, the fact that \(Z^u_t\) is constant for \(t \geq u\) and integration by parts together with the previous inequality we get

\[
\langle Z_u, Z^v \rangle_t \leq ((Z^u) \langle Z^v \rangle)_{t \wedge u \wedge v} = \left(\int_0^{t \wedge u \wedge v} (Z^u)_s d\langle Z^v \rangle_s + \int_0^{t \wedge u \wedge v} (Z^v)_s d\langle Z^u \rangle_s\right)^{1/2}
\]

\[
\leq \int_0^t \mathbb{1}_{\{s \leq u \wedge v\}} \sigma^2_s e^{-\kappa(u-s)} e^{-\kappa(v-s)} ds.
\]

Therefore by Fubini, for any \(0 \leq t \leq T\)

\[
\langle M^T \rangle_t = \int_0^T \int_0^T \langle Z^u, Z^v \rangle_t du dv
\]

\[
\leq \int_0^T \int_0^T \int_0^t \mathbb{1}_{\{s \leq u \wedge v\}} \sigma^2_s e^{-\kappa(u-s)} e^{-\kappa(v-s)} ds du dv
\]

\[
= \int_0^t \sigma^2_s \left(\int_s^T e^{-\kappa(u-s)} du\right)^2 ds.
\]

which is the result. \(\square\)
3.3 Martingales

Proposition 3.8. If $X$ is a square integrable real-valued martingale then

$$\langle M^T \rangle_T = 2 \int_0^T (T - u)\langle X \rangle_u \, du,$$

$$\Delta M^T_t = \int_0^T \Delta X_u^u \, du = \int_0^t \Delta X_u \, du.$$

Proof. Since $Z^u_t = E^{F_t} X_u = X_{t \wedge u}$ we have by the properties of predictable quadratic variation under stopping

$$\langle Z^u, Z^v \rangle_t = \langle X \rangle_{t \wedge u \wedge v}$$

so that by (3.4)

$$\langle M^T \rangle_T = \int_0^T \int_0^T \langle X \rangle_{T \wedge u \wedge v} \, dv \, du = 2 \int_0^T \int_u^T \langle X \rangle_{T \wedge u \wedge v} \, dv \, du$$

$$= 2 \int_0^T \int_u^T \langle X \rangle_u \, dv \, du = 2 \int_0^T (T - u)\langle X \rangle_u \, du$$

and by (3.2)

$$\Delta M^T_t = \int_0^T \Delta X^u_u \, du = \int_0^t \Delta X_u \, du.$$

\[ \square \]

3.4 Markov Processes

Consider a continuous-time Markov process $(Y_t)_{t \geq 0}$ with natural filtration $(F_t)_{t \geq 0}$, taking values in a Polish space $E$ and with trajectories that are right-continuous with left limits. Denote $\mathcal{B}$ the set of Borel functions on $\mathbb{R}^+ \times E$.

Fix a function $f \in \mathcal{B}$ such that $\sup_t E[f(t, Y_t)^2] < \infty$ so that we are in the setting of Section 3.2 with $X_t = f(t, Y_t)$, $S_T = \int_0^T f(u, X_u) \, du$ and $M^T_t = E^{F_t} S_T$. Define the two-parameter semigroup $(P_{t,u}f)_{0 \leq t \leq u}$ on $E$ by

$$P_{t,u}f(y) = E[f(u, Y_u)|Y_t = y].$$

Suppose that there is a set $\mathcal{D}(\Gamma) \subset \mathcal{B} \times \mathcal{B}$ and a map $\Gamma : \mathcal{D}(\Gamma) \to \mathcal{B}$ such that for each $(f, g) \in \mathcal{D}(\Gamma)$ there is a local martingale $M_t^{fg}$ with

$$f(t, Y_t)g(t, Y_t) - f(0, Y_0)g(0, Y_0) - 2 \int_0^t \Gamma(f, g)(s, Y_s) \, ds = M^{fg}_t, \quad t \geq 0. \quad (3.5)$$
Proposition 3.9. If \((P_{u,f}, P_{v,f})_{0 \leq u, v \leq T} \in \mathcal{D}(\Gamma)\) we have for \(0 \leq t \leq T\)
\[
d\langle M_T \rangle_t = 2 \int_t^T \int_t^T \Gamma(P_{t,u,f}, P_{t,v,f})(t, Y_t) \, du \, dv \, dt, \tag{3.6}
\]
\[
\Delta M^T_t = \int_t^T P_{t,u,f}(Y_t) - P_{t,u,f}(Y_{t-}) \, du. \tag{3.7}
\]

