ARTICLE IN PRESS

Available online at www.sciencedirect.com

ScienceDirect

Stochastic Processes and their Applications ▮ (▮▮▮▮) ▮▮▮–▮▮▮

stochastic processes and their applications

www.elsevier.com/locate/spa

Estimation error for occupation time functionals of stationary Markov processes

Randolf Altmeyer*, Jakub Chorowski

Institut für Mathematik, Humboldt-Universität zu Berlin, Unter den Linden 6, 10099 Berlin, Germany

Received 18 October 2016; received in revised form 3 June 2017; accepted 18 August 2017
Available online xxxx

Abstract

The approximation of integral functionals with respect to a stationary Markov process by a Riemann sum estimator is studied. Stationarity and the functional calculus of the infinitesimal generator of the process are used to explicitly calculate the estimation error and to prove a general finite sample error bound. The presented approach admits general integrands and gives a unifying explanation for different rates obtained in the literature. Several examples demonstrate how the general bound can be related to well-known function spaces.

MSC 2010: primary 62M05; secondary 60J55; 60J35

Keywords: Markov processes; Integral functionals; Occupation time; Sobolev spaces; Infinitesimal generator; Ornstein–Uhlenbeck

1. Introduction

Statistics for continuous-time Markov processes is usually based on the observation of a sample path. Typically, only discrete-time observations are available. An important task is the estimation of integral functionals such as

$$\Gamma_T(f) = \int_0^T f(X_r)\,dr, \quad T \geq 0.$$

* Corresponding author.
E-mail addresses: altmeyrx@math.hu-berlin.de (R. Altmeyer), chorowsj@math.hu-berlin.de (J. Chorowski).

https://doi.org/10.1016/j.spa.2017.08.013

$X = (X_r)_{r\geq 0}$ is an $\mathcal{S}$-valued Markov process on some probability space $(\Omega, \mathcal{F}, \mathbb{P})$ for a Polish space $\mathcal{S}$ equipped with its Borel-σ-field and $f : \mathcal{S} \to \mathbb{R}$ is a given function such that $\Gamma_T(f)$ is well-defined. Functional operators of this form appear in many problems. For instance, in mathematical finance they are used to model path dependent derivatives (see Hugonnier [20], Chesney et al. [7]). In evolutionary dynamics the value $f(x)$ is often associated with the cost of staying in a state x (Pollett [26]). The most important case for applications is $f = \mathbf{1}_A$ for a Borel set A. $\Gamma_T(f)$ is then known as the *occupation time* of X in A and measures the time X spends in A until T. For general f and when X is ergodic with stationary measure μ, integral functionals are also important for studying the long term behavior of the process as $T \to \infty$, since $T^{-1}\Gamma_T(f) \to \int f d\mu$ by the ergodic theorem. Furthermore, the smoothness properties of $x \mapsto \int_0^T f(x + X_r)dr$ play an important role for solving ordinary differential equations, for example in combination with the phenomenon of regularization by noise (Catellier and Gubinelli [5]).

Our goal is to approximate $\Gamma_T(f)$ given the equidistant observations X_{t_k} at $t_k = k\Delta_n$, where $\Delta_n = T/n$ and $k = 0, \ldots, n-1$, using the Riemann-sum estimator

$$\hat{\Gamma}_{T,n}(f) = \Delta_n \sum_{k=1}^{n} f(X_{t_{k-1}}).$$

We study the $L^2(\mathbb{P})$-error of this approximation. Consistency follows from Riemann approximation already under weak assumptions on f and X. The rate of this convergence, however, depends on f, X, T and Δ_n. We are interested in finite sample bounds making these dependencies explicit, independent of Δ_n being fixed or tending to 0, and independent of T being fixed or tending to infinity. The Riemann sum estimator has appeared in many places in the literature, mostly for estimating the occupation time (Chorowski [9, Theorem 42], Gobet and Matulewicz [17, Section 2]), or as a proxy for approximating the local time of X in the diffusion case, such as in Hoffmann [19, Proposition 3]. For general f see also Dion and Genon-Catalot [11, Section 5]. Error bounds are usually derived ad-hoc, leading to suboptimal bounds or without explicit constants for the dependence on parameters.

The approximation error clearly depends on the smoothness of f. Note, however, when measuring the approximation error in the $L^2(\mathbb{P})$-sense, it is not possible, in general, to have a faster rate of convergence than $T^{1/2}\Delta_n$, even with smooth f. Indeed, it can be checked for X being a Brownian motion and f the identity that $\|\Gamma_T(f) - \hat{\Gamma}_{T,n}(f)\|_{L^2(\mathbb{P})} = T^{1/2}\Delta_n/\sqrt{3}$. A systematic study of this approximation problem has started only recently. For one-dimensional diffusion processes with smooth coefficients Ngo and Ogawa [24, Theorem 2.2] obtain the rate $T^{1/2}\Delta_n^{3/4}$ for the special case of indicator functions $f = \mathbf{1}_{[K,\infty)}$, $K \in \mathbb{R}$. Interestingly, they also provide a lower bound with this rate in the $L^2(\mathbb{P})$-sense. The specific analysis for indicator functions, however, cannot explain which rates we can expect for more general f. For Markov processes in $\mathbb{R}^d$ with transition kernels satisfying certain heat kernel bounds Ganychenko et al. [15] prove for bounded f the $L^p(\mathbb{P})$-bound $\|f\|_\infty T^{1/2}(\Delta_n \log n)^{1/2}$, $p \geq 2$. Results for α-Hölder functions for $0 < \alpha \leq 1$ are obtained by Kohatsu-Higa et al. [22, Theorem 2.3], again for one-dimensional diffusions with smooth coefficients, and by Ganychenko [14], for the same class of Markov processes as in Ganychenko et al. [15]. They obtain $\|f\|_\alpha T^{1/2}\Delta_n^{(1+\alpha)/2}$ as upper bound, both of them losing an additional $\log n$-factor when $\alpha = 1$. Surprisingly, indicators obey the same bounds as $1/2$-Hölder functions.

The aim of this paper is twofold. First, we want to study the approximation of $\Gamma_T(f)$ by $\hat{\Gamma}_{T,n}(f)$ for more general functions f on arbitrary state spaces $\mathcal{S}$ and to find a unifying mathematical explanation for the different rates obtained in the literature. Second, we want to identify the key quantities driving the estimation error. For this, we focus on the special but important case of stationary Markov processes, because this allows us to calculate the error explicitly in terms of the associated semigroup. The main insight of our results is that the discretization error depends on the action of fractional powers of the infinitesimal generator applied to f. Several examples demonstrate how this can be related to more familiar L^2-Sobolev norms of f. These norms are the key to explain both the rates for Hölder and indicator functions by suitable interpolation. The dependence on T in the error is explicit. This allows us, when X is ergodic and when $T \to \infty$, to approximate integral functionals with respect to the stationary measure of X under weaker conditions than the ones commonly used in the literature. Our approach is based on the functional calculus of the generator. We therefore consider only stationary Markov processes whose generators are normal operators. While stationarity is assumed to hold in many applications and statistical procedures (for instance in several of the works mentioned above), our method still applies when the initial distribution is only absolutely continuous with respect to the stationary measure.

While stationarity makes explicit calculations possible, the rates of convergence obtained here also hold for L^2-Sobolev functions and Markov processes without invariant distributions or even non-Markovian processes. For more details the reader is referred to Altmeyer [2].

The paper is organized as follows. In Section 2 we state a general upper bound for the $L^2(\mathbb{P})$-approximation error. We apply it to approximate integral functionals with respect to the stationary measure, when X is ergodic. In Section 3 we study several concrete examples of processes and functions. We also discuss the important example of Brownian motion, which is not stationary, but which can be approximated by reflected stationary diffusions. Proofs can be found in Section 4. The Appendix contains a brief summary of the most important facts about semigroups and the functional calculus for normal operators.

2. A general upper bound

In the following, X is a continuous-time Markov process on some probability space $(\Omega, \mathcal{F}, \mathbb{P})$ and with Polish state space $\mathcal{S}$. For any measure μ on $\mathcal{S}$ denote by $L^2(\mu) := L^2(\mathcal{S}, \mu)$ the space of square integrable functions $f : \mathcal{S} \to \mathbb{R}$ with respect to μ and with norm $\|f\|_\mu = (\int f^2 d\mu)^{1/2}$. $\|\cdot\|_{\infty,\mu}$ denotes the sup-norm in $L^\infty(\mu)$ and $|\cdot|$ is the Euclidean norm on $\mathbb{R}^d$ or $\mathbb{R}^{d\times d}$. $a \lesssim b$ for $a, b \in \mathbb{R}$ means $a \leq Cb$ for some constant $C > 0$. $Z_n = O_{\mathbb{P}}(a_n)$ for a sequence of random variables $(Z_n)_{n\geq 1}$ and real numbers $(a_n)_{n\geq 1}$ means that $a_n^{-1} Z_n$ is tight. For the basic concepts of semigroup theory and functional calculus refer to the Appendix. Our main assumptions are the following:

Assumption 2.1. X is a stationary time-homogeneous Markov process with stationary measure μ. The associated semigroup $(P_r)_{r\geq 0}$ is Feller and its infinitesimal generator L with respect to $L^2(\mu()$ is a normal operator.

These assumptions are satisfied for many important processes. A leading example is the standard Ornstein–Uhlenbeck process. Note that for $f \in L^2(\mu)$ both $\Gamma_T(f)$ and $\hat{\Gamma}_{T,n}(f)$ are μ-almost surely well-defined random variables in $L^2(\mathbb{P})$. Consider the operators $|L|^{s/2}$, $s \geq 0$, which are defined via the functional calculus of L. They have domains $\mathcal{D}(|L|^{s/2}) \subset L^2(\mu)$ and thus contain all $f \in L^2(\mu)$ with $\||L|^{s/2} f\|_\mu < \infty$. If X is an Ornstein–Uhlenbeck process, then the related spaces $\mathcal{D}((I - L)^{s/2}) \subset \mathcal{D}^s(L)$ are known as Bessel potential spaces and play an important role in Malliavin calculus (Watanabe [28]). We are now ready to state the general upper bound.

Please cite this article in press as: R. Altmeyer, J. Chorowski, Estimation error for occupation time functionals of stationary Markov processes, Stochastic Processes and their Applications (2017), https://doi.org/10.1016/j.spa.2017.08.013.

