A HOEFFDING’S INEQUALITY FOR UNIFORMLY ERGODIC DIFFUSION PROCESS

MICHAEL C.H. CHOI, EVELYN LI

ABSTRACT. In this note, we present a version of Hoeffding’s inequality in a continuous-time setting, where the data stream comes from a uniformly ergodic diffusion process. Similar to the well-studied case of Hoeffding’s inequality for discrete-time uniformly ergodic Markov chain, the proof relies on techniques ranging from martingale theory to classical Hoeffding’s lemma as well as the notion of deviation kernel of diffusion process. We present two examples to illustrate our results. In the first example we consider large deviation probability on the occupation time of the Jacobi diffusion, a popular process used in modelling of exchange rates in mathematical finance, while in the second example we look at the exponential functional of a finite interval analogue of the Ornstein-Uhlenbeck process introduced by Kessler and Sørensen (1999).

AMS 2010 subject classifications: 60F10, 60J10, 60G44
Keywords: Diffusion process; Hoeffding’s inequality; large deviations

1. INTRODUCTION AND MAIN RESULTS

The seminal work of Hoeffding (1963), which gives bound on large deviation probability of sum of bounded random variables, has now became one among many classical tools in probability theory. In particular, it has far reaching applications in statistics and machine learning, see for instance Devroye et al. (1996) and the references therein. Hoeffding’s inequality has then been refined and extended to various settings. For example, motivated by applications in Markov decision processes and reinforcement learning Glynn and Ormoneit (2002) derives a Hoeffding’s inequality for uniformly ergodic Markov chain, while Boucher (2009) presents another method to prove Hoeffding’s inequality in terms of the Drazin inverse of Markov chain.

Inspired by the work cited above, we aim at extending Hoeffding’s inequality to the setting of diffusion process. Contrary to the classical setting, we assume that the data stream arrives continuously from a uniformly ergodic diffusion process. The major difficulty of the analysis is then twofold. First, as we have a continuous data stream instead of discrete data points, previous analysis does not carry over to this setting easily. In addition, the dependency within the data stream complicates the situation. To overcome these difficulties, we employ classical martingale techniques for diffusion process as well as the notion of deviation kernel to aid our analysis. Comparing our result with the existing literature on concentration inequalities for diffusion processes Galtchouk and Pergamenshchikov (2007), we argue that our proof is conceptually simpler since it utilizes similar techniques as in the discrete-time Markov chain case Boucher (2009); Glynn and Ormoneit (2002). In addition, as we shall see in Corollary 1.1 below, it is readily applicable as long as we have the relevant eigenvalue information of the generator of the diffusion.

To this end, we fix our notation and introduce the tools we need for our main result Theorem 1.1 below. Let \((\Omega, \mathcal{F}, (\mathcal{F}_t)_{t \geq 0}, \mathbb{P})\) be a filtered probability space satisfying the usual conditions. Suppose that we have an ergodic diffusion process \(X = (X_t)_{t \geq 0}\) on state space \(S\) with transition kernel \(P\), transition density \(P(x, dy)\) and stationary distribution \(\pi\). We write \(\mathbb{P}_x(\cdot) := \mathbb{P}(\cdot | X_0 = x)\) and
$\mathbb{E}_x (\cdot) := \mathbb{E}[\cdot \mid X_0 = x]$ to be the conditional probability and expectation when the process is initialized at $X_0 = x \in S$. $X$ is characterized by the infinitesimal generator $A$, which acts on the space of twice differentiable functions and is defined to be

$$A := \mu(x) \frac{d}{dx} + \frac{1}{2} \sigma^2(x) \frac{d^2}{dx^2},$$

where $\mu(x)$ and $\sigma^2(x)$ are respectively known as the drift and diffusion coefficient of $X$. A tool that we will use in the main result below is the deviation kernel $Q^2$ of $X$, which is defined as

$$Q^2 := \int_0^\infty (P^t - \Pi) \, dt,$$

where $\Pi$ is the projection kernel with density $\Pi(x, dy) = \pi(dy)$ for all $x \in S$. It is well-known that the function $\hat{f} := Q^2 f$ solves the Poisson equation $-A \hat{f} = f$, see e.g. Glynn and Meyn (1996). For bounded function $f$, we define the supremum norm to be $\|f\| := \sup_x |f(x)|$. We also write $\|Q^2\| := \sup_{\|f\| \leq 1} \|Q^2 f\|$ to be the induced operator norm of $Q^2$ on the space of bounded functions. For further references on $Q^2$, we refer readers to the work of Cheng and Mao (2015); Mao (2002); Whitt (1992).

