Electronic Journal of Statistics
Vol. 13 (2019) 1329–1358
ISSN: 1935-7524
https://doi.org/10.1214/19-EJS1545

The nonparametric LAN expansion for discretely observed diffusions

Sven Wang

Centre for Mathematical Sciences, Wilberforce Road
Cambridge CB3 0WB, United Kingdom
e-mail: sven.wang@statslab.cam.ac.uk

Abstract: Consider a scalar reflected diffusion $(X_t : t \geq 0)$, where the unknown drift function b is modelled nonparametrically. We show that in the low frequency sampling case, when the sample consists of $(X_0, X_\Delta, ..., X_{n\Delta})$ for some fixed sampling distance $\Delta > 0$, the model satisfies the local asymptotic normality (LAN) property, assuming that b satisfies some mild regularity assumptions. This is established by using the connections of diffusion processes to elliptic and parabolic PDEs. The key tools used are regularity estimates for certain parabolic PDEs as well as a detailed analysis of the spectral properties of the elliptic differential operator related to $(X_t : t \geq 0)$.

MSC 2010 subject classifications: 62M99.
Keywords and phrases: Nonparametric diffusion model, LAN property, parabolic PDE.

Received March 2018.

1. Introduction

Consider a scalar diffusion, described by a stochastic differential equation (SDE)

$$dX_t = b(X_t)dt + \sqrt{2}dW_t, \qquad t \geq 0,$$

where $(W_t : t \geq 0)$ is a standard Brownian motion and b is the unknown *drift function* that is to be estimated. We investigate the so-called *low frequency* observation scheme, where the data consists of states

$$X^{(n)} = (X_0, X_\Delta ..., X_{n\Delta}) \tag{1}$$

of one sample path of $(X_t : t \geq 0)$, where $\Delta > 0$ is the *fixed* time difference between measurements. To ensure ergodicity and to limit technical difficulties, we follow [12] and [24] and consider a version of the model where the diffusion takes values on $[0, 1]$ with reflection at the boundary points $\{0, 1\}$, see Section 2.1 for the precise definition.

The nonparametric estimation of the coefficients of a diffusion process has attracted a great deal of attention in the past. For the low-frequency sampling scheme (1), Gobet, Hoffmann and Reiss [12] determined the minimax rate of estimation for both the drift and diffusion coefficient and also devised a spectral estimation method which achieves this rate. Thereafter, Nickl and Söhl [24] proved that the Bayesian posterior distribution contracts at the minimax rate, giving a frequentist justification for the use of Bayesian methods. In other sampling schemes, various methods have been studied, see e.g. [15] for a frequentist approach, [13, 16, 28, 1, 22] for recent posterior consistency and contraction rate results for Bayesian methods as well as [25, 26] for MCMC methodology for the computation of the Bayesian posterior.

However, often one desires a more detailed understanding of the performance of both frequentist and Bayesian methods, e.g. by establishing semi-parametric efficiency bounds or by proving a nonparametric Bernstein-von Mises theorem (BvM), which would give a frequentist justification for the use of Bayesian credible sets as confidence sets (see [10], Chapter 7.3). Nonparametric BvMs have been explored in the papers [5, 6] and have recently been proven for a number of statistical inverse problems [21, 23, 20], by Nickl and co-authors. In a diffusion model with continuous observations $(X_t : t \leq T)$, Nickl and Ray [22] recently proved a nonparametric BvM for estimating the drift b.

To order to achieve such a detailed understanding, a key step lies in studying the local information geometry of the parameter space, which in terms of semi-parametric efficiency theory (see e.g. [27], Chapter 25) involves finding the LAN expansion and the corresponding (Fisher) information operator. This in turn determines the Cramér-Rao lower bound for estimating a certain class of functionals of the parameter of interest. While in the Gaussian white noise model with direct observations, the LAN expansion of the log-likelihood ratio is exact and given by the Cameron-Martin theorem, in inverse problems proving the LAN property is often not straightforward.

In a finite-dimensional (parametric) model for multidimensional diffusions which are sampled at high frequency, where the sample consists of states

$$X^{(n)} = (X_0, X_{\Delta_n} ..., X_{n\Delta_n})$$

with asymptotics such that $\Delta_n \to 0$ and $n\Delta_n \to \infty$, the LAN property was shown by Gobet [11] by use of Malliavin calculus.

The main contribution of this paper is to prove that also with *low* frequency observations, the reflected diffusion model satisfies the LAN property, under mild regularity assumptions on the drift b. If the transition densities of the Markov chain $(X_{i\Delta} : i \in \mathbb{N})$ are denoted by $p_{\Delta,b}$, then the log-likelihood of the sample (1) is approximately equal to

$$\ell_b(X^{(n)}) \approx \sum_{i=1}^{n} \log p_{\Delta,b}(X_{(i-1)\Delta}, X_{i\Delta}),$$

from which one can see the necessity of two ingredients to show the LAN expansion:

- The first is a result on the differentiability of the transition densities $b \mapsto p_{\Delta,b}(x, y)$, which guarantees that we can form the second-order Taylor expansion of the log-likelihood in certain ‘directions’ $h/\sqrt{n}$ with sufficiently good control over the remainder. See Theorem 1 for the precise statement, where we importantly also obtain an explicit form for the first derivative A_b, the ‘score operator’.

- The second main ingredient consists of two well known limit theorems, the central limit theorem for martingale difference sequences [4] and the ergodic theorem, which ensure the right limits for the first and second order terms in the Taylor expansion respectively.

In view of this, the main work done in this paper lies in establishing the regularity needed for $p_{\Delta,b}(x, y)$, see Theorem 1 below. As there is no explicit formula for $p_{\Delta,b}(x, y)$ in terms of b, our approach relies on techniques from the theory of parabolic PDE and spectral theory. We use a PDE perturbation argument, based on the fact that the transition densities of a diffusion process can naturally be viewed as the fundamental solution to a related parabolic PDE.

The main difficulty in the proofs lies in the singular behaviour of $p_{t,b}(x, y)$ as (t, x) approaches $(0, y)$, which is why standard PDE results cannot be applied directly, but only in a regularised setting. Thus the arguments will first be carried out for any fixed regularisation parameter $\delta > 0$, where the analysis needs to be done carefully in order to ensure that the estimates obtained are uniform in $\delta > 0$ and hence still valid in the limit $\delta \to 0$.

In the context of a statistical inverse problem for the (elliptic) Schrödinger equation [21, 18], where the above singular behaviour is not present, PDE perturbation arguments have previously been used to linearize the log-likelihood.

We also remark that the use of more probabilistic proof techniques like in [11] would have been conceivable, too. However, we found the PDE approach employed here to be more naturally suited to dealing with boundary conditions, and it avoids dealing with pathwise properties of the diffusions by working with the transitions densities directly, which are ultimately the objects of interest for analyzing the likelihood.

Potential applications of the LAN expansion presented in Theorem 2 include the study of semiparametric efficiency for a certain class of functionals of b which is implicitly defined by the range of the 'information operator' $A_b^* A_b$ (where A_b is the score operator (9)), as well as an infinite-dimensional Bernstein-von-Mises theorem similar to [20, 21, 22, 23]. However, studying the properties of $A_b^* A_b$ needed for this poses a highly non-trivial challenge which still has to be overcome, see Section 2.4 for a more detailed discussion.

In Section 2, we state and prove the LAN expansion. Section 3 is devoted to proving Theorem 1. Finally, in Section 4, we derive the spectral properties of the differential operator $\mathcal{L}_b$ and the transition semigroup $(P_{t,b} : t \geq 0)$ needed throughout the proofs.

2. Main results

2.1. A reflected diffusion model

We shall work with boundary reflected diffusions on the interval $[0, 1]$, following [12, 24]. Consider the stochastic process $(X_t : t \geq 0)$, whose evolution is described by the stochastic differential equation (SDE)

$$dX_t = b(X_t)dt + \sqrt{2}dW_t \ + \ dK_t(X), \qquad X_t \in [0, 1], \quad t \geq 0. \tag{2}$$

Here $(W_t : t \geq 0)$ is a standard Brownian motion, $(K_t(X) : t \geq 0)$ is a non-anticipative finite variation process that only changes when $X_t \in \{0, 1\}$ and

$$b : [0, 1] \to \mathbb{R}$$

is the unknown drift function. We note that $K(X)$, which accounts for the reflecting boundary behaviour, is part of a solution to (2) and is in fact given by the difference of the local times of X at 0 and 1.