Theorem 2.2. *Let X be a Markov process satisfying Assumption 2.1 with $X_0 \overset{d}{\sim} \mu$. There exists a universal constant C such that for all $f \in \mathcal{D}(|L|^{s/2})$, $0 \le s \le 1$,*

$$\left\| \Gamma_T(f) - \hat{\Gamma}_{T,n}(f) \right\|_{L^2(\mathbb{P})} \le C \| |L|^{s/2} f \|_\mu T^{1/2} \Delta_n^{\frac{1+s}{2}}.$$

The proof of this theorem is remarkably short. For $s = 0$ it follows that $\mathcal{D}(|L|^0) = L^2(\mu)$ and the rate is $T^{1/2}\Delta_n^{1/2}$. This improves on Ganychenko et al. [15] by removing an additional $\sqrt{\log n}$. Since $\mathcal{D}(|L|^{s/2}) \subset \mathcal{D}(|L|^{1/2})$ for $s \ge 1$, the rate is never better than $T^{1/2}\Delta_n$. For $0 < s < 1$ the bound interpolates between the two extreme cases. A deeper understanding of the spaces $\mathcal{D}(|L|^{s/2})$ requires more explicit knowledge about the generator. For example, if L is self-adjoint, then $|L| = -L$ and thus $\| |L|^{1/2} f \|_\mu^2 = \|(-L)^{1/2} f\|_\mu^2 = \langle -Lf, f\rangle_\mu$. This is the Dirichlet form associated with L and μ. It is typically easier to analyze than studying $\mathcal{D}(|L|^{1/2})$ directly in terms of the functional calculus. Important examples are diffusions on $\mathbb{R}^d$ such that for sufficiently smooth functions f the Dirichlet form is bounded by $\| |\nabla f| \|_\mu^2$, the $L^2(\mu)$-norm of the gradient of f. This immediately leads to upper bounds for Hölder and indicator functions. Up to some additional conditions, we will show that α-Hölder functions lie in $\mathcal{D}(|L|^{\alpha/2})$ and indicator functions of certain cylinder sets of $\mathbb{R}^d$ lie in $\mathcal{D}(|L|^{1/4})$. This gives a unifying explanation for the different rates (see also Remark 3.8). These and other examples will be discussed in Section 3.

The assumption of starting in the stationary distribution can be relaxed to some extent. Indeed, if the initial distribution is absolutely continuous with respect to μ, then the result of Theorem 2.2 remains valid. More generally, if the distribution of X_{T_0}, $T_0 \ge 0$, is absolutely continuous with respect to μ, then the result still applies, if instead of $\Gamma_T(f)$ the modified occupation time functional $\Gamma_{T_n,T}(f) = \int_{T_n}^T f(X_r)dr$ is estimated by $\hat{\Gamma}_{T_n,T,n}(f) = \Delta_n \sum_{k=T_n\Delta_n^{-1}+1}^n f(X_{t_{k-1}})$, where $T_n = \lceil T_0/\Delta_n^{-1} \rceil \Delta_n$. Clearly, $\Gamma_T(f) = \Gamma_{0,T}(f)$ and $\hat{\Gamma}_{T,n}(f) = \hat{\Gamma}_{0,T,n}(f)$. This yields the following corollary.

Corollary 2.3. *Let X be a Markov process satisfying Assumption 2.1. Assume that $X_{T_0} \overset{d}{\sim} \eta$, $T_0 \ge 0$, for a probability measure η such that $\eta \ll \mu$ with density $d\eta/d\mu$. There exists a universal constant C such that for all $f \in \mathcal{D}(|L|^{s/2})$, $0 \le s \le 1$,*

$$\left\| \Gamma_{T_n,T}(f) - \hat{\Gamma}_{T_n,T,n}(f) \right\|_{L^2(\mathbb{P})} \le C \left\| \frac{d\eta}{d\mu} \right\|_{\infty,\mu}^{1/2} \| |L|^{s/2} f \|_\mu T^{1/2} \Delta_n^{\frac{1+s}{2}}.$$

As an example consider the Ornstein–Uhlenbeck process which is a Gaussian process and therefore every X_{T_0} for $T_0 > 0$ is normally distributed and therefore the distribution of X_{T_0} is absolutely continuous with respect to the stationary measure of X (cf. Example 3.1).

Instead of $\Gamma_T(f)$, a different target functional is often $\int f d\mu$. It is well-known that $T^{-1}\Gamma_T(f)$ is $T^{1/2}$-consistent for $\int f d\mu$, i.e. $T^{-1}\Gamma_T(f) - \int f d\mu = O_{\mathbb{P}}(T^{-1/2})$, when L is self-adjoint and $f \in \mathcal{D}((-L)^{-1/2})$ (see e.g. Kipnis and Varadhan [21, Theorem 1.8]). By Theorem 2.2 this can be extended to the estimator $T^{-1}\hat{\Gamma}_{T,n}(f)$ and more general $L^2(\mu)$-functions.

Theorem 2.4. *Let X be a Markov process satisfying Assumption 2.1 with $X_0 \overset{d}{\sim} \mu$. There exists a universal constant C such that for all $f \in L^2(\mu)$ with $f_0 \in \mathcal{D}(|L|^{-1/2})$, $f_0 = f - \int f d\mu$,*

$$\left\| T^{-1}\hat{\Gamma}_{T,n}(f) - \int_{\mathcal{S}} f(x)\, d\mu(x) \right\|_{L^2(\mathbb{P})} \le C T^{-1/2} \left(\|f\|_\mu \Delta_n^{1/2} + \left\| |L|^{-1/2} f_0 \right\|_\mu \right).$$

Using Corollary 2.3 the assumption of starting in the stationary distribution can be relaxed. As an example for $\mathcal{D}(|L|^{-1/2})$ being non-trivial, assume that 0 is a simple eigenvalue of L and that L has a spectral gap, i.e. $s_0 > 0$, where $s_0 = \sup\{r > 0 : B(0,r) \cap \sigma(L) = \{0\}\}$ and $B(0,r) = \{z \in \mathbb{C} : |z| \le r\}$. In that case, X is ergodic and it can be shown that $f_0 \in \mathcal{D}(|L|^{-1/2})$ is satisfied whenever f is non-constant (Bakry et al. [3, Section 4.2.1]). Furthermore, the upper bound of the theorem simplifies, since

$$\left\| |L|^{-1/2} f_0 \right\|_\mu \le s_0^{-1/2} \|f_0\|_\mu \le s_0^{-1/2} \|f\|_\mu.$$

A concrete example of a process with spectral gap is the Ornstein–Uhlenbeck process (Bakry et al. [3, Chapter 4]). In general, Theorem 2.4 shows that in order to achieve the rate $T^{1/2}$ as $n, T \to \infty$ there is essentially no gain in the high-frequency case, i.e. $\Delta_n \to 0$, compared to the low-frequency case with Δ_n fixed. The error bound improves on the commonly used condition in the literature that $T\Delta_n \lesssim 1$ to achieve $T^{1/2}$-consistency (see e.g. Dion and Genon-Catalot [11, Section 5]). It is interesting to note that Theorem 2.4 depends on negative powers of $|L|$, while Theorem 2.2 depends on positive powers of $|L|$.

3. Examples

In this section, the general bound from Theorem 2.2 is applied to several important examples. We first study Markov jump processes, i.e. continuous time Markov processes with countable state spaces. We then consider a special class of diffusion processes for which the spaces $\mathcal{D}(|L|^{s/2})$ can be described via the Dirichlet form $\langle -Lf, f\rangle_\mu$. After that, we show for the one-dimensional Brownian motion how the assumption of stationarity can be removed. Finally, we show that our method also applies to infinite dimensional diffusions.

3.1. *Markov-jump processes*

Consider a continuous-time Markov process $(X_r)_{r\ge 0}$ on a countable state space $\mathcal{S}$. Such a process can always be realized as $X_r = Y_{N_r}$ for a Markov chain $(Y_s)_{s\in\mathcal{S}}$ starting in some initial distribution μ with transition probabilities $(P_{xy})_{x,y\in S}$ and an independent Poisson process $(N_r)_{r\ge 0}$ with intensity $0 < \lambda < \infty$ (Ethier and Kurtz [13, Section 4.2]). Observing a path of X at the discrete times $0, \Delta_n, 2\Delta_n, \ldots, (n-1)\Delta_n$, the jump times can be identified with Δ_n precision. Hence, if the function f is bounded, then every jump contributes at most $2\|f\|_\infty$ to the estimation error $|\Gamma_T(f) - \hat{\Gamma}_{T,n}(f)|$. This yields the bound

$$\|\Gamma_T(f) - \hat{\Gamma}_{T,n}(f)\|_{L^2(\mathbb{P})} \le 2\|f\|_\infty \mathbb{E}\left[N_T^2\right]^{1/2} \Delta_n = 2\|f\|_\infty \left(\lambda T + (\lambda T)^2\right)^{1/2} \Delta_n.$$

This gives the optimal rate Δ_n but requires the function f to be bounded. Moreover, the error grows linearly in T as opposed to $T^{1/2}$ in Theorem 2.2. This can be improved, if X is stationary with stationary measure μ and reversible, i.e. $P^\top = P$. Then the infinitesimal generator $L = \lambda(P - I)$ is a bounded, non-negative self-adjoint operator. Therefore, $\|(-L)^{1/2} f\|_\mu \le \|(-L)^{1/2}\| \|f\|_\mu \le \lambda^{1/2} \|f\|_\mu$ with operator norm $\|(-L)^{1/2}\|$. It follows that $\mathcal{D}((-L)^{1/2}) = L^2(\mu)$ and Theorem 2.2 implies

$$\left\| \Gamma_T(f) - \hat{\Gamma}_{T,n}(f) \right\|_{L^2(\mathbb{P})} \le C\lambda^{1/2} \|f\|_\mu T^{1/2} \Delta_n.$$

Note that the results of Ganychenko et al. [15], Ganychenko [14] do not apply here, because the state space is countable and therefore heat kernel bounds are not available.

Please cite this article in press as: R. Altmeyer, J. Chorowski, Estimation error for occupation time functionals of stationary Markov processes, Stochastic Processes and their Applications (2017), https://doi.org/10.1016/j.spa.2017.08.013.