Proof. For all \(0 \leq t \leq u \leq T\) we have
\[
Z^u_t = \mathbb{E}^{F_t} f(u, Y_u) = P_{t,u,f}(Y_t).
\]
Since \(f\) can depend on \(u\) we can always consider \(f(u, y) - P_{0,u,f}(y)\) instead of \(f\) and therefore without loss of generality suppose that \(Z^u_0 = P_{0,u,f}(Y_0) = 0\). By (3.5) for \(0 \leq t \leq u \land v \leq T\)
\[
Z^u_t Z^v_t = P_{t,u,f}(Y_t)P_{t,v,f}(Y_t) = 2 \int_0^t \Gamma(P_{s,u,f}, P_{s,v,f})(s, Y_s) ds + \text{loc. mart.}
\]
and since \(Z^u, Z^v\) are martingales we can identify
\[
d(Z^u, Z^v)_t = 2\Gamma(P_{t,u,f}, P_{t,v,f})(t, Y_t) dt, \quad 0 \leq t \leq u \land v \leq T.
\]
For \(t \geq v \land u\) either \(Z^u_t\) or \(Z^v_t\) is constant so that
\[
d(Z^u, Z^v)_t = 0, \quad u \land v \leq t \leq T,
\]
and thus
\[
\langle Z^u, Z^v \rangle_t = \int_0^t 1_{\{s \leq u \land v\}} 2\Gamma(P_{s,u,f}, P_{s,v,f})(s, Y_s) ds.
\]
By (3.4) and Fubini’s theorem
\[
\langle M^T \rangle_t = \int_0^t \int_0^t \int_0^t 1_{\{s \leq u \land v\}} 2\Gamma(P_{s,u,f}, P_{s,v,f})(s, Y_s) ds du dv
\]
\[
= \int_0^t \int_s^T \int_s^T 2\Gamma(P_{s,u,f}, P_{s,v,f})(s, Y_s) du dv ds.
\]
This proves the first equality (3.6). Equality (3.7) follows directly from (3.2) together with the observation that \(Z^u_t\) is constant for \(t \geq u\) and the fact that \(P_{t,u,f}\) is continuous in \(t\). \qed

Remark 3.10. When \(Y\) is a Markov process with infinitesimal generator \((L, \mathcal{D}(L) \subset \mathcal{B})\) then \(\Gamma\) in (3.6) corresponds to the usual squared field operator whenever the latter is well-defined,
\[
\Gamma(f, g) = \frac{1}{2}(Lf g - f Lg - g Lf), \quad f \in \mathcal{D}(L), g \in \mathcal{D}(L), f g \in \mathcal{D}(L).
\]
Indeed, suppose that for \( f \in \mathcal{D}(L) \)
\[
f(t, Y_t) - f(0, Y_0) - \int_0^t (\partial_s f + Lf(s, \cdot))(s, Y_s) \, ds
\]  
(3.8)
is a local martingale. As before we can assume \( P_{0,u}f(Y_0) = 0 \). Now if \( P_{t,u}f, P_{t,v}f \) and their product \( P_{t,u}fP_{t,v}f \) are in \( \mathcal{D}(L) \) then
\[
P_{t,u}f(Y_t)P_{t,v}f(Y_t) - \int_0^t \left( \partial_s (P_{s,u}fP_{s,v}f) + L(P_{s,u}fP_{s,v}f) \right)(s, Y_s) \, ds
\]  
(3.9)
is a local martingale. Since we assumed \( P_{t,u}f \) to be in \( \mathcal{D}(L) \) it solves the Kolmogorov backward equation
\[
\partial_t P_{t,u}f(y) = -LP_{t,u}(t, y), \quad y \in \mathbb{E}, \quad 0 \leq t \leq u
\]and the same holds true for \( P_{t,v}f \). Thus
\[
\partial_t (P_{t,u}fP_{t,v}f) = P_{t,u}f\partial_t P_{t,v}f + P_{t,v}f\partial_t P_{t,u}f = -P_{t,u}fLP_{t,v}f - P_{t,v}fLP_{t,u}f.
\]Substituting this into the integral in (3.9) shows that indeed
\[
P_{t,u}f(Y_t)P_{t,v}f(Y_t) - 2 \int_0^t \frac{1}{2} \left( L(P_{s,u}fP_{s,v}f) - P_{s,u}fLP_{s,v}f - P_{s,v}fLP_{s,u}f \right)(s, Y_s) \, ds
\]is a local martingale.