For any integer $s \geq 0$, let $C^s = C^s((0, 1))$ and $H^s = H^s((0, 1))$ denote the spaces of s-times continuously differentiable functions and s-times weakly differentiable functions with L^2-derivatives, respectively, endowed with the usual norms. We also define the subspace

$$C_0^1 := \{f \in C^1 : f(0) = f(1) = 0\}.$$

We assume throughout that for some $B < \infty$, b lies in the C_0^1-ball

$$\Theta := \big\{f \in C_0^1 : \|f\|_{C^1} := \|f\|_\infty + \|f'\|_\infty \leq B\big\}. \tag{3}$$

This ensures the existence of a pathwise solution $(X_t : t \geq 0)$ to (2) which can be constructed by a reflection argument, see e.g. Section I.§23 in [9] or [24]. For some $\Delta > 0$, which we assume to be *fixed* throughout the paper, our sample consists of measurements $X^{(n)} = (X_0, X_\Delta ..., X_{n\Delta})$ of one sample path, with asymptotics $n \to \infty$.

The process $(X_t : t \geq 0)$ forms an ergodic Markov process with invariant distribution $\mu_b = \mu$, whose Lebesgue density (which we also denote by μ_b) is identified by b via

$$\mu_b(x) = \frac{e^{\int_0^x b(y)dy}}{\int_0^1 e^{\int_0^u b(y)dy} du}, \quad x \in [0, 1], \tag{4}$$

see e.g. Chapter 4 in [3]. Moreover, we denote the Lebesgue transition densities and the semigroup associated to $(X_t : t \geq 0)$ by $p_{t,b}$ and $P_{t,b}$ respectively:

$$p_{t,b} : [0, 1]^2 \to \mathbb{R}, \quad p_{t,b}(x, y) = \mathbb{P}_x(X_t \in dy), \quad t > 0, \tag{5}$$

$$P_{t,b} f(x) = \mathbb{E}_x[f(X_t)] = \int_0^1 p_{t,b}(x, z) f(z) dz, \quad t > 0, \quad f \in L^2. \tag{6}$$

Here, by Proposition 9 in [24], the transition densities are well–defined as well as bounded above and below for each $t > 0$, so that (6) is well-defined, too.

Let $\mathbb{P}_b$ denote the law of $(X_{i\Delta} : i \geq 0)$ on $[0, 1]^{\mathbb{N}}$. For ease of exposition, we assume throughout that $X_0 \sim \mu_b$ under $\mathbb{P}_b$, a common assumption (cf. [12, 24]) which we make due to the uniform spectral gap over $b \in \Theta$ guaranteed by Lemma 12 below, which yields exponentially fast convergence of X_t to μ_b. Then under any $\mathbb{P}_b$, $b \in \Theta$, the law of $X^{(n)}$ from (1) on $[0, 1]^{n+1}$ is absolutely continuous with respect to the $n + 1$– dimensional Lebesgue measure, and the log–density, which also constitutes the *log–likelihood* (when viewed as a function of b), is given by

$$\log d\mathbb{P}_b(X^{(n)}) = \log \mu_b(X_0) + \sum_{i=1}^{n} \log p_{\Delta,b}(X_{(i-1)\Delta}, X_{i\Delta}). \tag{7}$$

We note that some of the above assumptions can be relaxed at the expense of further technicalities in the proofs: Firstly, the assumption $X_0 \sim \mu_b$ could be replaced by $X_0 \sim \pi_b$ (under $\mathbb{P}_b$), for any measures π_b with Lebesgue densities such that for all $b \in \Theta$, $\log d\pi_{\tilde{b}}(X_0) - \log d\pi_b(X_0) = o_{\mathbb{P}_b}(1)$ as $\|\tilde{b} - b\|_{H^1} \to 0$. Secondly, it is conceivable that the main Theorems 1 and 2 below can be generalized to all $b \in H^1$ and $h \in \{f \in H^1 : f(0) = f(1) = 0\}$, which we shall not pursue further here, however.

2.2. Differentiability of the transition densities

In order to prove the LAN property, we need to differentiate the log-likelihood (7) at any drift parameter $b \in \Theta$, and the following theorem shows that for any $x, y \in [0,1]$, maps of the form $b \mapsto p_{\Delta,b}(x,y)$ are infinitely differentiable in 'directions' $h \in C_0^1$ (and in fact, Fréchet differentiable). For $b, h \in C_0^1$, $\eta \in \mathbb{R}$ and $x, y \in [0,1]$, for convenience we introduce the notation

$$\Phi_{b,h,x,y} = \Phi : \mathbb{R} \to \mathbb{R}, \qquad \Phi(\eta) = p_{\Delta,b+\eta h}(x,y).$$

Theorem 1. *For all $b, h \in C_0^1$ and $x, y \in [0,1]$, $\Phi = \Phi_{b,h,x,y}$ is a smooth (in fact, real analytic) function on $\mathbb{R}$, and we have*

$$\Phi'(0) = \int_0^\Delta P_{\Delta-s,b}[h\partial_1 p_{s,b}(\cdot,y)](x)ds. \tag{8}$$

Moreover, for each integer $k \geq 1$, we have the following bound on the k-th derivative of Φ at 0:

$$\sup_{b\in\Theta} \sup_{h\in C_0^1, \|h\|_{H^1}\leq 1} \sup_{x,y\in[0,1]} \left|\Phi^{(k)}(0)\right| < \infty.$$

Section 3 is devoted to the proof of Theorem 1.

Heuristically speaking, the right hand side of (8) has the form of a solution to an inhomogeneous parabolic PDE (cf. Proposition 4), and this PDE perspective will be key in the proofs. However, one has to be careful with such an interpretation, as the singular 'source term' $h\partial_1 p_{b,t}(\cdot, y)$ does not fall within the scope of classical PDE theory. Therefore, the above intuition needs to be made rigorous via a regularisation argument, see Section 3.

2.3. LAN expansion

By Lemma 15, for each $b \in \Theta$, $p_{\Delta,b}(\cdot,\cdot)$ is bounded above and below. Hence by Theorem 1 and the chain rule, the score operator is given by

$$A_b: \quad C_0^1([0,1]) \;\rightarrow\; L^2([0,1]\times[0,1]), \qquad A_b h(x,y) = \frac{[\Phi_{b,h,x,y}]'(0)}{p_{\Delta,b}(x,y)}. \tag{9}$$

For any $f, g \in L^2([0,1]\times[0,1])$, we also define the corresponding 'LAN inner product' and 'LAN norm' as follows:

$$\begin{aligned} \langle f,g\rangle_{L^2(p_{\Delta,b}\mu_b)} &:= \int_0^1\int_0^1 f(x,y)g(x,y)\mu_b(x)p_{\Delta,b}(x,y)dxdy, \\ \langle f,g\rangle_{LAN} &:= \langle A_b f, A_b g\rangle_{L^2(p_{\Delta,b}\mu_b)}, \qquad \|f\|_{LAN}^2 := \langle f,f\rangle_{LAN}. \end{aligned} \tag{10}$$

Here is our main result, the proof can be found in Section 2.5.

Theorem 2 (LAN expansion)**.** *For any $b, h \in C_0^1$, we have that*

$$\log\frac{d\mathbb{P}_{b+h/\sqrt{n}}}{d\mathbb{P}_b}(X^{(n)}) = \frac{1}{\sqrt{n}}\sum_{i=1}^n A_b h(X_{(i-1)\Delta}, X_{i\Delta}) - \frac{1}{2}\|h\|_{LAN}^2 + o_{\mathbb{P}_b}(1) \tag{11}$$

as $n \to \infty$ and

$$\frac{1}{\sqrt{n}}\sum_{i=1}^n A_b h(X_{(i-1)\Delta}, X_{i\Delta}) \xrightarrow{n\to\infty}{}^{d}\; N(0, \|h\|_{LAN}^2). \tag{12}$$

Note that due to the nature of the non-i.i.d. Markov chain data at hand, A_b necessarily needs to map into a function space of two variables, as the overall log-likelihood cannot be formed as a sum of functions of single states of the chain, but only of increments of the chain.

2.4. Potential statistical applications of Theorem 2

The LAN expansion can be used to obtain semiparametric lower bounds for the estimation of certain linear functionals $L(b)$ for which there exists a Riesz representer $\Psi \in C_0^1$ such that $L(\cdot) = \langle \Psi, \cdot\rangle_{LAN}$, and can potentially further be used to prove a non-parametric Bernstein-von-Mises theorem.

To make this more precise, we define the 'information operator' (which generalizes the Fisher information) by $I_b := A_b^* A_b : C_0^1 \to L^2$, where A_b from (9) is viewed as a densely defined operator on L^2 with domain C_0^1 and A_b^* is the adjoint of A_b with respect to the inner products $\langle\cdot,\cdot\rangle_{L^2}$ and $\langle\cdot,\cdot\rangle_{L^2(p_{\Delta,b}\mu_b)}$. Then, for example, to study semiparametric Cramér-Rao lower bounds for functionals of the form $L(b) = \langle\psi, b\rangle_{L^2}$, $\psi \in L^2$, one needs that there is some $\Psi \in C_0^1$ such that

$$\forall w \in C_0^1: \quad \langle\psi, w\rangle_{L^2} = \langle\Psi, w\rangle_{LAN} = \langle I_b\Psi, w\rangle_{L^2}.$$

Hence one needs that ψ lies in the range $R(I_b)$ of I_b (or at least of $R(A_b^*)$), see p.372-373 in [27] for a detailed discussion. Assuming the injectivity of I_b, the 'optimal asymptotic variance' for estimators of $L(b)$ is then given by

$$\|\Psi\|_{LAN}^2 = \langle A_b\Psi, A_b\Psi\rangle_{L^2(p_{\Delta,b}\mu_b)},$$

which may intuitively be understood as an 'inverse Fisher information $\langle\psi, I_b^{-1}\psi\rangle_{L^2}$', in analogy to the parametric setting.