3.2. Diffusions with generator in divergence form

Let L be an elliptic operator in divergence form (cf. Bass [4, Chapter VII]) and let $(X_r)_{r\geq 0}$ be the associated diffusion process (in the sense of Bass [4, Section I.2]) with or without reflection arising as the solution of some stochastic differential equation. Assume that the process is stationary and takes its values in some closed subset $U \subset \mathcal{S} := \mathbb{R}^d$. Then the stationary measure μ has support in U. In case $U \subsetneq \mathbb{R}^d$ we think of μ as a measure on $\mathbb{R}^d$ and embed the domain of the infinitesimal generator $\mathcal{D}(L) \subset L^2(U, \mu)$ canonically into $L^2(\mathbb{R}^d, \mu)$ by letting $Lf := L\tilde{f}$, whenever $f|_U = \tilde{f}$ for $f \in L^2(\mathbb{R}^d, \mu)$, $\tilde{f} \in L^2(U, \mu)$. Finally, assume that L satisfies

$$\langle -Lf, g\rangle_\mu = \int_{\mathbb{R}^d} \langle A(x)\nabla f(x), \nabla g(x)\rangle_{\mathbb{R}^d}\, d\mu(x), \quad f, g \in \mathcal{D}(L) \cap C^2(\mathbb{R}^d), \tag{1}$$

for a measurable function $x \mapsto A(x) \in \mathbb{R}^{d\times d}$ such that $A(x)$ is symmetric, positive definite for all $x \in \mathbb{R}^d$ and such that $\|\,|A|\,\|_{\infty,\mu}$ is finite, where $|\cdot|$ is any matrix norm. Observe that the right hand side of the last line is also well-defined for $L^2(\mathbb{R}^d, \mu)$-integrable functions $f, g \in C^1(\mathbb{R}^d)$. An operator L satisfying (1) is self-adjoint on $\mathcal{D}(L) \cap C^2(\mathbb{R}^d)$. Observe for $f \in \mathcal{D}(L) \cap C^2(\mathbb{R}^d) \subset \mathcal{D}(L) \subset \mathcal{D}(|L|^{s/2})$ and $0 \leq s \leq 1$ that

$$\begin{aligned}\|\,|L|^{s/2} f\|_\mu^2 &= \|(-L)^{s/2} f\|_\mu^2 \leq \|(I-L)^{s/2} f\|_\mu^2 \leq \|(I-L)^{1/2} f\|_\mu^2 \\ &= \langle f - Lf, f\rangle_\mu \leq \|f\|_\mu^2 + \|\,|A|\,\|_{\infty,\mu} \|\,|\nabla f|\,\|_\mu^2 \\ &\leq \max\left(1, \|\,|A|\,\|_{\infty,\mu}\right) \|f\|_{H^1(\mu)}^2,\end{aligned} \tag{2}$$

where

$$\|f\|_{H^1(\mu)} = \|f\|_\mu + \|\,|\nabla f|\,\|_\mu$$

is the μ-weighted Sobolev norm. Combining this with Theorem 2.2 yields

$$\begin{aligned}&\left\|\Gamma_T(f) - \hat{\Gamma}_{T,n}(f)\right\|_{L^2(\mathbb{P})} \\ &\leq \begin{cases} C \max\left(1, \|\,|A|\,\|_{\infty,\mu}^{1/2}\right) \|f\|_{H^1(\mu)} T^{1/2} \Delta_n, & f \in \mathcal{D}(L) \cap C^2(\mathbb{R}^d), \\ C\|f\|_\mu T^{1/2} \Delta_n^{1/2}, & f \in L^2(\mu). \end{cases}\end{aligned} \tag{3}$$

By interpolating between the two cases $f \in L^2(\mu)$ and $f \in \mathcal{D}(L) \cap C^2(\mathbb{R}^d)$ we will study Hölder and indicator functions. Before doing this let us discuss some important examples where (1) holds.

Example 3.1 (*Ornstein–Uhlenbeck Process*). Assume that $(X_r)_{r\geq 0}$ satisfies the stochastic differential equation

$$dX_r = -X_r dr + \sqrt{2} dW_r$$

in $\mathbb{R}^d$ where $(W_r)_{r\geq 0}$ is a d-dimensional Brownian motion. If $X_0 \overset{d}{\sim} \mu$, where μ has Lebesgue density $d\mu(x)/d\lambda = (2\pi)^{-d/2} \exp(-|x|^2/2)$, then X is stationary with stationary measure μ. The infinitesimal generator L satisfies

$$Lf(x) = -\langle x, \nabla f(x)\rangle_{\mathbb{R}^d} + \Delta f(x), \quad x \in \mathbb{R}^d, \tag{4}$$

with $f \in \mathcal{D}(L) = H^2(\mu)$, the μ-weighted Sobolev space of twice weakly differentiable functions with all partial derivatives up to order two belonging to $L^2(\mu)$. Using integration by parts it follows that

$$\langle -Lf, g\rangle_\mu = \int_{\mathbb{R}^d} \langle \nabla f(x), \nabla g(x)\rangle_{\mathbb{R}^d}\, d\mu(x), \quad f, g \in C^2(\mathbb{R}^d) \tag{5}$$

(cf. Pavliotis [25, Section 4.4]). Hence, L is a self-adjoint operator of the form (1) with $A_{jk} = \mathbf{1}(j = k)$ for all $1 \le j, k \le d$. This example can be generalized considerably (see Chojnowska-Michalik and Goldys [8] and Section 3.4 below).

Example 3.2 (*Scalar Diffusion with Possibly Attracting Boundaries*). Fix some boundaries $-\infty \le \beta < \rho \le \infty$. Assume that $(X_r)_{r\ge 0}$ is a stationary diffusion process on $[\beta, \rho]$ solving the one-dimensional stochastic differential equation

$$dX_r = b(X_r)dr + \sigma(X_r)dW_r, \tag{6}$$

for a continuous drift $b : [\beta, \rho] \to \mathbb{R}$, strictly positive continuous volatility $\sigma : [\beta, \rho] \to (0, \infty)$ and a one-dimensional Brownian motion $(W_r)_{r\ge 0}$. Sufficient conditions for the existence of such a process can be found in Hansen et al. [18, Section 3.1]. In particular, stationarity is guaranteed if the *speed density*

$$m(x) = \frac{1}{\sigma^2(x)} \exp\Big(\int_{x_0}^x \frac{2b(y)}{\sigma^2(y)} dy\Big), \quad \beta \le x_0, x \le \rho,$$

is integrable on $[\beta, \rho]$. Then the stationary measure has the density

$$\frac{d\mu(x)}{d\lambda} = C_0 m(x)\, \mathbf{1}(\beta < x < \rho),$$

where C_0 is a normalizing constant. The infinitesimal generator L satisfies

$$\begin{aligned} Lf(x) &= b(x)f'(x) + \frac{\sigma^2(x)}{2} f''(x) \\ &= \frac{1}{2}\Big(\frac{d\mu(x)}{d\lambda}\Big)^{-1} \Big(f'(x)\sigma^2(x)\frac{d\mu(x)}{d\lambda}\Big)', \quad \beta < x < \rho, \end{aligned}$$

with $f \in \mathcal{D}(L)$, where

$$\begin{aligned} \mathcal{D}(L) = \Big\{ & f \in L^2([\beta, \rho], \mu) : f \text{ and } f' \text{ are absolutely continuous with} \\ & \lim_{x \searrow \lambda} f'(x)m(x)\sigma^2(x) = \lim_{x \nearrow \rho} f'(x)m(x)\sigma^2(x) = 0 \text{ and} \\ & Lf \in L^2([\beta, \rho], \mu) \Big\}. \end{aligned}$$

For details see Section 3.3 of Hansen et al. [18]. Embedding the domain into $L^2(\mu)$ as mentioned before and integrating by parts it follows that

$$\langle -Lf, g\rangle_\mu = \int_{\mathbb{R}} f'(x)g'(x)\sigma^2(x)\, d\mu(x), \quad f, g \in \mathcal{D}(L) \cap C^2(\mathbb{R}), \tag{7}$$

which is of the form (1) with $A = \sigma^2$. For $b(x) = -x$ and $\sigma(x) = \sqrt{2}$, X is just the one-dimensional Ornstein–Uhlenbeck process.

Example 3.3 (*Reflected Diffusion*). Fix some boundaries $-\infty < \beta < \rho < \infty$. Assume that X is a one-dimensional reflected diffusion on $[\beta, \rho]$. By this we mean that X satisfies the Skorokhod type stochastic differential equation

$$dX_r = b(X_r)dr + \sigma(X_r)dW_r + dK_r, \tag{8}$$

for a bounded measurable drift $b : [\beta, \rho] \to \mathbb{R}$, strictly positive continuous volatility $\sigma : [\beta, \rho] \to (0, \infty)$, $(W_r)_{r\geq 0}$ is a Brownian motion and $(K_r)_{r\geq 0}$ is an adapted continuous process with finite variation starting from 0 and such that for every $r \geq 0$

$$\int_0^r \mathbf{1}_{(\beta,\rho)}(X_s) dK_s = 0.$$

The stationary measure and the generator L are as in the last example. Since $[\beta, \rho]$ is compact, the domain simplifies to

$$\mathcal{D}(L) = \Big\{ f \in L^2([\beta, \rho], \mu) : f \text{ and } f' \text{ are absolutely continuous with} \\ f'(\lambda) = f'(\rho) = 0 \text{ and } Lf \in L^2([\beta, \rho], \mu) \Big\}.$$

Therefore (1) holds here, as well. For more details see Chorowski [9, Section 1.1].

3.2.1. Hölder functions

Consider an α-Hölder continuous function $f : \mathbb{R}^d \to \mathbb{R}$, $0 \leq \alpha \leq 1$, with finite Hölder-norm

$$\|f\|_\alpha = \sup_{x \neq y \in \mathbb{R}^d} \frac{|f(x) - f(y)|}{|x-y|^\alpha},$$

and such that $f \in L^2(\mu)$. Let $(\varphi_\varepsilon)_{\varepsilon\geq 0}$ be a non-negative smooth kernel, i.e. $\varphi_\varepsilon(x) = \varepsilon^{-1}\varphi(\varepsilon^{-1}x)$, $0 \leq \varphi \in C_c^\infty(\mathbb{R}^d)$, $\operatorname{supp}(\varphi) \subset [-1, 1]^d$, $\int_{\mathbb{R}^d} \varphi(x)\,dx = 1$. Then the convolution $f_\varepsilon = f * \varphi_\varepsilon$ lies in $C^\infty(\mathbb{R}^d)$ and has bounded derivatives. Hence $f_\varepsilon \in L^2(\mu) \cap C^2(\mathbb{R}^d)$. It is not clear that $f_\varepsilon \in \mathcal{D}(L)$ due to possible boundary conditions as in the examples above. In order to extend (3) assume the following:

Assumption 3.4. $\mathcal{D}(L) \cap C^2(\mathbb{R}^d)$ is dense in $L^2(\mu) \cap C^1(\mathbb{R}^d)$ with respect to $\|\cdot\|_{H^1(\mu)}$.