**Corollary 3.11.** If \( Y \) has continuous trajectories and \( Y_0 \) is constant then we have the following equality for all \( \lambda \in \mathbb{R}, T > 0 \):
\[
\mathbb{E} \exp \left[ \lambda (S_T f - \mathbb{E}S_T f) - \lambda^2 \int_0^T \int_T^T \Gamma(P_{t,u}f, P_{t,v}f)(t, Y_t) \, du \, dv \, dt \right] = 1.
\]

**Proof.** Since \( Y_0 \) is constant we have
\[
S_T f - \mathbb{E}S_T = S_T f - \mathbb{E}F_0 S_T = M_T^T - M_0^T
\]and the result follows from the fact that the Doléans-Dade exponential
\[
\exp \left[ \lambda^2 (M_t^T - M_0^T) - \frac{\lambda^2}{2} (M_t^T)^2 \right]
\]is a local martingale. \( \square \)

**Remark 3.12.** If we set \( R_t^2(x) = \int_t^T P_{t,u}f(x) \, du \) we can rewrite (3.6) as
\[
d(M_t^T)^2 = 2\Gamma(R_t^2)(Y_t)
\]
using the bilinearity of \( \Gamma \) (assuming everything is sufficiently regular). In particular \( R_T^T \) solves the PDE
\[
-(\partial_t + L)R_T^T = f.
\]
Whereas the use of \( R_T^T \) in this context seems to be new, many well-known results on concentration inequalities involve an expression of the form \( \Gamma(R) \) where the resolvent \( R(x) = \int_0^\infty P_0,uf(x)\,du \) solves the Poisson equation
\[
-\mathcal{L}R = f.
\]
See for example Section 5 in [GGW14]. The approaches based on the Poisson equation require the semigroup to be time-homogeneous and \( f \) to be centered with respect to a stationary measure of the process, whereas our method has no need for stationarity and only imposes mixing conditions up to time \( T \).

4 Examples

4.1 Squared Ornstein-Uhlenbeck Process

Let \( X \) be the Ornstein-Uhlenbeck process, solution to the following SDE on \( \mathbb{R} \):
\[
dX_t = -\theta X_t\,dt + dB_t
\]
for \( \theta > 0 \) and \( B \) a standard Brownian motion on \( \mathbb{R} \). Let \( f(x) = x^2 \) and
\[
S_T = \int_0^T f(X_t)\,dt = \int_0^T X_t^2\,dt.
\]
Let \( \Lambda(\lambda) = \log \mathbb{E}\exp(\lambda (\mathbb{E}X_0 - \mathbb{E}X_T)) \). Using the method of self-normalized martingales, we find the following Bernstein-type concentration inequality.