On one-dimensional state space $S = (l, u)$, we now recall two fundamental notions associated with the diffusion $X$, namely the scale function $S(x)$ and the speed function $M(x)$. For $x \in X$, these functions are defined by

$$S(x) := \int_{x_0}^x \exp \left\{ - \int_{x_0}^y \frac{2\mu(z)}{\sigma^2(z)} dz \right\} dy, \quad M(x) := \int_l^x \frac{2}{\sigma^2(y)} \exp \left\{ \int_{x_0}^y \frac{2\mu(z)}{\sigma^2(z)} dz \right\} dy,$$

where $x_0 \in S$ is a fixed and arbitrary reference point. Their respective densities are given by

$$s(x) := \frac{d}{dx} S(x) = \exp \left\{ - \int_{x_0}^x \frac{2\mu(z)}{\sigma^2(z)} dz \right\}, \quad m(x) := \frac{d}{dx} M(x) = \frac{2}{\sigma^2(x) s(x)}.$$

In this note, we are primarily interested in uniformly ergodic diffusions. That is, it is the class of ergodic diffusions such that the convergence to equilibrium in total variation distance is uniformly bounded by, for $t \geq 0$ and some constants $C < \infty$, $\beta > 0$,

$$\sup_{x \in S} \|P^t(x, \cdot) - \pi\|_{TV} \leq C e^{-\beta t},$$

where $\|P^t(x, \cdot) - \pi\|_{TV} := \sup_A |P^t(x, A) - \pi(A)|$ is the total variation distance between $P^t(x, \cdot)$ and $\pi$. We write $\tau_y := \inf\{t \geq 0; X_t = y\}$ to be the first hitting time of $y$ and

$$t_{av} := \int_S \mathbb{E}_x [\tau_y] \pi(dx) \pi(dy)$$

to be the average hitting time of $X$. While verifying uniform ergodicity can be quite difficult, it turns out that, according to (Cheng and Mao, 2015, Theorem 2.2), uniform ergodicity for diffusion on $(0, u)$ with reflecting boundary at 0 is equivalent to a few readily checkable conditions on $t_{av}$, $s(x)$ and $m(x)$:

**Proposition 1.1** ( Necessary and sufficient conditions for uniform ergodicity Cheng and Mao (2015)). Given an ergodic diffusion $X$ on $S = (0, u)$ with reflecting boundary at 0, the following statements are equivalent:

1. $X$ is uniformly ergodic;
2. $\int_S m([x, u]) s(x) \, dx < \infty$;
3. $t_{av} < \infty$;
(4) $\sigma_{\text{ess}}(A) = \emptyset$ and

$$\sum_{i \geq 1} \frac{1}{\lambda_i} < \infty,$$

where $\sigma_{\text{ess}}(A)$ is the essential spectrum of $A$ and $(\lambda_i)_{i \geq 1}$ are the non-zero eigenvalues of $-A$.

At times it may be easier to check item (2) as it depends on $\mu(x)$ and $\sigma^2(x)$ through $s(x)$ and $m(x)$, while at other times when eigenvalues information are available perhaps it is more convenient to check item (4). As a simple illustration of item (2), we consider the class of diffusions with $\mu(x) = 0, \sigma^2(x) = 2(1 + x)^{\gamma}$ and $S = (0, \infty)$, where $\gamma > 2$ is a parameter. This class is first studied in Mao (2002). It is easy to see that $s(x) = 1$ and $m(x) = (1 + x)^{-\gamma}$. As a result, item (2) now reads

$$\int_S m([x, \infty]) s(x) \, dx = \frac{1}{(\gamma - 1)(\gamma - 2)} < \infty,$$

and so this class of diffusions with $\gamma > 2$ are uniformly ergodic. For illustration of item (4), we defer the readers to Corollary 1.1 when we discuss the Jacobi process. In view of Proposition 1.1, for uniformly ergodic $X$ we have

$$||Q^2|| \leq 2t_{av} < \infty,$$

where the first inequality follows from (Choi, 2018, Theorem 1.1). In other words, for uniformly ergodic diffusion the induced operator norm $||Q^2||$ is finite. Such a term will appear in our version of Hoeffding’s inequality Theorem 1.1 below.