When $R(I_b)$ is known to contain at least a 'nice' subspace of functions, e.g. C_c^∞, I_b can be inverted on that subspace, and if key mapping properties of I_b^{-1} are known, then along the lines of [21, 23, 22, 20], one can further try to prove a nonparametric BvM. This would assert the convergence of infinite-dimensional posterior distributions to a Gaussian limit measure $\mathcal{G}$ whose covariance is given by the LAN inner product via $\mathrm{Cov}[G(\psi_1), G(\psi_2)] = \langle\Psi_1, \Psi_2\rangle_{LAN}$, cf. (28) in [21].

The identification of $R(I_b)$ in the present case of diffusions sampled at low frequency, as well as the study of mapping properties of I_b, remain challenging open problems.

2.5. Proof of the LAN expansion

We now give the proof of Theorem 2, assuming the validity of Theorem 1 which is proven in Section 3 below. Besides Theorem 1, the other key ingredient for Theorem 2 is the following CLT for martingale difference sequences. It is due to Brown (building on ideas of Billingsley and Lévy) and follows immediately from the special case $t = 1$ in Theorem 2 of [4].

Proposition 3 (cf. [4]). *Suppose $(\Omega, \mathcal{F}, (\mathcal{F}_n : n \geq 0), \mathbb{P})$ is a filtered probability space and let $(M_n : n \in \mathbb{N})$ be a $\mathcal{F}_n$-martingale with $M_0 = 0$. For $n \geq 1$, define the increments $Y_n := M_n - M_{n-1}$ and let*

$$\sigma_n^2 := \mathbb{E}\left[Y_n^2 \middle| \mathcal{F}_{n-1}\right], \qquad V_n^2 := \sum_{i=1}^n \sigma_i^2, \qquad s_n^2 := \mathbb{E}[V_n^2].$$

Suppose that $V_n^2 s_n^{-2} \xrightarrow{n\to\infty} 1$ in probability and that for all $\epsilon > 0$,

$$\frac{1}{s_n^2}\sum_{i=1}^n \mathbb{E}\left[Y_i^2 \mathbb{1}\{|Y_i| \geq \epsilon s_n\}\right] \xrightarrow{n\to\infty} 0 \tag{13}$$

in probability. Then, as $n \to \infty$, we have

$$M_n/s_n \xrightarrow{d} \mathcal{N}(0,1).$$

Proof of Theorem 2. Fix $b, h \in C_0^1$. Due to the spectral gap of the generator $\mathcal{L}_b$ (see Lemma 12), the Markov chain $(X_{n\Delta} : n \in \mathbb{N})$ originating from the diffusion (2) with initial distribution $X_0 \sim \mu_b$, is stationary and geometrically ergodic – we will use this fact repeatedly.

For notational convenience, we write

$$f(\eta,x,y)=\log p_{\Delta,b+\eta h}(x,y),\qquad g(\eta,x,y)=p_{\Delta,b+\eta h}(x,y).$$

By Theorem 1, f is smooth in η on a neighbourhood of 0, and for some $C<\infty$, the second order Taylor remainder satisfies

$$R_f(\eta):=\sup_{x,y\in[0,1]}|f(\eta,x,y)-f(0,x,y)-\eta\partial_\eta f(0,x,y)-\frac{\eta^2}{2}\partial_\eta^2 f(0,x,y)|\le C|\eta|^3. \tag{14}$$

Thus, Taylor-expanding the log-likelihood (7) in direction $h/\sqrt{n}$ yields that

$$\begin{aligned}\log&\frac{d\mathbb{P}^n_{b+h/\sqrt{n}}}{d\mathbb{P}^n_b}(X_0,...,X_{n_\Delta})\\ &=\big(\log\mu_{b+h/\sqrt{n}}(X_0)-\log\mu_b(X_0)\big)+\frac{1}{\sqrt{n}}\sum_{i=1}^n\partial_\eta f(0,X_{(i-1)\Delta},X_{i\Delta})\\ &\quad+\frac{1}{2n}\sum_{i=1}^n\partial_\eta^2 f(0,X_{(i-1)\Delta},X_{i\Delta})+D_n\\ &=:A_n+B_n+C_n+D_n.\end{aligned} \tag{15}$$

For the remainder term D_n, we immediately see from (14) that $|D_n|\le nR_f(n^{-1/2})$, whence $D_n=o_{\mathbb{P}_b}(1)$ as $n\to\infty$.

For C_n, observe that the function $\partial_\eta^2 f(0,\cdot,\cdot)$ is bounded by Theorem 1, such that the almost sure ergodic theorem yields that

$$C_n\xrightarrow{n\to\infty}\frac{1}{2}\mathbb{E}_b[\partial_\eta^2 f(0,X_0,X_\Delta)]\quad\text{a.s.},$$

where $\mathbb{E}_b$ denotes the expectation with respect to $\mathbb{P}_b$. Moreover, we have

$$\partial_\eta^2 f(0,X_0,X_\Delta)=\frac{\partial_\eta^2 g(0,X_0,X_\Delta)}{g(0,X_0,X_\Delta)}-\big(\partial_\eta f(0,X_0,X_\Delta)\big)^2=:I+II,$$

and by interchanging differentiation and integration (which is possible by Theorem 1), we see that

$$\mathbb{E}[I]=\int_0^1\int_0^1\partial_\eta^2 g(0,x,y)\mu_b(x)dxdy=0,$$

and hence $\mathbb{E}_b[\partial_\eta^2 f(0,X_0,X_\Delta)]=-\langle A_bh,A_bh\rangle_{L^2(p_{\Delta,b}\mu_b)}=-\|h\|^2_{LAN}$.

We next treat B_n. Let $(\mathcal{F}_n:n\ge 0)$ denote the natural filtration of $(X_{\Delta n}:n\ge 0)$. In view of Proposition 3, let us write

$$Y_n=\partial_\eta f(0,X_{(n-1)\Delta},X_{n\Delta}),\qquad M_n:=\sqrt{n}B_n=\sum_{i=1}^n Y_n,$$

$$\sigma_n^2=\mathbb{E}\left[Y_n^2\middle|X_{(n-1)\Delta}\right],\qquad V_n^2=\sum_{i=1}^n\sigma_i^2,\qquad s_n^2=\mathbb{E}[V_n^2].$$

Then, using dominated convergence and the Markov property, we see that $M_0 = 0$ and that $(M_n : n \geq 0)$ is a martingale:

$$\begin{aligned}\mathbb{E}[Y_n|\mathcal{F}_{n-1}] &= \int_0^1 \partial_\eta f(0, X_{(n-1)\Delta}, y) p_{\Delta,b}(X_{(n-1)\Delta}, y) dy \\ &= \int_0^1 \partial_\eta g(0, X_{(n-1)\Delta}, y) dy \\ &= \partial_\eta \int_0^1 p_{\Delta, b+\eta h}(X_{(n-1)\Delta}, y) dy \,\Big|_{\eta=0} = 0.\end{aligned}$$

Moreover, we have that $\sigma_n^2 = \tilde{\sigma}^2(X_{(n-1)\Delta})$ for some bounded measurable function $\tilde{\sigma}^2 : [0,1] \to [0,\infty)$ and by the stationarity of $(X_{i\Delta} : i \geq 0)$, we have $s_n^2 = n\mathbb{E}_b[\tilde{\sigma}^2(X_0)] = n\|h\|_{LAN}^2$, whence the ergodic theorem yields that $\mathbb{P}_b-$ a.s.,

$$V_n^2 s_n^{-2} = \frac{1}{n\|h\|_{LAN}^2} \sum_{i=1}^n \tilde{\sigma}^2(X_{(n-1)\Delta}) \xrightarrow{n\to\infty} \|h\|_{LAN}^{-2} \mathbb{E}_b[(\partial_\eta f(0, X_0, X_1))^2] = 1.$$

Lastly, as the Y_i's are bounded random variables, the condition (13) is fulfilled. Hence Proposition 3 yields that $B_n \to^d \mathcal{N}(0, \|h\|_{LAN}^2)$.