This assumption is relatively weak and is satisfied in all the examples above. In particular, if there are no boundary conditions for $f \in \mathcal{D}(L)$, then $L^2(\mu) \cap C^2(\mathbb{R}^d) = \mathcal{D}(L) \cap C^2(\mathbb{R}^d)$, as is the case for the Ornstein–Uhlenbeck process. By approximation (3) can thus be extended to

$$\left\| \Gamma_T(f) - \hat{\Gamma}_{T,n}(f) \right\|_{L^2(\mathbb{P})} \\ \leq \begin{cases} C \max\left(1, \|\,|A|\,\|_{\infty,\mu}^{1/2}\right) \|f\|_{H^1(\mu)} T^{1/2} \Delta_n, & f \in L^2(\mu) \cap C^1(\mathbb{R}^d), \\ C\|f\|_\mu T^{1/2} \Delta_n^{1/2}, & f \in L^2(\mu). \end{cases} \tag{9}$$

Note that $f_\varepsilon \in L^2(\mu) \cap C^1(\mathbb{R}^d)$. Using $\int \varphi(x)\,dx = 1$ and $\int \nabla\varphi(x)dx = 0$, it follows that

$$\|f - f_\varepsilon\|_\mu^2 = \int \left| \int (f(x) - f(x + \varepsilon y))\, \varphi(y)\, dy \right|^2 d\mu(x) \lesssim \|f\|_\alpha^2 \varepsilon^{2\alpha},$$

$$\begin{aligned} \|\,|\nabla f_\varepsilon|\,\|_\mu^2 &= \int \left| \frac{f(x)}{\varepsilon} \int \nabla\varphi(y)\, dy - \nabla f_\varepsilon(x) \right|^2 d\mu(x) \\ &= \frac{1}{\varepsilon^2} \int \left| \int (f(x) - f(x + \varepsilon y))\, \nabla\varphi(y)\, dy \right|^2 d\mu(x) \\ &\lesssim \|f\|_\alpha^2 \varepsilon^{2\alpha-2}. \end{aligned}$$

Please cite this article in press as: R. Altmeyer, J. Chorowski, Estimation error for occupation time functionals of stationary Markov processes, Stochastic Processes and their Applications (2017), https://doi.org/10.1016/j.spa.2017.08.013.

From $\|f_\varepsilon\|_{H^1(\mu)} \lesssim \|f_\varepsilon\|_\mu + \||\nabla f_\varepsilon|\|_\mu \leq \|f - f_\varepsilon\|_\mu + \|f\|_\mu + \||\nabla f_\varepsilon|\|_\mu$ this yields with (9) that

$$\begin{aligned}
&\left\|\Gamma_T(f) - \hat{\Gamma}_{T,n}(f)\right\|_{L^2(\mathbb{P})} \\
&\quad\leq \left\|\Gamma_T(f - f_\varepsilon) - \hat{\Gamma}_{T,n}(f - f_\varepsilon)\right\|_{L^2(\mathbb{P})} + \left\|\Gamma_T(f_\varepsilon) - \hat{\Gamma}_{T,n}(f_\varepsilon)\right\|_{L^2(\mathbb{P})} \\
&\quad\lesssim \|f\|_\alpha T^{1/2}\Delta_n^{1/2}\varepsilon^\alpha + \|f\|_\alpha T^{1/2}\Delta_n\varepsilon^{\alpha-1} + \|f\|_\mu T^{1/2}\Delta_n.
\end{aligned}$$

Choosing $\varepsilon = \Delta_n^{1/2}$ implies the bound $\|f\|_\alpha T^{1/2}\Delta_n^{(1+\alpha)/2} + \|f\|_\mu T^{1/2}\Delta_n$. Up to the second term, which is of smaller order as long as $\alpha < 1$, these are the rates obtained by Kohatsu-Higa et al. [22] and Ganychenko [14]. This can be improved, if L satisfies a Poincaré type inequality, i.e. if there exists a constant $c < \infty$ such that for all $f \in \mathcal{D}(L)$

$$\|f_0\|_\mu \leq c\| |\nabla f| \|_\mu, \tag{10}$$

where $f_0 = f - \int f d\mu$. Let $f_{0,\varepsilon} = f_0 * \varphi_\varepsilon$. Then $\|f_\varepsilon\|_\alpha = \|f_{0,\varepsilon}\|_\alpha$ and it follows that $\|f_{0,\varepsilon}\|_{H^1(\mu)} \lesssim \||\nabla f_\varepsilon|\|_\mu \lesssim \|f\|_\alpha\varepsilon^{\alpha-1}$. With $\varepsilon = \Delta_n^{1/2}$ this implies

$$\left\|\Gamma_T(f) - \hat{\Gamma}_{T,n}(f)\right\|_{L^2(\mathbb{P})} = \left\|\Gamma_T(f_0) - \hat{\Gamma}_{T,n}(f_0)\right\|_{L^2(\mathbb{P})} \lesssim \|f\|_\alpha T^{1/2}\Delta_n^{\frac{1+\alpha}{2}}. \tag{11}$$

Poincaré inequalities hold for many stationary measures μ, for example for the Ornstein–Uhlenbeck process and in Example 3.2 when $m(x)$ is uniformly bounded from above and below. For other examples see Bakry et al. [3, Chapter 4] and Chen [6]. Observe that for $\alpha = 1$ the upper bound is $\|f\|_1 T^{1/2}\Delta_n$, removing an additional $\sqrt{\log n}$ term present in the results of Kohatsu-Higa et al. [22] and Ganychenko [14]. In summary, we have shown the following:

Theorem 3.5. *Let X be a stationary diffusion with values in $\mathbb{R}^d$ and stationary measure μ, whose generator L satisfies (1) and Assumption 3.4. There exists a constant $C < \infty$ such that for all α-Hölder continuous functions f, $0 \leq \alpha \leq 1$,*

$$\left\|\Gamma_T(f) - \hat{\Gamma}_{T,n}(f)\right\|_{L^2(\mathbb{P})} \leq C\left(\|f\|_\alpha T^{1/2}\Delta_n^{\frac{1+\alpha}{2}} + \|f\|_\mu T^{1/2}\Delta_n\right).$$

If L satisfies a Poincaré type inequality as in (10) for some $c < \infty$, then the upper bound is $C\|f\|_\alpha T^{1/2}\Delta_n^{\frac{1+\alpha}{2}}$.

3.2.2. Indicator functions

Let $d = 1$ and consider $f = \mathbf{1}_{[K,\infty)}$, $K \in \mathbb{R}$, such that $f \in L^2(\mu)$. Let $(\varphi_\varepsilon)_{\varepsilon>0}$ be a non-negative smooth kernel as in the previous example. Then $f_\varepsilon = f * \varphi_\varepsilon$ is bounded by 1 and lies in $L^2(\mu) \cap C^2(\mathbb{R})$. $f - f_\varepsilon$ has support in $[K - \varepsilon, K + \varepsilon]$ such that

$$\|f - f_\varepsilon\|_\mu^2 \leq \int_{K-\varepsilon}^{K+\varepsilon} d\mu,$$

$$\begin{aligned}
\|f_\varepsilon'\|_\mu^2 &= \int \left|\frac{f(x)}{\varepsilon}\int \varphi'(y)\,dy - f_\varepsilon'(x)\right|^2 d\mu(x) \\
&= \frac{1}{\varepsilon^2}\int_{K-\varepsilon}^{K+\varepsilon} \left|\int (f(x) - f(x+\varepsilon y))\,\varphi'(y)\,dy\right|^2 d\mu(x) \\
&\lesssim \frac{1}{\varepsilon^2}\int_{K-\varepsilon}^{K+\varepsilon} d\mu.
\end{aligned}$$

Please cite this article in press as: R. Altmeyer, J. Chorowski, Estimation error for occupation time functionals of stationary Markov processes, Stochastic Processes and their Applications (2017), https://doi.org/10.1016/j.spa.2017.08.013.

As before, $\|f_\varepsilon\|_\mu \le \|f - f_\varepsilon\|_\mu + \|f\|_\mu$. If μ is absolutely continuous with respect to the Lebesgue measure λ with bounded density $d\mu/d\lambda$, then $\varepsilon^{-1}\int_{K-\varepsilon}^{K+\varepsilon} d\mu$ is bounded and in that case it follows from (3), uniformly in K, that

$$\left\|\Gamma_T(f) - \hat{\Gamma}_{T,n}(f)\right\|_{L^2(\mathbb{P})} \lesssim T^{1/2}(\Delta_n\varepsilon)^{1/2} + T^{1/2}\Delta_n\varepsilon^{-1/2} + T^{1/2}\Delta_n. \tag{12}$$

The last term is of lower order compared to the first two. Hence, choosing $\varepsilon = \Delta_n^{1/2}$ yields the rate $T^{1/2}\Delta_n^{3/4}$ obtained by Ngo and Ogawa [24] and Kohatsu-Higa et al. [22] for one-dimensional diffusions. However, now the rate is uniform in K with explicit dependence on T. These arguments can easily be extended to general dimension d implying the following theorem.

Theorem 3.6. *Let X be a stationary diffusion with values in $\mathbb{R}^d$ and stationary measure μ, whose generator L satisfies (1) and Assumption 3.4. Assume that μ has bounded Lebesgue density. If f is an indicator function in $\mathbb{R}^d$ of $[K_1, L_1) \times \cdots \times [K_d, L_d)$, $-\infty < K_j < L_j \le \infty$, $1 \le j \le d$, then*

$$\left\|\Gamma_T(f) - \hat{\Gamma}_{T,n}(f)\right\|_{L^2(\mathbb{P})} \lesssim T^{1/2}\Delta_n^{3/4},$$

uniformly in K_j, L_j.

The same rate clearly holds up to constants for finite linear combinations of such indicators.

3.2.3. Sobolev functions

Surprisingly, the rate for indicator functions in Theorem 3.6 corresponds to the rate in Theorem 3.5 for 1/2-Hölder functions. In this section we will show that this follows from a more general result for Sobolev functions. The closure of $L^2(\mu) \cap C^1(\mathbb{R}^d)$ with respect to $\|\cdot\|_{H^1(\mu)}$ yields the space $H^1(\mu)$, a μ-weighted Sobolev space. This is not a Banach space in general (Kufner [23]). In order to avoid this issue, assume in the following that μ has a bounded Lebesgue density $d\mu/d\lambda$. Then $L^2(\mathbb{R}^d) := L^2(\mathbb{R}^d, \lambda) \subset L^2(\mathbb{R}^d, \mu)$, where λ is the Lebesgue measure, and

$$\|f\|_{H^1(\mu)} \le \|\frac{d\mu}{d\lambda}\|_\infty \|f\|_{H^1}, \quad f \in L^2(\mathbb{R}^d) \cap C^1(\mathbb{R}^d).$$