With the above notation, we are now ready to state our main result. It follows from the classical ergodic theorem that for bounded function $f$, the time average $\frac{1}{t} \int_0^t f(X_s) \, ds$ converges almost surely to the space average $\pi(f) := \int_S f(x) \pi(dx)$ as $t \to \infty$, see e.g. (Bhattacharya and Waymire, 2009, Theorem 12.2). In our main result below, we present non-asymptotic probabilistic error bound of such convergence:

**Theorem 1.1.** Suppose that $X = (X_t)_{t \geq 0}$ is a uniformly ergodic diffusion and $f$ is a bounded function.

For any $\varepsilon > 0$, $t > \frac{2\|f\|\|Q^2\|}{\varepsilon}$, $x \in S$,

$$\mathbb{P}_x \left( \frac{1}{t} \int_0^t f(X_s) \, ds - \pi(f) \geq \varepsilon \right) \leq \exp \left\{ \frac{-2 (t\varepsilon - 2 \|f\| \|Q^2\|)^2}{(t + 1) \|f\|^2 (2 \|Q^2\| + 1)^2} \right\},$$

where $Q^2$ is the deviation kernel of the process $X$.

**Remark 1.1** (On the assumption of bounded $f$). As usual in the Hoeffding’s inequality literature, our main result Theorem 1.1 requires the function $f$ to be bounded. This assumption is crucial when we apply the classical Hoeffding’s lemma (Devroye et al., 1996, Lemma 8.1) to certain martingale difference sequence in (2.6) and (2.7) below, which only holds when the random variable of interest is bounded. Although there is an extension of the Hoeffding’s lemma to non-negative random variable with finite mean Bentkus (2008), this result is however difficult to apply in our setting as one need to find random variables that stochastically dominate the martingale difference sequence. We leave this question of extending the main result to unbounded $f$ as future work.

As our first example to illustrate our main result Theorem 1.1, we investigate the Jacobi process $X = (X_t)_{t \geq 0}$ on the state space $S = (0, 1)$. The generator of this process is given by

$$A^f = (a - bx) \frac{d}{dx} + \frac{\sigma^2}{2} x (1 - x) \frac{d^2}{dx^2},$$
where \(a, b, \sigma \in \mathbb{R}\) are parameters of \(X\) and are assumed to take on values such that \(\alpha := \frac{2b}{\sigma^2} - 2\sigma^2 - 1 > -1\) and \(\beta := \frac{2a}{\sigma^2} - 1 > -1\), i.e. \(b > a > 0\) and \(\sigma \in \mathbb{R}\). With these choices of parameters, the stationary distribution of Jacobi process is the Beta distribution with parameters \(\alpha + 1\) and \(\beta + 1\), where its density is governed by \(\pi(x) = \frac{x^{\alpha}(1-x)^{\beta}}{B(\alpha+1,\beta+1)}\), and \(B(\cdot, \cdot)\) denotes the Beta function. According to (Albanese and Kuznetsov, 2009, Appendix B.3) and (Forman and Sørensen, 2008, Section 2.1) \(X\) is ergodic with \(\sigma_{\text{ess}}(A) = 0\) and
\[
\lambda_i = \frac{\sigma^2}{2} i \left( i - 1 + \frac{2b}{\sigma^2} \right), \quad \sum_{i=1}^{\infty} \frac{1}{\lambda_i} = \frac{2}{\sigma^2} \sum_{i=1}^{\infty} \frac{1}{i \left( i - 1 + \frac{2b}{\sigma^2} \right)} < \infty.
\]
In view of Proposition 1.1 item (4), \(X\) is thus uniformly ergodic. One major motivation for us to study such a process stems from its usage in financial modelling, where a more general form of Jacobi process has been employed to model exchange rates in a target zone, see Larsen and Sørensen (2007) and the references therein. In these models, one is often interested in the long-run average of the occupation time of the process in certain regions, say the occupation time of the exchange rate above or below a threshold. Unfortunately, distributional information on the functional \(\int_0^t 1_A(X_s) \, ds\) is often inaccessible, where \(1_A\) is the indicator function of the set \(A\). In practice, one may resort to the space average \(\pi(A)\) as a natural approximation of the quantity of interest \(\int_0^t 1_A(X_s) \, ds\), where the former is often easier to access than the latter. Our main result in Theorem 1.1 thus provides an invaluable tool and can be used to give non-asymptotic probabilistic error bounds on such approximation. Another situation where Theorem 1.1 is needed is about constructing confidence interval of the functional \(\int_0^t 1_A(X_s) \, ds\). One can easily construct confidence band based on these large deviation probability. With these motivations in mind, we now apply Theorem 1.1 to the Jacobi process with \(f = 1_A\) that gives:

\textbf{Corollary 1.1.} Suppose that \(X = (X_t)_{t \geq 0}\) is the Jacobi process which is uniformly ergodic with generator given by (1.1) and parameters \(b > a > 0, \sigma \in \mathbb{R}\). For any \(\varepsilon > 0, t > \frac{4t_{av}}{c}\) and measurable subset \(A \subseteq (0, 1)\), we have
\[
\mathbb{P}_x \left( \frac{1}{t} \int_0^t 1_A(X_s) \, ds - \pi(1_A) \geq \varepsilon \right) \leq \exp \left\{ \frac{-2(t\varepsilon - 4t_{av})^2}{(t+1)(4t_{av}+1)^2} \right\},
\]
where
\[
t_{av} = \frac{2}{\sigma^2} \sum_{i=1}^{\infty} \frac{1}{i \left( i - 1 + \frac{2b}{\sigma^2} \right)}
\]
is the average hitting time of the Jacobi process.

As our first remark, we note that the upper bound in (1.2) can be quite loose since it does not depend on the size of \(A\). Such a bound indeed holds as long as we have \(\|f\| \leq 1\) in Theorem 1.1. In addition, we see that this upper bound depends only on the parameters \(b\) and \(\sigma\) through \(t_{av}\) but not on \(a\).

In our second example, we introduce the finite interval analogue of the Ornstein-Uhlenbeck process first studied by Kessler and Sørensen (1999) on the state space \(S = (-\pi/2, \pi/2)\), where we take the drift to be \(\mu(x) = -\rho \tan(x)\), the diffusion coefficient to be \(\sigma^2(x) = 1\) and \(\rho \geq 1/2\) to be a parameter. That is, the generator \(A^O\) is written as
\[
A^O = -\rho \tan(x) \frac{d}{dx} + \frac{1}{2} \frac{d^2}{dx^2}.
\]
According to (Forman and Sørensen, 2008, Section 2.1) $X$ is ergodic with $\sigma_{ess}(A^O) = \emptyset$ and

\[ \lambda_i = i (\rho + i/2), \quad \sum_{i=1}^\infty \frac{1}{\lambda_i} = \sum_{i=1}^\infty \frac{1}{i (\rho + i/2)} < \infty. \] (1.4)

In view of Proposition 1.1 item (4), $X$ is thus uniformly ergodic for any $\rho \geq 1/2$. Specializing into the case $\rho = 1/2$, we see that the stationary distribution has density given by

$$\pi(x) = \frac{\cos(x)}{2} 1_{x \in (-\pi/2, \pi/2)}.$$

For $u \in \mathbb{R}$, if we take $f(x) = e^{ux} 1_{x \in (-\pi/2, \pi/2)}$ with $\|f\| \leq e^{u\pi/2}$ in Theorem 1.1, the time integral becomes

$$\int_0^t f(X_s) \, ds = \int_0^t e^{ux} \, ds,$$

the exponential functional associated with $X$. Often distributional information of exponential functionals are difficult to obtain, see for instance the book Yor (2001). One may approximate such functional by means of their space average $\pi(f)$, and our results come in handy since they give probabilistic error bound on such approximation. Theorem 1.1 now reads