Finally, we observe that the term A_n in (15) from the invariant measure is of order $o_{\mathbb{P}_b}(1)$, as it can be bounded uniformly over b, h using (4):

$$\begin{aligned}&|\log \mu_{b+h/\sqrt{n}}(X_0) - \log \mu_b(X_0)| \lesssim \|\mu_{b+h/\sqrt{n}} - \mu_b\|_\infty \\ &\lesssim \Big\| \frac{e^{\int_0^\cdot (b+h/\sqrt{n})(y)dy}}{\int_0^1 e^{\int_0^x (b+h/\sqrt{n})(y)dy} dx} - \frac{e^{\int_0^\cdot b(y)dy}}{\int_0^1 e^{\int_0^x b(y)dy} dx} \Big\|_\infty \xrightarrow{n\to\infty} 0.\end{aligned}$$

□

3. Local approximation of transition densities

In this section, we study the differentiability properties of $p_{t,b}(x, y)$ as a function of the drift b, and the main goal is to prove Theorem 1. For technical reasons, we first prove a regularized version of it (Lemma 8 in Section 3.2) and then let the regularization parameter $\delta > 0$ tend to 0 to obtain Theorem 1 (Section 3.3).

3.1. Preliminaries and notation

We begin by introducing some notation and important classical results.

3.1.1. Some function spaces

For any integer $s \geq 0$, we equip the Sobolev space $H^s = H^s((0,1))$ with the inner product

$$\langle g_1, g_2 \rangle_{H^s} = \langle g_1, g_2 \rangle_{L^2} + \langle g_1^{(s)}, g_2^{(s)} \rangle_{L^2}, \tag{16}$$

where L^2 is the usual space of square integrable functions with respect to Lebesgue measure. Occasionally it will be convenient to replace the L^2–inner product above by the $L^2(\mu)$–inner product, where μ is the invariant measure of $(X_t : t \geq 0)$, which by (24) induces a norm which is equivalent to the norm induced by (16).

We will also use the fractional Sobolev spaces H^s for real $s \geq 0$, which are obtained by interpolation, see [17]. For $s > \frac{1}{2}$, the Sobolev embedding (19) implies that any function $f \in H^s$ extends uniquely to a continuous function on $[0, 1]$. The following standard interpolation equalities and embeddings (see [17], p.44-45) will be used throughout. For all $s_1, s_2 \geq 0$ and $\theta \in (0, 1)$, we have

$$\forall f \in H^{s_1} \cap H^{s_2} : \quad \|f\|_{H^{\theta s_1 + (1-\theta) s_2}} \leq C(\theta, s_1, s_2) \|f\|^{\theta}_{H^{s_1}} \|f\|^{1-\theta}_{H^{s_2}}, \tag{17}$$

and for each $s > 1/2$, we have the multiplicative inequality

$$\forall f, g \in H^s : \quad \|fg\|_{H^s} \lesssim C(s) \|f\|_{H^s} \|g\|_{H^s} \tag{18}$$

as well as the continuous embedding

$$H^s \subseteq C([0, 1]), \quad \|f\|_\infty \leq C(s) \|f\|_{H^s}, \tag{19}$$

where $C([0, 1])$ denotes the space of continuous functions on $[0, 1]$. Moreover, for any $s > 0$, we define the negative order Sobolev space H^{-s} as the topological dual space of H^s, where for any $f \in L^2$, the norm can be written as

$$\|f\|_{H^{-s}} = \sup_{\psi \in H^s, \|\psi\|_{H^s} \leq 1} \Big| \int_0^1 f \psi \Big|.$$

For any $T > 0$, any Banach space $(X, \|\cdot\|)$ and any integer $k \geq 0$, we denote by $C^k([0, T], X)$ the k-times continuously differentiable functions from $[0, T]$ to X, equipped with the norm

$$\|f\|_{C^k([0,T],X)} = \sum_{i=0}^{k} \sup_{t \in [0,T]} \Big\| \frac{d^i}{dt^i} f(t) \Big\|.$$

For $\alpha > 0$ with $\alpha \notin \mathbb{N}$, we denote the space of α-Hölder continuous functions $f : [0, T] \to X$ by $C^\alpha([0, T], X)$ and equip it with the usual norm

$$\|f\|_{C^\alpha([0,T],X)} = \|f\|_{C^{\lfloor \alpha \rfloor}([0,T],X)} + \sup_{s,t \in [0,T], s \neq t} \frac{\Big\| \frac{d^{\lfloor \alpha \rfloor}}{dt^{\lfloor \alpha \rfloor}} f(t) - \frac{d^{\lfloor \alpha \rfloor}}{dt^{\lfloor \alpha \rfloor}} f(s) \Big\|}{\big| t - s \big|^{\alpha - \lfloor \alpha \rfloor}}.$$

We will frequently, without further comment, interpret functions $f : [0, T] \times [0, 1] \to \mathbb{R}$ as L^2-valued maps $f : [0, T] \to L^2$, $f(t) = f(t, \cdot)$, and vice versa.

3.1.2. The differential operator $\mathcal{L}_b$

For any drift function $b \in C^1_0$, we define the differential operator

$$\mathcal{L}_b f(x) := f''(x) + b(x) f'(x) \qquad \text{for } f \in \mathcal{D},$$
$$\mathcal{D} := \left\{ f \in H^2 : f'(0) = f'(1) = 0 \right\}.$$

It is well-known that $\mathcal{L}_b$ is the infinitesimal generator of the semigroup $(P_{t,b} : t \geq 0)$ defined in (6), so that we get by the usual functional calculus that $P_{t,b} = e^{t\mathcal{L}_b}$ for all $t \geq 0$ (with the convention $e^0 = Id$). The fact that the domain $\mathcal{D}$ of $\mathcal{L}_b$ is equipped with Neumann boundary conditions corresponds to the diffusion being reflected at the boundary, see [14] for a detailed discussion. We equip $\mathcal{D}$ with the graph norm

$$\|f\|_{\mathcal{L}_b} := \left(\|\mathcal{L}_b f\|^2_{L^2(\mu_b)} + \|f\|^2_{L^2(\mu_b)}\right)^{1/2},$$

which by Lemma 13 is equivalent to the H^2-norm on $\mathcal{D}$. Moreover, for $h \in H^1$, we define the first order differential operator

$$L_h f(x) = h(x)f'(x), \qquad f \in H^1. \tag{20}$$

The operator $\mathcal{L}_b$ has a purely discrete spectrum $\mathrm{Spec}(\mathcal{L}_b) \subseteq (-\infty, 0]$ (see [8], Theorem 7.2.2). We will denote by $(u_{j,b})_{j\geq 0}$ the $L^2(\mu_b)$-normalized orthogonal basis of $L^2(\mu_b)$ consisting of the eigenfunctions $u_{j,b} \in \mathcal{D}$ of $\mathcal{L}_b$, ordered such that the corresponding eigenvalues $(\lambda_{j,b})_{j\geq 0}$ are non-increasing. When there is no ambiguity, we will often simply write λ_j and u_j. We will use throughout the spectral decomposition

$$p_{t,b}(x,y) = \sum_{j\geq 0} e^{\lambda_j t} u_j(x)u_j(y)\mu(y), \quad x, y \in [0,1],\ t > 0, \tag{21}$$

see e.g. p. 101 in [2], and the spectral representations

$$\mathcal{L}_b f = \sum_{j\geq 0} \lambda_j \langle f, u_j\rangle_{L^2(\mu)} u_j, \quad f \in \mathcal{D}, \tag{22}$$

$$P_{t,b} f = \sum_{j\geq 0} e^{t\lambda_j} \langle f, u_j\rangle_{L^2(\mu)} u_j, \quad f \in L^2,\ t > 0. \tag{23}$$

We also note that (4) immediately yields that there exist constants $0 < C < C' < \infty$ such that for all $b \in \Theta$ and all $x \in [0,1]$,

$$C \leq \mu_b(x) \leq C'. \tag{24}$$

3.1.3. *A key PDE result*

For any $f \in C([0,T], L^2)$ and $u_0 \in \mathcal{D}$, consider the inhomogeneous parabolic equation

$$\begin{cases} \frac{d}{dt}u(t) = \mathcal{L}_b u(t) + f(t) & \text{for all } t \in [0,T], \\ u(0) = u_0. \end{cases} \tag{25}$$

We say that a function $u : [0,T] \to L^2$ is a solution to (25) if $u \in C^1([0,T], L^2) \cap C([0,T], \mathcal{D})$ and (25) holds. The next proposition regarding the existence, uniqueness and regularity properties of solutions to (25) will play a key role for the proofs in the rest of Section 3. To state the result, we need the following interpolation spaces $\mathcal{D}(\alpha), 0 \le \alpha \le 1$, between L^2 and $\mathcal{D}$:

$$\mathcal{D}(\alpha) := \left\{ f \in L^2 : \omega(t) := t^{-\alpha} \|P_{t,b} f - f\|_{L^2(\mu_b)} \text{ is bounded on } t \in (0,1] \right\},$$
$$\|f\|_{\mathcal{D}(\alpha)} := \|f\|_{L^2(\mu_b)} + \sup_{t\in[0,1]} \omega(t). \tag{26}$$