Here, $\|\cdot\|_\infty := \|\cdot\|_{\infty,\lambda}$ is the λ-sup-norm and $\|f\|_{H^1} := \|f\|_{H^1(\lambda)}$ is the classical Sobolev norm with respect to λ. Taking the closure of $L^2(\mu) \cap C^2(\mathbb{R}^d)$ with respect to $\|\cdot\|_{H^1}$ leads to the Sobolev space $H^1(\mathbb{R}^d)$ of weakly differentiable functions with square integrable partial derivatives. This yields, instead of (9),

$$\begin{aligned}&\left\|\Gamma_T(f) - \hat{\Gamma}_{T,n}(f)\right\|_{L^2(\mathbb{P})} \\ &\le \begin{cases} C \max\left(1, \|\,|A|\,\|_{\infty,\mu}^{1/2}\right) \|\frac{d\mu}{d\lambda}\|_\infty^{1/2} \|f\|_{H^1} T^{1/2}\Delta_n, & f \in H^1(\mathbb{R}^d), \\ C\|\frac{d\mu}{d\lambda}\|_\infty^{1/2}\|f\|_\lambda T^{1/2}\Delta_n^{1/2}, & f \in L^2(\mathbb{R}^d). \end{cases}\end{aligned} \tag{13}$$

The operator norm of $\Gamma_T - \hat{\Gamma}_{T,n}$ as operator from $H^1(\mathbb{R}^d)$ to $L^2(\mathbb{P})$ is therefore bounded by $C\max(1, \|\,|A|\,\|_{\infty,\mu}^{1/2})\|\frac{d\mu}{d\lambda}\|_\infty^{1/2} T^{1/2}\Delta_n$. Similarly, it is bounded by $C\|\frac{d\mu}{d\lambda}\|_\infty^{1/2} T^{1/2}\Delta_n^{1/2}$ when considering $\Gamma_T - \hat{\Gamma}_{T,n}$ as operator from $L^2(\mathbb{R}^d)$ to $L^2(\mathbb{P})$. We want to derive an upper bound on the operator norm of $\Gamma_T - \hat{\Gamma}_{T,n}$ for functions that are "between" $H^1(\mathbb{R}^d)$ and $L^2(\mathbb{R}^d)$. Interpolation between these two spaces leads to the fractional Sobolev spaces $H^s(\mathbb{R}^d)$, $0 \le s \le 1$, with finite norm

$$\|f\|_{H^s} = \left(\frac{1}{(2\pi)^{d/2}}\int_{\mathbb{R}^d}(1+|u|)^{2s}|\mathcal{F}f(u)|^2 du\right)^{1/2},$$

where $\mathcal{F}f$ is the $L^2(\mathbb{R}^d)$-Fourier transform of f (cf. Adams and Fournier [1, Section 7.50ff]). The exact interpolation theorem (Adams and Fournier [1, Section 7.2]) applied to $\Gamma_T - \hat{\Gamma}_{T,n}$ yields an upper bound on its operator norm when considered as operator from $H^s(\mathbb{R}^d)$ to $L^2(\mathbb{P})$ by interpolating the operator norms from above. This gives the following result.

Theorem 3.7. *Let X be a stationary diffusion with values in $\mathbb{R}^d$ and stationary measure μ, whose generator L satisfies (1) and Assumption 3.4. Assume that μ has bounded Lebesgue density. There exists a constant $C < \infty$ such that for any $f \in H^s(\mathbb{R}^d)$, $0 \le s \le 1$,*

$$\left\|\Gamma_T(f) - \hat{\Gamma}_{T,n}(f)\right\|_{L^2(\mathbb{P})} \le C \max\left(1, \|\,|A|\,\|_{\infty,\mu}^{1/2}\right) \|\frac{d\mu}{d\lambda}\|_{\infty}^{1/2} \|f\|_{H^s} T^{1/2} \Delta_n^{\frac{1+s}{2}}.$$

If f is an α-Hölder continuous function with $0 < \alpha \le 1$ and compact support in $\mathbb{R}^d$, then it is well-known that $f \in H^{\alpha-}(\mathbb{R}^d)$. The theorem thus implies the rate $T^{1/2}\Delta_n^{(1+\alpha)/2-}$. Using large deviation bounds for X, this extends to general α-Hölder continuous f, losing a $\log n$-factor in the rate. For the one-dimensional indicator function $f = \mathbf{1}_{[K,L)}$, $-\infty < K < L < \infty$, (12) remains true, but it also follows that $f \in H^{1/2-}(\mathbb{R})$. The theorem above therefore yields up to a constant the bound $\|f\|_{H^{1/2-\varepsilon}} T^{1/2}\Delta_n^{3/4-\varepsilon}$ for any small $\varepsilon > 0$. This is not uniform in K, L anymore, but still describes well how the error depends on f. Compared to Theorems 3.5 and 3.6 the explicit approximation via (9) yields in general sharper bounds than Theorem 3.7, but the latter one is easier to apply since it only requires bounding the Fourier transform of f.

Remark 3.8.

(i) Assume that $H^1(\mu)$ is a Banach space. This is true, for instance, in Examples 3.1 and 3.2, when $m(x)$ is uniformly bounded from above and below. In that case interpolation (in the sense of Adams and Fournier [1, Chapter 7]) between $H^1(\mu)$ and $L^2(\mu)$ is possible and yields a similar bound as Theorem 3.7, but with $\|\cdot\|_{H^s}$ replaced by an appropriate interpolation norm. The results in Theorems 3.5 and 3.6 are explicit cases of this. Up to boundary conditions this implies that α-Hölder functions lie in $\mathcal{D}((-L)^{\alpha/2})$, $0 \le \alpha \le 1$, and indicator functions $f = \mathbf{1}_{[K,L)}$ lie in $\mathcal{D}((-L)^{1/4})$, $-\infty < K < L \le \infty$.

(ii) Depending on the boundary conditions for $f \in \mathcal{D}(L)$ and assuming that μ has bounded Lebesgue density, it can be shown in many examples that $H^1(\mathbb{R}^d)$ embeds continuously into $\mathcal{D}((I-L)^{1/2}) \subset \mathcal{D}((-L)^{1/2})$. This holds, for instance, for the Ornstein–Uhlenbeck process and for the reflected diffusions in Example 3.3. Since $L^2(\mathbb{R}^d) \subset \mathcal{D}((-L)^0) = L^2(\mathbb{R}^d, \mu)$, interpolation implies that $H^s(\mathbb{R}^d)$ embeds continuously into $\mathcal{D}((I-L)^{s/2}) \subset \mathcal{D}((-L)^{s/2})$. In particular, the indicator functions $f = \mathbf{1}_{[K,L)}$ lie in $\mathcal{D}((-L)^{1/4-\varepsilon})$ for any small $\varepsilon > 0$.

(iii) Arguing like in the proof of Corollary 2.3 the strict stationarity assumption can be relaxed. This is also true for Theorems 3.5 and 3.6.

(iv) The theorem applies to some extent also to $f \in H^s(\mathbb{R}^d)$ with $s < 0$. For example, let $d = 1$ and assume that the quadratic variation of X is $\langle X \rangle_t = 1$ on $[0, T]$ (consider for instance Example 3.3 with $\sigma = 1$). Let $f = \delta_a$ be the Dirac delta function in $a \in \mathbb{R}$. Formally, it follows that $\Gamma_T(f)$ is equal to L_T^a, the local time of X in a until T. Note that $\delta_a \in H^{-1/2-}(\mathbb{R})$ has negative regularity, i.e. it is a distribution in the sense of Schwartz.

Please cite this article in press as: R. Altmeyer, J. Chorowski, Estimation error for occupation time functionals of stationary Markov processes, Stochastic Processes and their Applications (2017), https://doi.org/10.1016/j.spa.2017.08.013.

From the theorem it therefore can be expected that the rate of convergence should be $\Delta_n^{1/4-}$ as $n \to \infty$. Indeed, using the theorem it can be shown that $\|L_T^a - \hat{\Gamma}_{T,n}(f_{a,\varepsilon})\|_{L^2(\mathbb{P})} \lesssim \Delta_n^{1/4-\rho}$ for any small $\rho > 0$, where $f_{a,\varepsilon}(x) = (2\varepsilon)^{-1}\mathbf{1}_{(a-\varepsilon,a+\varepsilon)}(x)$ (cf. Kohatsu-Higa et al. [22, Theorem 2.6] and Altmeyer [2, Section 3.3]).

3.3. Brownian motion and related diffusions

In some situations the procedure of the last subsection can be adapted to non-stationary diffusions, as long as the generator satisfies (1). As an important example for this assume that $(B_r)_{0\leq r\leq T}$ is a one-dimensional Brownian motion. Its generator is $Lf = (1/2)\Delta f$ with domain $\mathcal{D}(L) = H^2(\mathbb{R})$ and (1) holds with $A = 1/2$ and μ being the Lebesgue measure on $\mathbb{R}$. Theorem 3.7 then extends to $(B_r)_{0\leq r\leq T}$, assuming that the initial distribution is absolutely continuous with compact support. The key idea is to approximate the process by stationary diffusions with reflecting boundaries.

Theorem 3.9. *Assume that $(X_r)_{r\geq 0}$ satisfies $X_r = X_0 + B_r$, where $(B_r)_{r\geq 0}$ is a one-dimensional Brownian motion, $X_0 \overset{d}{\sim} \eta$ is independent of $(B_r)_{r\geq 0}$ and η has Lebesgue density with compact support. There exists a constant $C < \infty$ such that for $f \in H^s(\mathbb{R})$, $0 \leq s \leq 1$,*

$$\left\| \Gamma_T(f) - \hat{\Gamma}_{T,n}(f) \right\|_{L^2(\mathbb{P})} \leq C \| \frac{d\eta}{d\lambda} \|_\infty^{1/2} \| f \|_{H^s} T^{1/2} \Delta_n^{\frac{1+s}{2}}.$$

Following a similar program as in Ngo and Ogawa [24] or Kohatsu-Higa et al. [22] this result can be extended to more general one-dimensional diffusion processes, using the Lamperti and Girsanov transforms.
For different proofs see Altmeyer [2, Theorem 3.6] or Altmeyer [2, Theorem 3.8]

3.4. Infinite dimensional diffusions

Since the general state space $\mathcal{S}$ of X is only assumed to be Polish, it is also possible to study infinite dimensional diffusions. The results of Ganychenko et al. [15], Ganychenko [14] do not apply then, because, in general, heat kernel bounds are not available in this setting. Example 3.1 can be generalized considerably. If X satisfies the stochastic differential equation

$$dX_r = AX_r dr + Q^{1/2} dW_r,$$

where A and Q are operators on a separable Hilbert space $\mathcal{H}$, with Q being bounded self-adjoint, then X is a Gaussian Markov process and the generator L satisfies a similar formula as in (4) with ∇ and Δ replaced by the corresponding Fréchet derivatives D and D^2. Under certain conditions on A and Q the generator is reversible and X has a stationary measure μ. The domain is again a μ-weighted Sobolev space and the associated Dirichlet form is

$$\langle -Lf, g\rangle_\mu = \frac{1}{2}\int_{\mathcal{H}} \langle Q^{1/2} Df(x), Q^{1/2} Dg(x)\rangle_{\mathcal{H}} \, d\mu(x).$$

The results of Section 3.2 therefore remain formally the same. For details see Chojnowska-Michalik and Goldys [8]. For a different kind of example consider an infinite dimensional system of the form

$$dX_r^{(i)} = \left(pV'\left(X_r^{(i+1)} - X_r^{(i)}\right) - qV'\left(X_r^{(i)} - X_r^{(i-1)}\right)\right) dr + dW_r^{(i)},$$

Please cite this article in press as: R. Altmeyer, J. Chorowski, Estimation error for occupation time functionals of stationary Markov processes, Stochastic Processes and their Applications (2017), https://doi.org/10.1016/j.spa.2017.08.013.

where $(r, i) \in [0, \infty) \times \mathbb{Z}$, $p, q \geq 0$ with $p = (1 + \sqrt{\varepsilon})/2$, $q = (1 - \sqrt{\varepsilon})/2$, $\varepsilon > 0$, and where $\{(W_r^{(i)})_{r\geq 0} : i \in \mathbb{Z}\}$ is an independent family of Brownian motions. Moreover, V is some potential function (Diehl et al. [10]). $X = (X_r^{(i)})_{r\geq 0, i\in\mathbb{Z}}$ is stationary and the infinitesimal generator $L^{(\varepsilon)} = L_S + \sqrt{\varepsilon} L_A$ can be studied via its symmetric and antisymmetric parts L_S and L_A. The Dirichlet form for the symmetric part is given in Lemma 2.1 of Diehl et al. [10]. If $\varepsilon = 0$, then the generator is symmetric and a similar analysis as in Section 3.2 can be applied.