**Proposition 4.1.** If \( \Lambda(\lambda) \leq \rho^2 \lambda^2/2 \) for some \( \rho > 0 \) then
\[
\mathbb{P}
\left(
\frac{1}{T} (\mathbb{E}X_T) > R \right) \leq I(R,T)^{-1/3}e^{-(3/2)I(R,T)}
\]
with
\[
I(R, T) = \frac{\theta^2 R^2 T}{2 \left( R + \frac{1}{\theta} + \frac{\mathbb{E}X_0 - \frac{1}{2\theta}}{\theta} + \frac{2\rho^2 \rho^2}{R} \right)}.
\]

**Remark 4.2.** If \( X_0 = x \in \mathbb{R} \) we can always take \( \rho = 0 \). If we consider the stationary case \( X_0 \sim \mu = \mathcal{N}(0, \frac{1}{2\theta}) \) then we have
\[
\Lambda(\lambda) = \log \int_{\mathbb{R}} \exp \lambda \left( \int_0^T \mathbb{E}X_t^2 - \frac{1}{2\theta} dt \right) \mu(dx) = \log \int_{\mathbb{R}} \exp \left( \frac{\lambda (\mu^2 - \frac{1}{2\theta} \mathbb{E}X_T^2)}{2\theta} \right) \mu(dx).
\]
which is finite only for \( \lambda < \lambda_0 := \frac{2}{\theta + e^{-2\theta}} \). This case would require a generalization of the inequality for self-normalized martingales used in Corollary 3.5 to the case where \( \Lambda(\lambda) < \infty \) only for \( \lambda \leq \lambda_0 \). See [LKP04] for some work in that direction.

**Remark 4.3.** According to results obtained using spectral methods and large deviation estimates (see [Lez01; GGW14] and references therein), the best constants in the exponential on the right-hand side of (4.1) would be achieved by replacing \( (3/2) I(R, T) \) by

\[
J(R, T) = \frac{\theta^2 R^2 T}{2 \left( R + \frac{1}{2g} + \frac{\theta X_t^2}{g T} + \frac{2g^2T^2}{T} \right)}.
\]

The factor \( 2/3 \) and the extra term \( \frac{1}{2g} \) in the denominator of \( (3/2) I \) can be traced back to the corresponding factor \( 2/3 \) and term \( \mathbb{E} \langle M \rangle_T \) in the martingale inequality used to obtain Corollary 3.5.

**Proof.** Since the SDE for \( X \) is linear we can find an explicit solution which reads

\[
X_t = X_0 e^{-\theta t} + \int_0^t e^{-\theta(t-s)} dB_s.
\]

For each \( t \), \( X_t \) is normally distributed with mean \( \mathbb{E} X_0 e^{-\theta t} \) and variance \( 1 - e^{-2\theta t} \). In particular \( \sup_t \mathbb{E} X_t^4 < \infty \) so that the results from Section 3 apply to \( X_t^2 \). We also get an explicit expression for the semigroup

\[
P_t f(x) = \mathbb{E} [ f(X_t) | X_0 = x ] = \mathbb{E} f \left( x e^{-\theta t} + \int_0^t e^{-\theta(t-s)} dB_s \right)
\]

and by differentiating under the expectation, still with \( f(x) = x^2 \),

\[
\partial_x P_t f(x) = \mathbb{E} f' \left( x e^{-\theta t} + \int_0^t e^{-\theta(t-s)} dB_s \right) e^{-\theta t} = 2xe^{-2\theta t}.
\]

The squared field operator is

\[
\Gamma(f, g) = \frac{1}{2} \partial_x f \partial_x g
\]

and the semi-group is time-homogeneous:

\[
P_{t,u} f = P_{u-t} f.
\]

If we let

\[
M_T^T = \mathbb{E}^S t S_T
\]

for the natural filtration \( (\mathcal{F}_t)_{t \geq 0} \) associated with \( X \) then we have by Proposition 3.9

\[
d(M^T)_t = \int_t^T \int_t^T 2 \Gamma(P_{t,u} f, P_{t,v} f)(X_t) dudv
\]

\[
= \left( \int_0^{T-t} \nabla P_u f(X_t) du \right)^2
\]

\[
= 4X_t^2 \left( \int_0^{T-t} e^{-2\theta u} du \right)^2
\]

\[
= X_t^2 \left( 1 - e^{-2\theta(T-t)} \right)^2.
\]
The key observation is that

$$\langle M^T \rangle_T = \frac{1}{\theta^2} \int_0^T X_t^2 \left(1 - e^{-2\theta(T-t)} \right)^2 \, dt \leq \frac{1}{\theta^2} S_T.$$  \hspace{1cm} (4.1)