Proposition 4. *Suppose $0 < \alpha < 1$, $f \in C^\alpha([0,T], L^2)$ and $u_0 \in \mathcal{D}$. Then there exists a unique solution u to (25), given by the Bochner integral*

$$u(t) = P_{t,b} u_0 + \int_0^t P_{t-s,b} f(s) ds, \quad t \in [0,T]. \tag{27}$$

If also $f(0) + \mathcal{L}_b u_0 \in \mathcal{D}(\alpha)$, then we have $u \in C^{1+\alpha}([0,T], L^2) \cap C^\alpha([0,T], \mathcal{D})$ and there exists $C < \infty$ so that for all such f and u_0,

$$\|u\|_{C^{1+\alpha}([0,T],L^2)} + \|u\|_{C^\alpha([0,T],\mathcal{D})}$$
$$\le C \left(\|f\|_{C^\alpha([0,T],L^2)} + \|u_0\|_{\mathcal{L}_b} + \|f(0) + \mathcal{L}_b u_0\|_{\mathcal{D}(\alpha)} \right).$$

Proof. This is a special case of Theorem 4.3.1 (iii) in [19] with $X = L^2(\mu_b)$ and $A = \mathcal{L}_b$, where we note that the integral formula (27) is given by Proposition 4.1.2 in the same reference. We also note that $\mathcal{D}(\alpha)$ coincides with the space $D_A(\alpha, \infty)$ from [19] with equivalent norms, see Proposition 2.2.4 in [19]. It therefore suffices to verify that the general theory for parabolic PDEs developed in [19] applies to our particular case. For that, we need to check that $(P_{t,b} : t \ge 0)$ is an analytic semigroup of operators on L^2 in the sense of [19], p.34, which requires the following.

1. For some $\theta \in (\pi/2, \pi)$ and $\omega \in \mathbb{R}$, the resolvent set $\rho(\mathcal{L}_b)$ of $\mathcal{L}_b$ contains the sector $S_{\theta,\omega} \subseteq \mathbb{C}$, where $S_{\theta,\omega}$ is defined by
$$S_{\theta,\omega} := \{\lambda \in \mathbb{C} : \lambda \ne \omega, |\arg(\lambda - \omega)| < \theta\}.$$
2. There exists some $M < \infty$ such that we have the resolvent estimate
$$\|R(\lambda, \mathcal{L}_b)\|_{L^2 \to L^2} \le M |\lambda - \omega|^{-1} \quad \forall \lambda \in S_{\theta,\omega}.$$

As $\mathcal{L}_b$ has a discrete spectrum contained in the non-positive half line, both of the above properties are easily checked with $\omega = 0$ and any $\theta \in (\frac{\pi}{2}, \pi)$. □

Definition 5 (Solution operator)**.** In what follows, we denote by $\mathcal{S} = (\frac{d}{dt} - \mathcal{L}_b)^{-1}$ the linear solution operator which maps any $f \in C^\alpha([0,T], L^2)$, $0 < \alpha < 1$, to the unique solution $u = \mathcal{S}(f)$ of the parabolic problem

$$\begin{cases} \left(\frac{d}{dt} - \mathcal{L}_b\right) u(t) = f(t), & t \in [0,T], \\ u(0) = 0. \end{cases} \tag{28}$$

3.2. Approximation of regularized transition densities

The main result of this section is Lemma 8, which can be viewed as a regularized version of Theorem 1. The main tools used to prove it are Proposition 4 as well as the spectral analysis of $\mathcal{L}_b$ and $P_{t,b}$ from Section 4.

In order to apply Proposition 4, we view the transition densities $p_{t,b}(x,y)$ as functions of the two variables $(t,x) \in [0,T] \times [0,1]$ with $y \in [0,1]$ fixed, where T is an arbitrary constant $T > \Delta > 0$ (with the convention that $p_{0,b}(\cdot,y)$ is the point mass at y). Due to the singular behaviour of $p_{t,b}(x,y)$ for $(t,x) \to (0,y)$, a regularisation argument is needed. For any $\delta > 0$ and $d \in C_0^1$, define the δ-regularized transition densities by

$$u_d^\delta : [0,\infty) \times [0,1] \to \mathbb{R}, \quad u_d^\delta(t,x) := P_{\delta,0}(p_{t,d}(x,\cdot))(y),$$

where $(P_{t,0} : t \geq 0)$ denotes the transition semigroup for $b = 0$, which corresponds to reflected Brownian motion.

3.2.1. Recursive definition of approximations

We now implicitly define the 'candidate' local approximations to u_d^δ as solutions to certain parabolic PDEs. To that end, we note that using (6), one easily checks that for all $t \geq 0$,

$$u_d^\delta(t) = P_{t,d}\varphi_\delta, \quad \text{where} \quad \varphi_\delta(x) := p_{\delta,0}(y,x). \tag{29}$$

Hence we can give the following crucial PDE interpretation to u_d^δ.

Lemma 6. *For any $d \in C_0^1$, we have that $u_d^\delta \in C^{3/2}([0,T],L^2) \cap C^{1/2}([0,T],\mathcal{D})$, and u_d^δ is the unique solution to the initial value problem*

$$\begin{cases} (\frac{d}{dt} - \mathcal{L}_d)u(t) = 0 \;\; \text{for all } t \in [0,T], \\ u(0) = \varphi_\delta. \end{cases} \tag{30}$$

Proof. We check that Proposition 4 applies with $\alpha = 1/2$. For this, we need that $\varphi_\delta \in \mathcal{D}$ and that $\mathcal{L}_d\varphi_\delta \in \mathcal{D}(1/2)$. Using the spectral decomposition (21) and the fact that $\mu_b = \mathrm{Leb}([0,1])$ for $b = 0$, we see by differentiating under the sum that $\varphi_\delta \in \mathcal{D}$. This is possible by Lemma 12 and the dominated convergence theorem. The same argument yields that $\varphi_\delta \in H^3$. Thus, we have that $\mathcal{L}_d\varphi_\delta \in H^1$, which is a subset of $\mathcal{D}(1/2)$ by the second part of Lemma 16. □

We now recursively define the functions $R_k^\delta[h]$ and $v_k^\delta[h]$, $k \geq 0$. The norm estimates in Section 3.2.2 justify that they are the correct remainder and approximating terms, respectively, in the k-th order Taylor expansion of $\eta \mapsto u_{b+\eta h}^\delta$.

Definition 7. Let $b, h \in C_0^1$ and $\delta > 0$.

1. For $k = 0$, we define the '0-th order local approximation' of $\eta \mapsto u_{b+\eta h}^\delta$ at 0, and the remainder of this approximation, by

$$v_0^\delta[h] = v_0^\delta := u_b^\delta, \qquad R_0^\delta[h] = R_0^\delta := u_{b+h}^\delta - u_b^\delta.$$

2. For $k \geq 1$, we recursively define the functions $R_k^\delta[h] = R_k^\delta, v_k^\delta[h] = v_k^\delta \in C^{3/2}([0,T], L^2) \cap C^{1/2}([0,T], \mathcal{D})$ by

$$R_k^\delta[h] := \mathcal{S}\big(L_h R_{k-1}^\delta[h]\big), \quad v_k^\delta[h] := R_{k-1}^\delta[h] - R_k^\delta[h], \tag{31}$$

where $\mathcal{S}$ is the solution operator defined in (28) and L_h was defined in (20).

We should justify why the definition (31) is admissible, and we do so by induction. By Lemma 6, we have $R_0^\delta[h] \in C^{3/2}([0,T], L^2) \cap C^{1/2}([0,T], \mathcal{D})$. Hence, using the definition of $R_k^\delta[h]$ and Proposition 4 inductively, we obtain that for all $k \geq 1$, $L_h R_{k-1}^\delta[h] \in C^{1/2}([0,T], H^1)$ as well as $L_h R_{k-1}^\delta[h](0) = 0$, so that R_k^δ, v_k^δ have the stated regularity. Thus, (31) is well-defined.