4. Proofs

4.1. *Proof of Theorem 2.2*

Proof. Assume first that $f \in L^2(\mu)$. Expanding the squared error yields

$$\| \Gamma_T(f) - \hat{\Gamma}_{T,n}(f) \|^2_{L^2(\mathbb{P})} = \mathbb{E}\Big[\Big| \sum_{k=1}^{n} \int_{t_{k-1}}^{t_k} \big(f(X_r) - f\big(X_{t_{k-1}}\big)\big) dr \Big|^2 \Big]$$

$$= \sum_{k,l=1}^{n} \int_{t_{k-1}}^{t_k} \int_{t_{l-1}}^{t_l} \mathbb{E}\big[\big(f(X_r) - f\big(X_{t_{k-1}}\big)\big)\big(f(X_h) - f\big(X_{t_{l-1}}\big)\big)\big] dr dh.$$

We bound the diagonal ($l = k$)) and off-diagonal ($l \neq k$) terms separately. Consider first the diagonal case and $t_{k-1} \leq r \leq h \leq t_k$. By the Markov property and stationarity of X the expectation above can be calculated explicitly. Indeed,

$$\begin{aligned}
&\mathbb{E}\big[\big(f(X_r) - f\big(X_{t_{k-1}}\big)\big)\big(f(X_h) - f\big(X_{t_{k-1}}\big)\big)\big] \\
&\qquad = \langle P_{h-r} f, f \rangle_\mu - \big\langle P_{r-t_{k-1}} f, f \big\rangle_\mu - \big\langle P_{h-t_{k-1}} f, f \big\rangle_\mu + \langle f, f \rangle_\mu \\
&\qquad = \big\langle (P_{h-r} - I) f + \big(I - P_{r-t_{k-1}}\big) f + \big(I - P_{h-t_{k-1}}\big) f, f \big\rangle_\mu .
\end{aligned}$$

Consequently, by symmetry in r, h,

$$\begin{aligned}
&\sum_{k=1}^{n} \int_{t_{k-1}}^{t_k} \int_{t_{k-1}}^{t_k} \mathbb{E}\big[\big(f(X_r) - f\big(X_{t_{k-1}}\big)\big)\big(f(X_h) - f\big(X_{t_{k-1}}\big)\big)\big] dr dh \\
&= 2 \sum_{k=1}^{n} \Big\langle \Big(\int_{t_{k-1}}^{t_k} \int_{t_{k-1}}^{h} (P_{h-r} - I) \, dr dh + \Delta_n \int_{t_{k-1}}^{t_k} \big(I - P_{h-t_{k-1}}\big) dh \Big) f, f \Big\rangle_\mu \\
&= 2n \Big\langle \Big(\int_0^{\Delta_n} \int_0^h (P_{h-r} - I) \, dr dh + \Delta_n \int_0^{\Delta_n} (I - P_h) \, dh \Big) f, f \Big\rangle_\mu .
\end{aligned}$$

Since the generator L is normal, by the functional calculus of L (see Section 4.4) this can be written as

$$\langle \Psi(L) f, f \rangle_\mu = \int_{\sigma(L)} \Psi(\lambda) \, d\langle E_\lambda f, f \rangle_\mu$$

with

$$\Psi(\lambda) = 2n \left(\int_0^{\Delta_n} \int_0^h \big(e^{\lambda(h-r)} - 1\big) dr dh + \Delta_n \int_0^{\Delta_n} \big(1 - e^{\lambda h}\big) dh \right), \quad \lambda \in \mathbb{C}.$$

Since L is the generator of a Feller semigroup, it follows that $\sigma(L) \subset \{\lambda \in \mathbb{C} : Re(\lambda) \leq 0\}$. Fix $0 \leq s \leq 1$ such that $|1 - e^z| \leq 2|z|^s$ for $z \in \{\lambda \in \mathbb{C} : Re(\lambda) \leq 0\}$. Then $|\Psi(\lambda)| \leq 8n\Delta_n^{2+s}|\lambda|^s$,

$\lambda \in \sigma(L)$. Hence the diagonal terms are bounded by

$$\int_{\sigma(L)} |\Psi(\lambda)|\, d\langle E_\lambda f, f\rangle_\mu \le 8T\Delta_n^{1+s} \int_{\sigma(L)} |\lambda|^s d\langle E_\lambda f, f\rangle_\mu \le 8\||L|^{s/2} f\|_\mu^2 T\Delta_n^{1+s}, \tag{14}$$

which is true as long as $f \in \mathcal{D}(|L|^{s/2})$. For the off-diagonal terms with $l \neq k$ consider $t_{l-1} \le r \le t_{k-1} \le h$. Then, similar as before

$$\begin{aligned} &\mathbb{E}\left[\left(f(X_h) - f\left(X_{t_{k-1}}\right)\right)\left(f(X_r) - f\left(X_{t_{l-1}}\right)\right)\right] \\ &\quad = \langle P_{h-r} f, f\rangle_\mu - \left\langle P_{h-t_{l-1}} f, f\right\rangle_\mu - \left\langle P_{t_{k-1}-r} f, f\right\rangle_\mu + \left\langle P_{t_{k-1}-t_{l-1}} f, f\right\rangle_\mu \\ &\quad = \left\langle P_{t_{k-1}-r}(P_{h-t_{k-1}} - I)(I - P_{r-t_{l-1}}) f, f\right\rangle_\mu. \end{aligned} \tag{15}$$

The off-diagonal terms are therefore equal to

$$\begin{aligned} &2\sum_{k>l=1}^{n} \int_{t_{k-1}}^{t_k} \int_{t_{l-1}}^{t_l} \mathbb{E}\left[\left(f(X_r) - f\left(X_{t_{k-1}}\right)\right)\left(f(X_h) - f\left(X_{t_{l-1}}\right)\right)\right] dr\,dh \\ &\quad = 2\sum_{k>l=1}^{n} \left\langle \left(\left(\int_{t_{k-1}}^{t_k} \int_{t_{l-1}}^{t_l} P_{t_{k-1}-r}\left(P_{h-t_{k-1}} - I\right)\left(I - P_{r-t_{l-1}}\right) dr\,dh\right) f, f\right\rangle_\mu \\ &\quad = \left\langle 2\left(\int_0^{\Delta_n} \int_0^{\Delta_n} \left(\sum_{k>l=1}^{n} P_{t_{k-1}-t_{l-1}-r}\right)\left(P_h - I\right)(I - P_r)\, dr\,dh\right) f, f\right\rangle_\mu \\ &\quad = \int_{\sigma(L)} \tilde{\Psi}(\lambda)\, d\langle E_\lambda f, f\rangle_\mu, \end{aligned}$$

where

$$\tilde{\Psi}(\lambda) = 2\int_0^{\Delta_n} \int_0^{\Delta_n} \left(\sum_{k>l=1}^{n} e^{\lambda\left(t_{k-1}-t_{l-1}-r\right)}\right)\left(e^{\lambda h} - 1\right)\left(1 - e^{\lambda r}\right) dr\,dh, \quad \lambda \in \mathbb{C}.$$

We will show that there exists a universal constant $\tilde{C} < \infty$ such that

$$\left|\tilde{\Psi}(\lambda)\right| \le \tilde{C}T|\lambda|^s \Delta_n^{1+s}, \quad \lambda \in \sigma(L). \tag{16}$$

As in (14), this implies that the off-diagonal terms are bounded by

$$\tilde{C}T\Delta_n^{1+s} \int_{\sigma(L)} |\lambda|^s d\langle E_\lambda f, f\rangle_\mu = \tilde{C}\||L|^{s/2} f\|_\mu^2 T\Delta_n^{1+s}$$

for $f \in \mathcal{D}(|L|^{s/2})$. Combining this with (14) yields the claim. In order to show (16) observe that $\tilde{\Psi}(\lambda) = 0$ for $\lambda = 0$. It is therefore sufficient to consider $\lambda \neq 0$. In order to bound $\tilde{\Psi}$ in that case note that

$$\sum_{k>l=1}^{n} e^{\lambda(k-l-1)\Delta_n} = \sum_{l=1}^{n} \frac{1 - e^{\lambda(n-l)\Delta_n}}{1 - e^{\lambda\Delta_n}} = \frac{n}{1 - e^{\lambda\Delta_n}} - \frac{1 - e^{\lambda n\Delta_n}}{\left(1 - e^{\lambda\Delta_n}\right)^2}.$$

Hence, again using $|1 - e^z| \le 2|z|^s$,

$$\left|\Delta_n^2\left(1 - e^{\lambda\Delta_n}\right)^2 \sum_{k>l=1}^{n} e^{\lambda(k-l-1)\Delta_n}\right| \le 2n\Delta_n^2|\lambda\Delta_n|^s + 2\Delta_n^2|\lambda n\Delta_n|^s \le 4T\Delta_n^{1+s}|\lambda|^s.$$

Therefore, (16) follows if

$$\left|\Delta_n^{-2}\left(1-e^{\lambda\Delta_n}\right)^{-2}\int_0^{\Delta_n}\int_0^{\Delta_n} e^{\lambda(\Delta_n-r)}\left(e^{\lambda h}-1\right)\left(1-e^{\lambda r}\right)drdh\right| \tag{17}$$

is bounded by a universal constant. To show this, let $z=\lambda\Delta_n$ and note that

$$\Delta_n^{-1}\left|\frac{\int_0^{\Delta_n}\left(1-e^{\lambda h}\right)dh}{1-e^{\lambda\Delta_n}}\right| = \Delta_n^{-1}\left|\frac{\Delta_n}{1-e^{\lambda\Delta_n}}-\lambda^{-1}\right| = \left|\frac{1}{z}-\frac{1}{e^z-1}\right|, \tag{18}$$

$$\Delta_n^{-1}\left|\int_0^{\Delta_n}\frac{\left(e^{\lambda(\Delta_n-r)}-e^{\lambda\Delta_n}\right)}{1-e^{\lambda\Delta_n}}dr\right| = \left|-\frac{1}{z}-\frac{e^z}{1-e^z}\right| = \left|\frac{1}{z}-\frac{1}{e^z-1}+1\right|. \tag{19}$$

(18) converges to 1 and (19) converges to 0 as $|z|\to\infty$. If $|z|\to 0$ and $z\in\{\lambda\in\mathbb{C}:\mathrm{Re}(\lambda)\le 0\}$, then (18) converges to $1/2$ and (19) converges to $3/2$. This implies a universal constant bounding (17), thereby proving (16). □

Remark 4.1.