By monotonicity of $x \mapsto \frac{x}{x+1}$, inequality (4.1) and Corollary 3.5

$$P(S_T - ES_T \geq RT) \leq P \left( \frac{S_T - ES_T}{\sqrt{\frac{2}{3} \left(\frac{S_T}{\theta^2} + ES_T + 2\rho^2 \right)}} \geq \phi(R, T) \right)$$

$$\leq P \left( \frac{S_T - ES_T}{\sqrt{\frac{2}{3} (\langle M^T \rangle_T + ES_T + 2\rho^2)}} \geq \phi(R, T) \right)$$

$$\leq (2/3)^{2/3} \phi(R, T)^{-2/3} e^{-\phi(R, T)^2/2}$$

with

$$\phi(R, T) = \frac{RT}{\sqrt{\frac{2}{3} \left( \frac{RT + ES_T}{\theta^2} + \langle M^T \rangle_T + 2\rho^2 \right)}}.$$ 

Since $EX_T^2 = (EX_0^2 - \frac{1}{2\theta})e^{-2\theta t} + \frac{1}{2\theta}$ and by (4.1) we have

$$E\langle M^T \rangle_T \leq \frac{1}{\theta^2} ES_T \leq \frac{1}{\theta^2} \left( \frac{EX_0^2}{\theta} - \frac{1}{2\theta} + T \right),$$

so that

$$\phi(R, T) \geq \frac{\theta R \sqrt{T}}{\sqrt{\frac{2}{3} \left( R + \frac{1}{\theta} + \frac{EX_0^2 - \frac{1}{2\theta}}{\theta^2} + \frac{2\rho^2 + \rho^2}{3/4} \right)}}$$

which is the result.

\[ \square \]

4.2 \textit{M/M/}\(\infty\) queue

The $M/M/\infty$ queue with parameter $\lambda$ is a space-inhomogeneous birth and death Markov process on $\mathbb{N}$ with infinitesimal generator given for any function $f : \mathbb{N} \to \mathbb{R}$ and $x \in \mathbb{N}$ by

$$Lf(x) = \lambda(f(x+1) - f(x)) + x(f(x-1) - f(x)).$$

Denote $X$ a realization of the $M/M/\infty$ queue and for simplicity of exposition suppose that $X_0 = x_0 \in \mathbb{N}$. The situation for other initial measures is similar to that in the previous section.
**Proposition 4.4.** For any functions $f : \mathbb{N} \to \mathbb{R}$ such that $\sup_{x \geq 1} |f(x) - f(x - 1)| \leq 1$ we have for $S_T = \int_0^T f(X_u) du$ that

$$
\mathbb{P}\left( \frac{|S_T - \mathbb{E}S_T|}{\sqrt{\frac{2}{T} \sum_{t \leq T} |\Delta X_t| + 2\lambda T + x_0 - \lambda}} > R \right) \leq \left(\frac{2}{3}\right)^{2/3} R^{-2/3} \exp\left(-\frac{R^2}{2}\right).
$$

**Proof.** Denote $P_{t,u} = P_{u-t}$ the semigroup associated to $L$. We have from [Oll09] that for all 1-Lipschitz functions $f$ and $x, y \in \mathbb{N}$

$$
|P_t f(x) - P_t f(y)| \leq e^{-t}|x - y|.
$$

As before let $S_T = \int_0^T f(X_s) ds$ and $M_T^T = \mathbb{E}F_t S_T$ for the natural filtration $\mathcal{F}_t$ associated to $X$. By Theorem 3.6

$$
\Delta M_T^T = \int_0^{T-t} P_u f(X_t) - P_u f(X_t^-) du \leq |\Delta X_t| \int_0^{T-t} e^{-u} du \leq |\Delta X_t|
$$

so that, noting that $|\Delta X_t|$ is either 0 or 1,

$$
[M^T]_T = \sum_{t \leq T} |\Delta M_T^T|^2 \leq \sum_{t \leq T} |\Delta X_t|.
$$