By definition of $\mathcal{L}_b$ and (30), we see that $R_0^\delta[h]$ is the unique solution to

$$(\frac{d}{dt} - \mathcal{L}_b) R_0^\delta(t) = L_h u_{b+h}^\delta(t) \ \ \forall t \in [0,T] \quad \text{and} \quad R_0^\delta(0) = 0, \tag{32}$$

and (31) yields that

$$u_{b+h}^\delta = \sum_{i=0}^{k} v_i^\delta[h] + R_k^\delta[h] \qquad \forall b, h \in C^1, \quad k \geq 0. \tag{33}$$

The regularity estimates for $R_k^\delta[h]$ in the next section will justify that (33) is in fact the Taylor approximation for $\eta \mapsto u_{b+\eta h}^\delta$. Before proceeding to this, we need to check that the $v_k^\delta[h]$ are homogeneous of degree k in h, i.e. that

$$\forall h \in C_0^1 \ \forall \eta \in \mathbb{R}: \quad v_k^\delta[\eta h] = \eta^k v_k^\delta[h]. \tag{34}$$

This is again seen by induction. For $k = 0$, we have that $v_0^\delta[\eta h] = u_b^\delta = v_0^\delta[h]$, and if (34) holds for some $k \geq 0$, then we have that

$$v_{k+1}^\delta[\eta h] = \mathcal{S}(L_{\eta h} v_k^\delta[\eta h]) = \eta^{k+1} \mathcal{S}(L_h v_k^\delta[h]),$$

where we have used that for each $k \in \mathbb{N} \cup \{0\}$,

$$\big(\frac{d}{dt} - \mathcal{L}_b\big) v_{k+1}^\delta = L_h(R_{k-1}^\delta - R_k^\delta) = L_h v_k^\delta.$$

3.2.2. Regularity estimates

We now derive norm estimates for the remainders $R_k^\delta[h]$ from (31) and (33), using Propsition 4 and the results from Section 4.

The following Lemma is the main result of Section 3. It can be viewed as a regularised version of Theorem 1. Crucially, the estimate below is uniform in $\delta > 0$ such that it can be preserved in the limit $\delta \to 0$.

Lemma 8. *For each $\epsilon > 0$, there exists $C > 0$ such that for all $b \in \Theta$ from (3), $h \in C_0^1$ with $\|h\|_{H^1} \leq 1$, $y \in [0,1]$, $k \in \mathbb{N} \cup \{0\}$ and $\delta > 0$,*

$$\|R_k^\delta[h](\Delta)\|_\infty \leq C^k \|h\|_{H^1}^{k+1/2-\epsilon}.$$

The rest of this section is concerned with proving Lemma 8. In what follows, when we write that an inequality is 'uniform' without further comment, or when we use the symbols $\lesssim, \gtrsim, \simeq$, we mean that the constants involved can be chosen uniformly over b, h, y, k and δ as in the statement of Lemma 8.

The proof of Lemma 8 consists of two separate lemmas, which establish an L^2-estimate (38) and an H^1-estimate (41) for $R_k^\delta[h](\Delta)$ respectively. Given these two estimates, Lemma 8 then immediately follows from interpolating – Indeed, taking C to be the larger of the two constants from (38) and (41) yields

$$\|R_k[h](\Delta)\|_\infty \lesssim \|R_k[h](\Delta)\|_{H^{\frac{1}{2}+\varepsilon}} \lesssim \|R_k(\Delta)\|_{L^2}^{\frac{1}{2}-\varepsilon} \|R_k(\Delta)\|_{H^1}^{\frac{1}{2}+\varepsilon} \leq C^k \|h\|_{H^1}^{k+\frac{1}{2}-\epsilon}.$$

The L^2-estimate To obtain estimates which are uniform in $\delta > 0$, we 'regularise' R_k^δ further by integrating in time. For $k \geq 0$, define

$$Q_k^\delta[h] : [0, T] \to L^2, \quad Q_k^\delta[h](t) := \int_0^t R_k^\delta[h](s) ds.$$

Here is the L^2-estimate.

Lemma 9. *1. Let $b, h \in C_0^1$, $\delta > 0$ and recall the definition (28) of $\mathcal{S}$. Then we have that*

$$Q_0^\delta[h] = \mathcal{S}\big(L_h \int_0^\cdot u_{b+h}^\delta(s) ds\big), \tag{35}$$

and for $k \geq 1$, we have that

$$Q_k^\delta[h] = \mathcal{S}\big(L_h Q_{k-1}^\delta[h]\big). \tag{36}$$

2. For all $\alpha < 1/4$, there exists $C < \infty$ such that for all b, h, y, k, δ as in Lemma 8,

$$\|Q_k^\delta[h]\|_{C^{1+\alpha}([0,T],L^2)} + \|Q_k^\delta[h]\|_{C^\alpha([0,T],\mathcal{D})} \leq C^k \|h\|_\infty^{k+1}. \tag{37}$$

In particular, we have that

$$\|R_k^\delta[h]\|_{C^\alpha([0,T],L^2)} \leq C^k \|h\|_\infty^{k+1}. \tag{38}$$

Proof. We first show (35). Using Riemann sums to approximate the integrals below, the closedness of the operators $\mathcal{L}_b$ and L_h as well as (32), we obtain that

$$\begin{aligned}(\frac{d}{dt} - \mathcal{L}_b) Q_0^\delta &= R_0^\delta(t) - \int_0^t \mathcal{L}_b R_0^\delta(s) ds = R_0^\delta(t) - R_0^\delta(0) - \int_0^t \mathcal{L}_b R_0^\delta(s) ds \\ &= \int_0^t \big(\frac{d}{ds} - \mathcal{L}_b\big) R_0^\delta(s) ds = L_h \int_0^t u_{b+h}^\delta(s) ds.\end{aligned} \tag{39}$$

Moreover, we have $Q_0^\delta(0) = 0$ and $Q_0^\delta \in C^{3/2}([0,T], L^2) \cap C^{1/2}([0,T], \mathcal{D})$, so that (35) follows from Proposition 4. For $k \geq 1$, (36) is proved in the same manner.

Next, we prove (37) for $k = 0$. Let $\alpha < 1/4$, $\delta > 0$, $b \in \Theta$, $\|h\| \in C_0^1$ with $\|h\|_{H^1} \leq 1$, and let us write

$$f(t) = \partial_x \Big(\int_0^t u_{b+h}^\delta(s) ds \Big).$$

In view of (35) and Proposition 4, and noting that $hf(0) = 0$, it suffices to show that $\|f\|_{C^\alpha([0,T],L^2)} \leq C$ for some uniform constant C. For all $t < t' \in [0,T]$, we have by the definition of u_{b+h}^δ and Fubini's theorem that

$$\begin{aligned} [f(t') - f(t)](x) =& \partial_x \int_t^{t'} \int_0^1 p_{s,b+h}(x,z)\varphi_\delta(z) dz ds \\ =& \partial_x \int_0^1 \Big(\int_t^{t'} p_{s,b+h}(x,z) ds \Big) \varphi_\delta(z) dz. \end{aligned}$$

For convenience, let us for now write μ for μ_{b+h} and $(\lambda_j, u_j)_{j \geq 0}$ for the eigenpairs of $\mathcal{L}_{b+h}$. Using the spectral decomposition (21) with $b + h$ in place of b and Fubini's theorem, integrating each summand separately yields that

$$\begin{aligned} [f(t') - f(t)](x) &= (t' - t)\partial_x \int_0^1 \varphi_\delta(z)\mu(z) dz \\ &\quad + \partial_x \int_0^1 \sum_{j \geq 1} \frac{1}{\lambda_j} (e^{t'\lambda_j} - e^{t\lambda_j}) u_j(x) u_j(z) \varphi_\delta(z) \mu(z) dz \\ &= \partial_x \sum_{j \geq 1} \frac{1}{\lambda_j} (e^{t'\lambda_j} - e^{t\lambda_j}) u_j(x) \langle u_j, \varphi_\delta \rangle_{L^2(\mu)}, \end{aligned} \tag{40}$$

where Fubini's theorem is applicable due to Lemma 12:

$$\sum_{j \geq 1} \Big| \frac{1}{\lambda_j} (e^{t'\lambda_j} - e^{t\lambda_j}) u_j(x) \langle u_j, \varphi_\delta \rangle_{L^2(\mu)} \Big| \lesssim \|\mu \varphi_\delta\|_{L^2} \sum_{j \geq 1} j^{-2} \|u_j\|_\infty \lesssim \sum_{j \geq 1} j^{-3/2+\varepsilon}.$$

From Lemma 14 and (40), it follows that

$$f(t') - f(t) = \partial_x \left(\mathcal{L}_{b+h}^{-1} \left(P_{t',b+h} - P_{t,b+h} \right) \varphi_\delta \right).$$