(i) If the generator is self-adjoint, then the operators P_u, $u\ge 0$, are self-adjoint as well. P_u is positive, P_u-I is negative semidefinite and $I-P_u$ is positive semidefinite. Therefore $P_{t_{k-1}-r}(P_{h-t_{k-1}}-I)(I-P_{r-t_{l-1}})$ is negative semidefinite and (15) is non-positive. In this case, the off-diagonal terms do not contribute to the estimation error.

(ii) The restriction $0\le s\le 1$ appears in (14) and (16) due to the Lipschitz bound $|1-e^z|\le 2|z|^s$.

4.2. *Proof of Corollary 2.3*

Proof. Denote by $\mathbb{E}_\nu$ for a measure ν the expectation with respect to X starting with ν as initial distribution. Let $g(x)=\mathbb{E}[|\Gamma_{T_n,T}(f)-\hat{\Gamma}_{T_n,T,n}(f)|^2|X_{T_n}=x]$. By conditioning on X_{T_0} the tower and Markov properties yield that

$$\begin{aligned}\left\|\Gamma_{T_n,T}(f)-\hat{\Gamma}_{T_n,T,n}(f)\right\|_{L^2(\mathbb{P})}^2 &= \mathbb{E}\left[\mathbb{E}\left[g\left(X_{T_n}\right)\middle| X_{T_0}\right]\right]\\ &= \mathbb{E}\left[P_{T_n-T_0}g\left(X_{T_0}\right)\right] = \int_S P_{T_n-T_0}g\left(x\right)d\eta(x)\\ &\le \left\|\frac{d\eta}{d\mu}\right\|_{\infty,\mu}\int_S P_{T_n-T_0}g\left(x\right)d\mu(x).\end{aligned}$$

Since μ is the invariant (stationary) measure of the semigroup, this is equal to (cf. Bakry et al. [3, Section 1.2.1])

$$\left\|\frac{d\eta}{d\mu}\right\|_{\infty,\mu}\int_S g\left(x\right)d\mu(x) = \left\|\frac{d\eta}{d\mu}\right\|_{\infty,\mu}\mathbb{E}_\mu\left[\left|\Gamma_{T-T_n}\left(f\right)-\hat{\Gamma}_{T-T_n,n}\left(f\right)\right|^2\right],$$

because

$$\begin{aligned}g(x) &= \mathbb{E}\left[\left|\int_0^{T-T_n}f\left(X_r\right)dr-\Delta_n\sum_{k=1}^{n-T_n}f\left(X_{t_k}\right)\right|^2\middle| X_0=x\right]\\ &= \mathbb{E}\left[\left|\Gamma_{T-T_n}\left(f\right)-\hat{\Gamma}_{T-T_n,n}\left(f\right)\right|^2\middle| X_0=x\right].\end{aligned}$$

Please cite this article in press as: R. Altmeyer, J. Chorowski, Estimation error for occupation time functionals of stationary Markov processes, Stochastic Processes and their Applications (2017), https://doi.org/10.1016/j.spa.2017.08.013.

The conclusion follows by a simple modification of Theorem 2.2, because the error is now considered on $[0, T - T_n]$ instead of $[0, T]$. □

4.3. Proof of Theorem 2.4

Proof. By the triangle inequality it follows from $f \in L^2(\mu) = \mathcal{D}(|L|^0)$ and Theorem 2.2 that

$$\begin{aligned}
&\left\| T^{-1} \hat{\Gamma}_{T,n}(f) - \int_{\mathcal{S}} f(x)\, d\mu(x) \right\|_{L^2(\mathbb{P})} \\
&\quad \leq T^{-1} \left\| \hat{\Gamma}_{T,n}(f) - \Gamma_T(f) \right\|_{L^2(\mathbb{P})} + \left\| T^{-1} \Gamma_T(f) - \int_{\mathcal{S}} f(x)\, d\mu(x) \right\|_{L^2(\mathbb{P})} \\
&\quad \leq C \|f\|_{L^2(\mu)} T^{-1/2} \Delta_n^{1/2} + \left\| T^{-1} \Gamma_T(f) - \int_{\mathcal{S}} f(x)\, d\mu(x) \right\|_{L^2(\mathbb{P})}
\end{aligned}$$

for a universal constant C. The claimed bound for the second term is well-known, but we give the proof here to complement the proof of Theorem 2.2. Consider f such that $f_0 = f - \int f d\mu \in \mathcal{D}(|L|^{-1/2})$. By linearity of the occupation time functional it follows that

$$T^{-1}\Gamma_T(f) - \int_{\mathcal{S}} f d\mu = T^{-1}\Gamma_T(f_0).$$

Fubini's theorem yields

$$\begin{aligned}
\mathbb{E}\left[\left| T^{-1} \Gamma_T(f_0) \right|^2 \right] &= T^{-2} \int_0^T \int_0^T \mathbb{E}\left[f_0(X_r) f_0(X_h) \right] dr dh \\
&= 2T^{-2} \int_0^T \int_0^h \langle P_{h-r} f_0, f_0 \rangle_\mu dr dh \\
&= \int_{\sigma(L)} \Psi(\lambda)\, d\langle E_\lambda f_0, f_0 \rangle_\mu,
\end{aligned}$$

where

$$\Psi(\lambda) = 2T^{-2} \int_0^T \int_0^h e^{\lambda(h-r)} dr dh = 2\frac{e^{\lambda T} - 1 - \lambda T}{\lambda^2 T^2} = 2\frac{(\lambda T)^{-1}(e^{\lambda T} - 1) - 1}{\lambda T},$$

and where $\Psi(0) = 1$ by continuous extension. Since $z \to z^{-1}(e^z - 1) - 1$ is bounded on the left half-plane $\{z \in \mathbb{C} : \mathrm{Re}(z) \leq 0\}$, there exists a constant $\tilde{C} < \infty$ such that

$$|\Psi(\lambda)| \leq \tilde{C} T^{-1} |\lambda|^{-1}, \quad \lambda \in \sigma(L).$$

Consequently,

$$\mathbb{E}\left[\left| T^{-1} \Gamma_T(f_0) \right|^2 \right] \leq C T^{-1} \int_{\sigma(L)} |\lambda|^{-1} d\langle E_\lambda f_0, f_0 \rangle_\mu = C T^{-1} \| |L|^{-1/2} f_0 \|_\mu^2. \quad \square$$

4.4. Proof of Theorem 3.9

Proof. The main idea of the proof is to approximate the process X by reflected processes for which Theorem 3.7 can be applied. Choose M big enough such that $\mathrm{supp}(\eta) \subset [-M, M]$ and let $\tau_M = \inf\{r > 0 : |X_r| > M\}$ the first time X exits from $[-M, M]$. By dominated convergence, for any fixed $T > 0$, it holds that

$$\left\| \Gamma_T(f) - \hat{\Gamma}_{T,n}(f) \right\|_{L^2(\mathbb{P})} = \lim_{M \to \infty} \left\| \left(\Gamma_T(f) - \hat{\Gamma}_{T,n}(f) \right) \mathbf{1}(T < \tau_M) \right\|_{L^2(\mathbb{P})}. \tag{20}$$

We will show that there exists a universal constant C, independent of M, such that

$$\left\| \left(\Gamma_T(f) - \hat{\Gamma}_{T,n}(f)\right)\mathbf{1}(t < \tau_M)\right\|_{L^2(\mathbb{P})} \leq C \left\| \frac{d\eta}{d\lambda}\right\|_{\infty,\mu}^{1/2} \|f\|_{H^s} T^{1/2} \Delta_n^{\frac{1+s}{2}}. \tag{21}$$

For this define

$$f_M(x) = \begin{cases} x - 4kM & : (4k-1)M \leq x < (4k+1)M \\ (4k+2)M - x & : (4k+1)M \leq x < (4k+3)M. \end{cases}$$

Applying the Itô-Tanaka formula the process $X_r^{(M)} := f_M(X_r)$ is a reflected Brownian motion with barriers at $-M, M$, i.e. $(X_r^{(M)})_{r\geq 0}$ satisfies (8) with $\beta = -M$, $\rho = M$, $b = 0$, $\sigma = 1$ and some process $(K_r)_{r\geq 0}$. See Gihman and Skorohod [16, Chapter I.23] for a similar construction of reflected diffusion processes. Furthermore, for all $0 \leq r \leq T \leq \tau_M$ it follows that

$$B_r = X_r^{(M)}. \tag{22}$$

In the following, denote by $\Gamma_T^{(M)}$, $\hat{\Gamma}_{T,n}^{(M)}$ the integral functional and the Riemann estimator with respect to the process $(X_r^{(M)})_{r\geq 0}$. (22) implies that

$$\begin{aligned}\left\| \left(\Gamma_T(f) - \hat{\Gamma}_{T,n}(f)\right)\mathbf{1}(T < \tau_M)\right\|_{L^2(\mathbb{P})} &= \left\| \left(\Gamma_T^{(M)} - \hat{\Gamma}_{T,n}^{(M)}(f)\right)\mathbf{1}(T < \tau_M)\right\|_{L^2(\mathbb{P})} \\ &\leq \left\| \Gamma_T^{(M)} - \hat{\Gamma}_{T,n}^{(M)}(f)\right\|_{L^2(\mathbb{P})}.\end{aligned}$$

Since $\mathrm{supp}(\eta) \subset [-M, M]$, it follows that $X_0 = X_0^{(M)}$, i.e. $X_0^{(M)}$ has distribution η. η is in general *not* the stationary distribution of $(X_r^{(M)})_{r\geq 0}$. From Example 3.3 we actually know that the stationary distribution μ_M of the reflected Brownian motion on $[-M, M]$ has Lebesgue density $d\mu_M/d\lambda = (2M)^{-1}\mathbf{1}_{[-M,M]}$. Theorem 3.7, with $A = 1$, therefore implies together with the argument from the proof of Corollary 2.3 that

$$\left\| \Gamma_T^{(M)} - \hat{\Gamma}_{T,n}^{(M)}(f)\right\|_{L^2(\mathbb{P})} \leq C \left\| \frac{d\mu_M}{d\lambda}\right\|_{\infty,\mu}^{1/2} \left\| \frac{d\eta}{d\mu_M}\right\|_{\infty,\eta}^{1/2} \|f\|_{H^s} T^{1/2} \Delta_n^{\frac{1+s}{2}}.$$

Observe that

$$\int g d\eta = \int_{-M}^{M} g \frac{d\eta}{d\lambda} d\lambda = 2M \int g \frac{d\eta}{d\lambda} d\mu_M$$

for any bounded continuous function g, i.e. $d\eta/d\mu_M = 2M\, d\eta/d\lambda$. In all, this means that

$$\left\| \Gamma_T^{(M)} - \hat{\Gamma}_{T,n}^{(M)}(f)\right\|_{L^2(\mathbb{P})} \leq C \left\| \frac{d\eta}{d\lambda}\right\|_{\infty,\eta}^{1/2} \|f\|_{H^s} T^{1/2} \Delta_n^{\frac{1+s}{2}},$$

which is (21). □

Acknowledgment

Both authors gratefully acknowledge the financial support of the DFG Research Training Group 1845 "Stochastic Analysis with Applications in Biology, Finance and Physics".