In particular, there exists a Poisson process $N$ (see for example Theorem 4.1 in [EK09]) such that

$$
\sum_{t \leq T} |\Delta X_t| = N\left( \int_0^T X_t + \lambda dt \right)
$$

so that

$$
\mathbb{E}[M^T]_T \leq \int_0^T \mathbb{E}X_t + \lambda dt = \int_0^T 2\lambda + (\mathbb{E}X_0 - \lambda) e^{-t} dt \leq 2\lambda T + \mathbb{E}X_0 - \lambda
$$

where we used $EX_t - \lambda = (EX_0 - \lambda) e^{-t}$ which follows from the fact that $x \mapsto x - \lambda$ is an eigenfunction of $-L$ with eigenvalue 1.

The result now follows directly by applying Corollary 3.5 with $H^0 = [M^T]$ and $\rho = 0$ (since we assumed $X_0 \sim \delta_{x_0}$).

**References**

[Ass+17] Roland Assaraf, Benjamin Jourdain, Tony Lelièvre, and Raphaël Roux. “Computation of sensitivities for the invariant measure of a parameter dependent diffusion”. In: *Stochastics and Partial Differential Equations: Analysis and Computations* (Oct. 16, 2017).
[JO10] Aldéric Joulin and Yann Ollivier. “Curvature, concentration and error estimates for Markov chain Monte Carlo”. In: The Annals of Probability 38.6 (Nov. 2010), pp. 2418–2442.

[Jou09] Aldéric Joulin. “A new Poisson-type deviation inequality for Markov jump processes with positive Wasserstein curvature”. In: Bernoulli 15.2 (May 2009), pp. 532–549.

[Lez01] Pascal Lezaud. “Chernoff and Berry–Esséen inequalities for Markov processes”. In: ESAIM: Probability and Statistics 5 (2001), pp. 183–201.

[LKP04] Tze Leung Lai, Michael J. Klass, and Victor H. de la Peña. “Self-normalized processes: Exponential inequalities, moment bounds and iterated logarithm laws”. In: The Annals of Probability 32.3 (July 2004), pp. 1902–1933.

[LL12] Eva Löcherbach and Dasha Loukianova. “Deviation Inequalities for Centered Additive Functionals of Recurrent Harris Processes Having General State Space”. In: Journal of Theoretical Probability 25.1 (Mar. 2012), pp. 231–261.

[Oll09] Yann Ollivier. “Ricci curvature of Markov chains on metric spaces”. In: Journal of Functional Analysis 256.3 (Feb. 2009), pp. 810–864.

[PP09] Victor de la Peña and Guodong Pang. “Exponential inequalities for self-normalized processes with applications”. In: Electronic Communications in Probability 14.0 (2009), pp. 372–381.

[Tha97] Anton Thalmaier. “On the differentiation of heat semigroups and Poisson integrals”. In: Stochastics: An International Journal of Probability and Stochastic Processes 61.3 (1997), pp. 297–321.

[Wan14] Feng-Yu Wang. Analysis for diffusion processes on Riemannian manifolds. World Scientific Pub. Co, 2014. 379 pp.

[WGD04] Liming Wu, Arnaud Guillin, and Hacene Djellout. “Transportation cost-information inequalities and applications to random dynamical systems and diffusions”. In: The Annals of Probability 32.3 (July 2004), pp. 2702–2732.

[Wu09] Liming Wu. “Gradient estimates of Poisson equations on Riemannian manifolds and applications”. In: Journal of Functional Analysis 257.12 (Dec. 2009), pp. 4015–4033.