Using this, (61), the self-adjointness of $P_{t,b+h}$ with respect to $\langle \cdot, \cdot \rangle_{L^2(\mu)}$ and (65), we obtain that

$$\begin{aligned} \|f(t') - f(t)\|_{L^2} &\leq \|\mathcal{L}_{b+h}^{-1} \left(P_{t',b+h} - P_{t,b+h} \right) \varphi_\delta\|_{H^1} \\ &\lesssim \|P_{t,b+h}(P_{t'-t,b+h} - Id)\varphi_\delta\|_{H^{-1}} \\ &\lesssim \sup_{\phi \in H^1, \|\phi\|_{H^1} \leq 1} \big| \langle (P_{t'-t,b+h} - Id) P_{t,b+h} \varphi_\delta, \phi \rangle_{L^2(\mu)} \big| \\ &= \sup_{\phi \in H^1, \|\phi\|_{H^1} \leq 1} \big| \langle P_{t,b+h} \varphi_\delta, P_{t'-t,b+h}\phi - \phi \rangle_{L^2(\mu)} \big| \\ &\lesssim \sup_{t > 0} \|P_{t,b+h}\varphi_\delta\|_{L^1} \sup_{\phi \in H^1, \|\phi\|_{H^1} \leq 1} \|P_{t'-t,b+h}\phi - \phi\|_\infty \\ &\lesssim \sup_{\phi \in H^1, \|\phi\|_{H^1} \leq 1} \|P_{t'-t,b+h}\phi - \phi\|_\infty \\ &\lesssim (t' - t)^\alpha. \end{aligned}$$

Hence, Proposition 4 and (35) imply (37) for $k = 0$. Choosing C large enough and inductively using Proposition 4 and (36), we also obtain (37) for $k \geq 1$:

$$\begin{aligned}\|Q_k^\delta\|_{C^\alpha([0,T],\mathcal{D})} + \|Q_k^\delta\|_{C^{1+\alpha}([0,T],L^2)} &\leq C\|L_h Q_{k-1}^\delta\|_{C^\alpha([0,T],L^2)}\\ &\leq C\|h\|_\infty \|Q_{k-1}^\delta\|_{C^\alpha([0,T],\mathcal{D})} \leq C^k \|h\|_\infty^{k+1}.\end{aligned}$$

Finally, (38) follows upon differentiating (37) in t. □

The H^1 estimate The H^1-estimate reads as follows.

Lemma 10. *Let $k \geq 0$ be an integer and $\Delta > 0$. Then there exists $C < \infty$ such that for all b, h, y, k, δ as in Lemma 8,*

$$\|R_k^\delta(\Delta)\|_{H^1} \leq C^k \|h\|_{H^1}^k. \tag{41}$$

To prove Lemma 10, we express $R_k^\delta[h]$ using (27) and decompose the integral into times close to 0 and times bounded away from 0. The following Lemma allows us to control the respective integrals.

Lemma 11. *For any $T > 0$, there exists $C < \infty$ such that for all b, h, y, k, δ as in Lemma 8 and all $\tilde{T} \in (0, T)$, we have the estimates*

$$\Big\| \int_0^{\tilde{T}} P_{T-s} L_h R_k^\delta(s) ds \Big\|_{H^1} \leq \frac{C}{(T-\tilde{T})^{5/4}} \|h\|_{H^1} \sup_{s \in [0,\tilde{T}]} \|R_k^\delta(s)\|_{L^2}, \tag{42}$$

$$\Big\| \int_{\tilde{T}}^{T} P_{T-s} L_h R_k^\delta(s) ds \Big\|_{H^1} \leq C\|h\|_\infty \sup_{s \in [\tilde{T},T]} \|R_k^\delta(s)\|_{H^1}. \tag{43}$$

Proof. We first show (43). By Lemma 13, we can estimate the $(-\mathcal{L}_b)^{1/2}$-graph norm instead of the H^1 norm. Using Lemma 12, we have

$$\begin{aligned}&\Big\|(-\mathcal{L}_b)^{1/2} \int_{\tilde{T}}^{T} P_{T-s} L_h R_k^\delta(s) ds\Big\|_{L^2(\mu)}^2\\ &= \sum_{j=1}^\infty \Big(\int_{\tilde{T}}^{T} |\lambda_j|^{\frac{1}{2}} e^{\lambda_j (T-s)} \langle u_j, h R_k^\delta(s)' \rangle_{L^2(\mu)} ds \Big)^2\\ &\lesssim \sum_{j=1}^\infty \Big(\int_{\tilde{T}}^{T} j e^{-cj^2(T-s)} \|h R_k^\delta(s)'\|_{L^2} ds \Big)^2\\ &\lesssim \|h\|_\infty^2 \sup_{s\in[\tilde{T},T]} \|R_k^\delta(s)\|_{H^1}^2 \sum_{j=1}^\infty \Big(j \int_{\tilde{T}}^{T} e^{-cj^2(T-s)} ds \Big)^2\\ &\lesssim \|h\|_\infty^2 \sup_{s\in[\tilde{T},T]} \|R_k^\delta(s)\|_{H^1}^2 \sum_{j=1}^\infty \frac{1}{j^2}.\end{aligned}$$

A similar calculation yields that

$$\Big\| \int_{\tilde{T}}^{T} P_{T-s} L_h R_k^\delta(s) ds \Big\|_{L^2(\mu)}^2 \lesssim \|h\|_\infty^2 \sup_{s\in[\tilde{T},T]} \|R_k^\delta(s)\|_{H^1}^2 \Big(T^2 + \sum_{j=1}^\infty \frac{1}{j^4} \Big).$$

Combining the last two displays completes the proof of (43).

Next, we prove (42). Using (66) with $\alpha = 1$, the boundary condition $h(0) = h(1) = 0$ to integrate by parts and (17), we obtain

$$\begin{aligned}\Big\| \int_0^{\tilde{T}} P_{T-s} L_h R_k^\delta(s) ds \Big\|_{H^1} &\lesssim \int_0^{\tilde{T}} (T-s)^{-\frac{5}{4}} \left\| L_h R_k^\delta(s) \right\|_{H^{-1}} ds \\ &\leq (T-\tilde{T})^{-\frac{5}{4}} \int_0^{\tilde{T}} \sup_{\psi \in C^\infty, \|\psi\|_{H^1} \leq 1} \Big| \int_0^1 (\psi h)' R_k^\delta(s) \Big| ds \\ &\lesssim (T-\tilde{T})^{-\frac{5}{4}} \|h\|_{H^1} \sup_{s \in [0,\tilde{T}]} \left\| R_k^\delta(s) \right\|_{L^2}. \end{aligned}$$ □

Proof of Lemma 10. The case $k = 0$ follows from Lemma 15. For $k \geq 1$, we iteratively apply the estimates (42) and (43). We first define the points Δ_j at which we will split the integrals involved below:

$$\Delta_j := \Delta \frac{1 + j/k}{2}, \quad j = 0, ..., k, \quad \text{and} \quad \eta_k := \frac{\Delta}{2k} = \Delta_k - \Delta_{k-1}.$$

Then, using (31) and (27), we can estimate

$$\begin{aligned}\|R_k^\delta(\Delta)\|_{H^1} &\leq \Big\| \int_0^{\Delta_{k-1}} P_{\Delta-s} L_h R_{k-1}^\delta(s) ds \Big\|_{H^1} + \Big\| \int_{\Delta_{k-1}}^{\Delta} P_{\Delta-s} L_h R_{k-1}^\delta(s) ds \Big\|_{H^1} \\ &=: I + II. \end{aligned}$$

Now let C be the largest of the constants from (38), (42) and (43). From (42) with $\tilde{T} = \Delta_{k-1}$ and (38), we obtain

$$I \leq C \eta_k^{-\frac{5}{4}} \|h\|_{H^1} \sup_{s \in [0, \Delta_{k-1}]} \left\| R_{k-1}^\delta(s) \right\|_{L^2} \leq C^k \eta_k^{-\frac{5}{4}} \|h\|_{H^1}^{k+1}.$$

For the second term, we apply (43) to obtain

$$II \leq C \|h\|_\infty \sup_{s \in [\Delta_{k-1}, \Delta]} \|R_{k-1}^\delta(s)\|_{H^1}.$$

To further estimate the right hand side, we can repeat the argument for any $s \in [\Delta_{k-1}, \Delta]$:

$$\begin{aligned}\|R_{k-1}^\delta(s)\|_{H^1} &\leq \Big\| \int_0^{\Delta_{k-2}} P_{\Delta-s} L_h R_{k-2}^\delta(s) ds \Big\|_{H^1} + \Big\| \int_{\Delta_{k-2}}^{s} P_{\Delta-u} L_h R_{k-2}^\delta(u) du \Big\|_{H^1} \\ &\leq C^k \eta_k^{-\frac{5}{4}} \|h\|_{H^1}^k + C \|h\|_\infty \sup_{s \in [\Delta_{k-2}, \Delta]} \|R_{k-2}^\delta(s)\|_{H^1}. \end{aligned}$$