Appendix. A short review of semigroup theory and functional calculus

We briefly recall the basic objects needed in the theory of semigroups and the functional calculus for normal operators. For more details see Bakry et al. [3], in particular Chapters 1.4.1 and A.4, Rudin [27, Chapter 13] or Engel and Nagel [12]. Let μ be any probability measure

Please cite this article in press as: R. Altmeyer, J. Chorowski, Estimation error for occupation time functionals of stationary Markov processes, Stochastic Processes and their Applications (2017), https://doi.org/10.1016/j.spa.2017.08.013.

on $(\mathcal{S}, \mathcal{B}(\mathcal{S}))$. On the induced Hilbert space $(L^2(\mu), \|\cdot\|_\mu)$ denote by $(P_r)_{r\geq 0}$ the Markov semigroup associated with the Markov process X which satisfies $P_r f(x) = \mathbb{E}[f(X_r)|X_0 = x]$ for $f \in L^2(\mu)$, $x \in \mathcal{S}$, and $P_{r+s} = P_r P_s$, $r, s \geq 0$. The *infinitesimal generator* of the semigroup is defined as

$$Lf = \lim_{r\to 0} \frac{P_r f - f}{r}, \quad f \in \mathcal{D}(L),$$

where the limit is taken with respect to $\|\cdot\|_\mu$ and where the *domain* $\mathcal{D}(L) \subset L^2(\mu)$ is the set of all functions f for which this limit exists. If $(P_r)_{r\geq 0}$ is strongly continuous, i.e. $P_r f \to f$ in $L^2(\mu)$ as $r \to 0$ for all $f \in L^2(\mu)$, then the semigroup is called *Feller*. This is true for most Markov processes in practice, including Lévy processes and many diffusions. In the Feller case, L is a densely defined closed linear and usually unbounded operator on its domain with spectrum $\sigma(L) \subset \{\lambda \in \mathbb{C} : \mathrm{Re}(\lambda) \leq 0\}$. In order to define fractional powers of the generator assume that the operator L (and thus the operators P_r) are *normal*, i.e. $LL^* = L^*L$, where L^* is the Hilbert space adjoint of L with respect to $L^2(\mu)$. In that case, the spectral theorem (Theorem 13.33 of Rudin [27]) guarantees the existence of a resolution of the identity or spectral measure $(E_A)_{A\in\mathcal{B}(\mathbb{C})}$ on $L^2(\mu)$. This means that $(E_A)_{A\in\mathcal{B}(\mathbb{C})}$ is a family of orthogonal projections $E_A : L^2(\mu) \to L^2(\mu)$ for Borel sets $A \subset \mathbb{C}$ such that for every $f, g \in L^2(\mu)$ the map $A \mapsto \langle E_A f, g\rangle_\mu$ is a complex measure supported on $\sigma(L)$. Moreover, $A \mapsto \langle E_A f, f\rangle_\mu$ is a positive measure with total variation $\langle E_{\mathbb{C}} f, f\rangle_\mu = \|f\|_\mu^2$. By the spectral theorem we can associate to any measurable function $\Psi : \mathbb{C} \mapsto \mathbb{C}$ a densely defined closed operator $\Psi(L)$ by the relation

$$\langle \Psi(L) f, g\rangle_\mu = \int_{\sigma(L)} \Psi(\lambda)\, d\langle E_\lambda f, g\rangle_\mu, \quad f, g \in L^2(\mu),$$

with domain $\mathcal{D}(\Psi(L)) = \{f \in L^2(\mu) : \int_{\sigma(L)} |\Psi(\lambda)|^2 d\langle E_\lambda f, f\rangle_\mu < \infty\}$. It satisfies $\|\Psi(L)f\|_\mu^2 = \int_{\sigma(L)} |\Psi(\lambda)|^2 d\langle E_\lambda f, f\rangle_\mu$. In particular, we can define the fractional operators $|L|^{s/2}$ on $\mathcal{D}(|L|^{s/2})$ for $0 \leq s \leq 1$. At last, by the spectral theorem for normal semigroups (Theorem 13.38 of Rudin [27]), the semigroup can be realized in its usual exponential form, i.e. $P_r = \Psi(L)$ with $\Psi(x) = e^{rx}$, $r \geq 0$.

References

[1] R.A. Adams, J.J.F. Fournier, Sobolev Spaces, in: Pure and Applied Mathematics, Elsevier Science, 2003.

[2] R. Altmeyer, Estimating occupation time functionals, 2017. ArXiv Preprint ArXiv:1706.03418.

[3] D. Bakry, I. Gentil, M. Ledoux, Analysis and Geometry of Markov Diffusion Operators, in: Grundlehren der mathematischen Wissenschaften, Springer International Publishing, 2013.

[4] R.F. Bass, Diffusions and Elliptic Operators, in: Probability and Its Applications, Springer New York, 2006.

[5] R. Catellier, M. Gubinelli, Averaging along irregular curves and regularisation of ODEs, Stochastic Process. Appl. 126 (8) (2016) 2323–2366.

[6] M.F. Chen, Eigenvalues, Inequalities, and Ergodic Theory, in: Probability and Its Applications, Springer London, 2006.

[7] M. Chesney, M. Jeanblanc-Picqué, M. Yor, Brownian excursions and Parisian barrier options, Adv. Appl. Probab. 29 (29) (1997).

[8] A. Chojnowska-Michalik, B. Goldys, Symmetric Ornstein-Uhlenbeck semigroups and their generators, Probab. Theory Related Fields 124 (4) (2002) 459–486.

[9] J. Chorowski, Nonparametric volatility estimation in scalar diffusions: Optimality across observation frequencies, 2015. ArXiv Preprint ArXiv:1507.07139, http://arxiv.org/abs/1507.07139 [arXiv:1507.07139], http://arxiv.org/abs/1507.07139.

[10] J. Diehl, M. Gubinelli, N. Perkowski, The Kardar–Parisi–Zhang Equation as Scaling Limit of Weakly Asymmetric Interacting Brownian Motions, Comm. Math. Phys. (2017) 1–41.

[11] C. Dion, V. Genon-Catalot, Bidimensional random effect estimation in mixed stochastic differential model, Stat. Inference Stoch. Process. 19 (2) (2016) 131–158.

[12] K.-J. Engel, R. Nagel, One-Parameter Semigroups for Linear Evolution Equations, in: Graduate Texts in Mathematics, Springer New York, 1999.

[13] S.N. Ethier, T.G. Kurtz, Markov Processes : Characterization and Convergence, Wiley, 1986, p. 534.

[14] I. Ganychenko, Fast L 2 -approximation of integral-type functionals of Markov processes, Modern Stochastics: Theory and Applications 2 (2015) 165–171.

[15] I. Ganychenko, V. Knopova, A. Kulik, Accuracy of discrete approximation for integral functionals of Markov processes, Mod. Stoch.: Theory Appl. 2 (4) (2015) 401–420.

[16] I. Gihman, A. Skorohod, Stochastic Differential Equations, Springer, 2015.

[17] E. Gobet, G. Matulewicz, Parameter estimation of Ornstein–Uhlenbeck process generating a stochastic graph, Stat. Inference Stoch. Process. (2016) 1–25.

[18] L.P. Hansen, J. Alexandre Scheinkman, N. Touzi, Spectral methods for identifying scalar diffusions, J. Econometrics 86 (1) (1998) 1–32.

[19] M. Hoffmann, Lp estimation of the diffusion coefficient, Bernoulli 5 (3) (1999) 447.

[20] J.-N. Hugonnier, The Feynman–Kac formula and pricing occupation time derivatives, Int. J. Theor. Appl. Finance 2 (02) (1999) 153–178.

[21] C. Kipnis, S.R.S. Varadhan, Central limit theorem for additive functionals of reversible Markov processes and applications to simple exclusions, Comm. Math. Phys. 104 (1) (1986) 1–19.

[22] A. Kohatsu-Higa, R. Makhlouf, H.-L. Ngo, Approximations of non-smooth integral type functionals of one dimensional diffusion processes, Stochastic Process. Appl. 124 (5) (2014) 1881–1909.

[23] A. Kufner, Weighted Sobolev Spaces, in: Teubner-Texte zur Mathematik, B.G. Teubner, 1985.

[24] H.-L. Ngo, S. Ogawa, On the discrete approximation of occupation time of diffusion processes, Electron. J. Stat. 5 (2011) 1374–1393.

[25] G.A. Pavliotis, Stochastic Processes and Applications: Diffusion Processes, the Fokker-Planck and Langevin Equations, in: Texts in Applied Mathematics, Springer New York, 2014.

[26] P. Pollett, Integrals for continuous-time Markov chains, Math. Biosci. 182 (2) (2003) 213–225.

[27] W. Rudin, Functional Analysis, in: International series in pure and applied mathematics, McGraw-Hill, 2006.

[28] S. Watanabe, Lectures on Stochastic Differential Equations and Malliavin Calculus, in: Lectures on Mathematics and Physics / Tata Institute of Fundamental Research: Mathematics, Springer, 1984.

Please cite this article in press as: R. Altmeyer, J. Chorowski, Estimation error for occupation time functionals of stationary Markov processes, Stochastic Processes and their Applications (2017), https://doi.org/10.1016/j.spa.2017.08.013.