**Corollary 1.2.** Suppose that $X = (X_t)_{t \geq 0}$ is the finite interval analogue of the Ornstein-Uhlenbeck process which is uniformly ergodic with generator given by (1.3) and parameter $\rho = 1/2$. For any $\varepsilon > 0$, $u \in \mathbb{R}$ and $t > \frac{4e^{u\pi/2}t_{av}}{\varepsilon}$, we have

\[ \mathbb{P}_x \left( \int_0^t e^{ux} \, ds - \frac{2t \cosh(u\pi/2)}{1 + u^2} \geq \varepsilon \right) \leq \exp \left\{ -2 \left( \frac{2t \varepsilon}{t_{av}} - 4e^{u\pi/2}t_{av} \right)^2 \right\}, \] (1.5)

where

$$t_{av} = \sum_{i=1}^\infty \frac{2}{i(i+1)} = 2$$

is the average hitting time of $X$.

For further concrete examples of uniformly ergodic diffusions with explicit eigenvalues information, we refer interested readers to the work of Forman and Sørensen (2008); Kessler and Sørensen (1999).

The rest of the paper is organized as follows. In Section 2, we first present the proof of the main result Theorem 1.1, followed by detailing the proof of Corollary 1.1 and the proof of Corollary 1.2.

2. **Proof of the main results**

2.1. **Proof of Theorem 1.1.** Suppose without loss of generality that the mean of $f$ with respect to $\pi$ is zero, that is, $\pi(f) = 0$. To begin with, it follows readily from the induced operator norm of $Q^t$ that we have

\[ \|\hat{f}\| \leq \|f\| \|Q^t\|, \] (2.1)

where we recall $\hat{f} = Q^t f$ is the solution to the Poisson equation. Now, for the large deviation probability, we see that

\[ \mathbb{P}_x \left( \frac{1}{t} \int_0^t f(X_s) \, ds - \pi(f) \geq \varepsilon \right) \leq e^{-\theta t \varepsilon} \mathbb{E}_x \left[ e^{\theta \int_0^t f(X_s) \, ds} \right] \] (2.2)
\[ (2.3) \quad e^{-\theta t} \mathbb{E}_x \left[ e^{-\int_0^t \mathcal{A} \hat{f} (X_s) \, ds} \right]. \]

In the above equation, (2.2) comes from Markov inequality, which holds for any \( \theta \geq 0 \), while (2.3) follows from the Poisson equation \(-\mathcal{A} \hat{f} = f \). Now, we explicitly construct a martingale that is useful in our analysis, namely

\[ (2.4) \quad \mathcal{M}_t^f := \hat{f} (X_t) - \hat{f} (X_0) - \int_0^t \mathcal{A} \hat{f} (X_s) \, ds. \]

Then by a classical result in (Bhattacharya and Waymire, 2009, Chapter 5 Theorem 2.3), we see that \( \mathcal{M}_t^f \) is a mean zero \( \{ \mathcal{F}_t \} \)-martingale, where we again recall \( \mathcal{F}_t = \sigma \{ X_u : 0 \leq u \leq t \} \) is the filtration of \( X \). Using (2.1) and (2.4), the tail bound in (2.3) is further upper bounded by

\[ (2.5) \quad e^{-\theta t} \mathbb{E}_x \left[ e^{-\int_0^t \mathcal{A} \hat{f} (X_s) \, ds} \right] = e^{-\theta t} \mathbb{E}_x \left[ e^{\theta (\mathcal{M}_t^f + f (X_0) - \hat{f} (X_t))} \right] \leq e^{-\theta t} e^{2 \| Q \| \| f \|} \mathbb{E}_x \left[ e^{\theta \mathcal{M}_t^f} \right]. \]

Now, we proceed to examine the bound for \( \mathbb{E}_x \left[ e^{\theta \mathcal{M}_t^f} \right] \). In order to use the classical Hoeffding’s lemma for bounded random variables (Devroye et al., 1996, Lemma 8.1), we write \( \mathcal{M}_t^f \) as

\[ \mathcal{M}_t^f = \sum_{s=1}^{[t]} (\mathcal{M}_s^f - \mathcal{M}_{s-1}^f) + \mathcal{M}_t^f - \mathcal{M}_{[t]}^f. \]

As a result, to bound the martingale \( \mathcal{M}_t^f \) it suffices to bound the martingale differences \( \mathcal{M}_s^f - \mathcal{M}_{s-1}^f \). Using the definition of \( \mathcal{M}_t^f \) in (2.4), these bounds are given by, for \( s = 1, 2, \ldots, [t] \),

\[ \mathcal{M}_s^f - \mathcal{M}_{s-1}^f = \hat{f} (X_s) - \hat{f} (X_{s-1}) - \int_{s-1}^s \mathcal{A} \hat{f} (X_s) \, ds \]

\[ \leq 2 \left\| \hat{f} \right\| + \left\| \int_{s-1}^s f (X_s) \, ds \right\| \leq 2 \left\| Q \right\| \| f \| + \| f \| = (2 \left\| Q \right\| + 1) \| f \|, \]

where we use the Poisson equation in the first inequality and (2.1) in the second inequality. Similarly,

\[ \mathcal{M}_t^f - \mathcal{M}_{[t]}^f = \hat{f} (X_t) - \hat{f} (X_{[t]}) - \int_{[t]}^t \mathcal{A} \hat{f} (X_s) \, ds \]

\[ \leq 2 \left\| \hat{f} \right\| + \| f \| \leq (2 \left\| Q \right\| + 1) \| f \|. \]

It follows from double expectation, (2.5), (2.6) and (2.7) that the upper bound in (2.3) becomes

\[ \mathbb{E}_x \left[ e^{\theta (\sum_{s=1}^{[t]} (\mathcal{M}_s^f - \mathcal{M}_{s-1}^f) + \mathcal{M}_t^f - \mathcal{M}_{[t]}^f))} \right] \leq \mathbb{E}_x \left[ e^{\theta \sum_{s=1}^{[t]} (\mathcal{M}_s^f - \mathcal{M}_{s-1}^f)} \right] \mathbb{E}_x \left[ e^{2 \| Q \| \| f \|} \right] e^{\theta (2 \left\| Q \right\| + 1) \| f \|^2} \]
\[ \leq e^{\frac{\theta^2(t+1)(2\|Q\|^2+1)^2\|f\|^2}{8}}, \]

where the first and second inequality follows from repeated applications of the Hoeffding’s lemma (Devroye et al., 1996, Lemma 8.1). Finally, collecting the above results the tail bound is given by (2.8)

\[ \mathbb{P}_x\left( \frac{1}{t} \int_0^t f(X_s)ds - \pi(f) \geq \varepsilon \right) \leq \exp\left( \frac{\theta^2(t+1)(2\|Q\|^2+1)^2\|f\|^2}{8} - \theta t\varepsilon + 2\theta \|Q\|^2\|f\|^2 \right), \]

which is minimized at \( \theta = \theta^* \) where

\[ \theta^* = \frac{4t\varepsilon - 8\|Q\|^2\|f\|}{(t+1)(2\|Q\|^2+1)^2\|f\|^2}. \]

Desired result follows by substituting \( \theta = \theta^* \) into (2.8).

### 2.2. Proof of Corollary 1.1.

Desired result follows from taking \( f = \mathbb{1}_A \) in Theorem 1.1 and utilizing the follow bound on the induced operator norm of the deviation kernel \( Q^\sharp \):

(2.9)

\[ \|Q^\sharp\| \leq 2t_{av}, \]

see e.g. (Choi, 2018, Theorem 1.1). As for the expression of the average hitting time \( t_{av} \), the eigentime identity Cheng and Mao (2015) gives

(2.10)

\[ t_{av} = \sum_{i=1}^{\infty} \frac{1}{\lambda_i} = \frac{2}{\sigma^2} \sum_{i=1}^{\infty} \frac{1}{i(i-1+2b/\sigma^2)} < \infty, \]

where \( 0 = \lambda_0 < \lambda_1 \leq \lambda_2 \leq \ldots \) are the eigenvalues of \( -A^J \) which are given by, for \( i = 0, 1, 2, \ldots \),

\[ \lambda_i = \frac{\sigma^2}{2} i \left( i - 1 + \frac{2b}{\sigma^2} \right), \]

see e.g. (Albanese and Kuznetsov, 2009, Appendix B.3).

### 2.3. Proof of Corollary 1.2.

Desired result follows from taking \( f(x) = e^{ux}\mathbb{1}_{x\in(-\pi/2,\pi/2)} \) in Theorem 1.1, and using

\[ \pi(f) = \frac{1}{2} \int_{-\pi/2}^{\pi/2} e^{ux} \cos(x) dx = \frac{2 \cosh(u\pi/2)}{1+u^2} \]

as well as the following bound on the induced operator norm of the deviation kernel \( Q^\sharp \):

\[ \|Q^\sharp\| \leq 2t_{av} = \sum_{i=1}^{\infty} \frac{4}{i(i+1)} = 4, \]

where again the first equality follows from Cheng and Mao (2015) with \( \lambda_i \) being given in (1.4) with parameter \( \rho = 1/2 \).

### Acknowledgement

Acknowledgement. We thank the anonymous referee for constructive comments that improve the presentation of the manuscript. This work is partially supported by the Chinese University of Hong Kong, Shenzhen grant PF01001143.
REFERENCES

C. Albanese and A. Kuznetsov. Transformations of Markov processes and classification scheme for solvable driftless diffusions. *Markov Process. Related Fields*, 15(4):563–574, 2009.

V. Bentkus. An extension of the Hoeffding inequality to unbounded random variables. *Lith. Math. J.*, 48(2):137–157, 2008.

R. N. Bhattacharya and E. C. Waymire. *Stochastic processes with applications*, volume 61. Siam, 2009.

T. R. Boucher. A Hoeffding inequality for Markov chains using a generalized inverse. *Statist. Probab. Lett.*, 79(8):1105–1107, 2009.

L.-J. Cheng and Y.-H. Mao. Eigentime identity for one-dimensional diffusion processes. *J. Appl. Probab.*, 52(1):224–237, 2015.

M. C. Choi. A scale function approach for Stein’s method of one-dimensional diffusion. *Submitted*, 2018.

L. Devroye, L. Györfi, and G. Lugosi. *A probabilistic theory of pattern recognition*, volume 31 of *Applications of Mathematics (New York)*. Springer-Verlag, New York, 1996.

J. L. Forman and M. Sørensen. The Pearson diffusions: a class of statistically tractable diffusion processes. *Scand. J. Statist.*, 35(3):438–465, 2008.

L. Galtchouk and S. Pergamenshchikov. Uniform concentration inequality for ergodic diffusion processes. *Stochastic Process. Appl.*, 117(7):830–839, 2007.

P. W. Glynn and S. P. Meyn. A Liapounov bound for solutions of the Poisson equation. *Ann. Probab.*, 24(2):916–931, 1996.

P. W. Glynn and D. Ormoneit. Hoeffding’s inequality for uniformly ergodic Markov chains. *Statist. Probab. Lett.*, 56(2):143–146, 2002.

W. Hoeffding. Probability inequalities for sums of bounded random variables. *J. Amer. Statist. Assoc.*, 58:13–30, 1963.

M. Kessler and M. Sørensen. Estimating equations based on eigenfunctions for a discretely observed diffusion process. *Bernoulli*, 5(2):299–314, 1999.

K. S. Larsen and M. Sørensen. Diffusion models for exchange rates in a target zone. *Math. Finance*, 17(2):285–306, 2007.

Y.-H. Mao. Deviation kernels for one-dimensional diffusion processes. *Preprint. Available at http://math.bnu.edu.cn/probab/Workshop2002/Preprints/MaoYH.pdf*, 2002.

W. Whitt. Asymptotic Formulas for Markov Processes with Applications to Simulation. *Operations Research*, 40(2):279–291, 1992.

M. Yor. *Exponential functionals of Brownian motion and related processes*. Springer Finance. Springer-Verlag, Berlin, 2001. With an introductory chapter by Hélyette Geman, Chapters 1, 3, 4, 8 translated from the French by Stephen S. Wilson.

INSTITUTE FOR DATA AND DECISION ANALYTICS, THE CHINESE UNIVERSITY OF HONG KONG, SHENZHEN, GUANGDONG, 518172, P.R. CHINA

E-mail address: michaelchoi@cuhk.edu.cn

SCHOOL OF MATHEMATICS, SUN YAT-SEN UNIVERSITY, GUANGZHOU, GUANGDONG, 510275, P.R. CHINA

E-mail address: liylin6@mail2.sysu.edu.cn