By iterating this argument k times, we obtain that for some larger constant $\tilde{C}$ independent of k,

$$\|R_k^\delta(\Delta)\|_{H^1} \leq k C^k \big(\frac{2k}{\Delta}\big)^{\frac{5}{4}} \|h\|_{H^1}^{k+1} + C^k \|h\|_\infty^k \sup_{s \in [\Delta/2, \Delta]} \|R_0^\delta(s)\|_{H^1} \leq \tilde{C}^k \|h\|_{H^1}^k,$$

where we used (64) in the last step. This completes the proof. □

3.3. Proof of Theorem 1

We now prove Theorem 1 by letting $\delta > 0$ in Lemma 8 tend to 0. Let us fix $b \in \Theta$, $h \in C_0^1$ with $\|h\|_{H^1} \leq 1$ and $x, y \in [0,1]$, and recall the notation $\Phi(\eta) := p_{\Delta,b+\eta h}(x,y)$ for $\eta \in \mathbb{R}$. For notational convenience, for any $\delta > 0, \eta \in \mathbb{R}$ and integer $k \geq 0$, define

$$\Phi^\delta(\eta) := u^\delta_{b+\eta h}(\Delta, x), \quad a^\delta_k := v^\delta_k[h](\Delta, x), \quad p^\delta_k(\eta) := \sum_{i=0}^{k} a^\delta_i \eta^i.$$

Then by Lemma 8 and (34), there exists $C < \infty$ such that for all $\delta > 0$, $k \geq 0$ and $\eta \in [-1, 1]$,

$$\left|\Phi^\delta(\eta) - p^\delta_k(\eta)\right| = \left|R^\delta_k[\eta h](\Delta, x)\right| \leq \|R^\delta_k[\eta h](\Delta)\|_\infty \leq C^k |\eta|^{k+1/4}. \tag{44}$$

Hence for all $\delta > 0$, on the interval $\eta \in [-\frac{1}{2C}, \frac{1}{2C}] \cap [-1, 1]$, Φ^δ is given by the power series $\Phi^\delta(\eta) = \sum_{i=0}^\infty a^\delta_i \eta^i$. We divide the rest of the proof into three steps. The first two steps imply an analogous power series for Φ, and the third proves the integral formula (8).

1. Convergence of $\Phi^\delta(\eta)$. Note that by the definition of $u^\delta_{b+\eta h}$, we have that

$$\forall \eta \in \mathbb{R}: \quad \left|\Phi^\delta(\eta) - \Phi(\eta)\right| = \left|P_{\delta,0}\big(p_{\Delta,b+\eta h}(x,\cdot)\big)(y) - p_{\Delta,b+\eta h}(x,y)\right|.$$

Moreover, by (64) we have for any $R > 0$ that

$$\sup_{x \in [0,1], \|d\|_{H^1} \leq R} \|p_{\Delta,d}(x,\cdot)\|_{H^1} < \infty. \tag{45}$$

Thus, using (65), it follows that for any $\alpha < 1/4$, there is $c < \infty$ such that for all $b \in \Theta, h \in C_0^1$ with $\|h\|_{H^1} \leq 1$ and $|\eta| \leq 1$,

$$\begin{aligned} &\left|P_{\delta,0}\big(p_{\Delta,b+\eta h}(x,\cdot)\big)(y) - p_{\Delta,b+\eta h}(x,y)\right| \\ &\quad \leq \sup_{x \in [0,1], \|d\|_{H^1} \leq B+1} \|P_{\delta,0} p_{\Delta,d}(x,\cdot) - p_{\Delta,d}(x,\cdot)\|_\infty \leq c\delta^\alpha \xrightarrow{\delta \to 0} 0. \end{aligned} \tag{46}$$

2. Convergence of a^δ_k. Fix some $\eta \neq 0$ and some sequence $\delta_n > 0$ tending to 0 as $n \to \infty$. Using (44), it is easily seen inductively that for all $k \geq 0$, the sequence $(a^{\delta_n}_k : n \in \mathbb{N})$ is bounded. Hence, by a diagonal argument there exists a subsequence $(\delta_{n_l} : l \in \mathbb{N})$ and some sequence $a_k \in \mathbb{R}$ such that for all k, $a^{\delta_{n_l}}_k \xrightarrow{l \to \infty} a_k$. Defining the polynomials

$$p_k(\eta) := \sum_{i=0}^{k} a_i \eta^i, \qquad \eta \in \mathbb{R}, \quad k = 0, 1, 2, ..., \tag{47}$$

we see that (44) still holds with Φ and p_k in place of Φ^δ and p^δ_k. Hence, Φ is analytic and $\Phi(\eta) = \sum_{k=0}^\infty a_i \eta^i$ holds for $\eta \in \left[-\frac{1}{2C}, \frac{1}{2C}\right] \cap [-1, 1]$.

3. Proof of (8). It remains to show the integral formula (8) for $\Phi'(0)$. By what precedes, we know that the constants a_0, a_1 from (47) satisfy

$$\forall \eta \in [-1,1] : |\Phi(\eta) - a_0 - \eta a_1| \leq C|\eta|^{5/4}, \quad \Phi(0) = a_0, \quad \Phi'(0) = a_1 = \lim_{\delta \to 0} a_1^\delta.$$

Moreover, by definition of $v_1^\delta[h]$, we have for all $\delta > 0$ that

$$a_1^\delta = v_1^\delta[h](\Delta, x) = \mathcal{S}(L_h u_b^\delta)(\Delta, x) = \int_0^\Delta \left[P_{\Delta-s,b} L_h u_b^\delta(s)\right](x) ds.$$

Therefore, (8) is proven if we can show that the following expression converges to 0 as $\delta \to 0$ (recall that φ_δ was defined in (29)):

$$\begin{aligned}
&\int_0^\Delta \left[P_{\Delta-s,b} L_h P_{s,b}\varphi_\delta\right](x) ds - \int_0^\Delta \left[P_{\Delta-s,b} L_h p_{s,b}(\cdot, y)\right](x) ds \\
&= \int_0^\Delta \int_0^1 p_{\Delta-s,b}(x,z) h(z) \partial_z \Big(\int_0^1 p_{s,b}(z,u)\varphi_\delta(u) du - p_{s,b}(z,y)\Big) dz ds \\
&= -\int_0^{\Delta/2} \int_0^1 \partial_z [p_{\Delta-s,b}(x,z) h(z)] \Big(\int_0^1 p_{s,b}(z,u)\varphi_\delta(u) du - p_{s,b}(z,y)\Big) dz ds \\
&\quad + \int_{\Delta/2}^\Delta \int_0^1 p_{\Delta-s,b}(x,z) h(z) \partial_z \Big(\int_0^1 p_{s,b}(z,u)\varphi_\delta(u) du - p_{s,b}(z,y)\Big) dz ds \\
&=: I + II.
\end{aligned}$$

Here we have integrated by parts and used that the boundary terms vanish due to $h(0) = h(1) = 0$. For the term I, by arguing as in (45)-(46) (with s and z in place of Δ and x), we have that

$$\forall s \in (0, \Delta/2] : \quad \sup_{z \in [0,1]} \Big| \int_0^1 p_{s,b}(z,u)\varphi_\delta(u) du - p_{s,b}(z,y)\Big| \xrightarrow{\delta \to 0} 0,$$

showing that the ds-integrand in I tends to 0 pointwise. By the heat kernel estimate (63) and (64), we can also bound the ds-integrand uniformly in δ by

$$\frac{2C}{\sqrt{s}} \|p_{\Delta-s,b}(x,\cdot)h\|_{H^1} \leq \frac{2C\|h\|_{H^1}}{\sqrt{s}} \sup_{x \in [0,1], s \in [\Delta/2, \Delta]} \|p_{s,b}(x,\cdot)\|_{H^1} < \infty,$$

where C is the constant from (63). Hence, we have by the dominated convergence theorem that $|I| \xrightarrow{\delta \to 0} 0$.

For II, we argue similarly. By Lemma 15, we have that

$$\sup_{s \in [\Delta/2, \Delta], z \in [0,1]} \|\partial_z p_{s,b}(z,\cdot)\|_{H^2} < \infty,$$

whence (65) yields that

$$\begin{aligned}&\Big|\partial_z\Big(\int_0^1 p_{s,b}(z,u)\varphi_\delta(u)du - p_{s,b}(z,y)\Big)\Big|\\&\quad= \Big|\int_0^1 \partial_z p_{s,b}(z,u)\varphi_\delta(u)du - \partial_z p_{s,b}(z,y)\Big|\\&\quad\le \big\|P_{\delta,0}\big(\partial_z p_{s,b}(z,\cdot)\big) - \partial_z p_{s,b}(z,\cdot)\big\|_\infty \xrightarrow{\delta\to 0} 0.\end{aligned}$$

Moreover, the ds-integrand is bounded by (cf. Lemma 15)

$$\frac{2C\|h\|_\infty}{\sqrt{\Delta - s}} \sup_{s\in[\Delta/2,\Delta], z\in[0,1]} \|p_{s,b}(z,\cdot)\|_{H^1},$$

such that by dominated convergence, we have $|II| \xrightarrow{\delta\to 0} 0$.

4. Spectral analysis of $\mathcal{L}_b$ and $(P_{t,b} : t \ge 0)$

In this section, we collect some properties of the generator $\mathcal{L}_b$, the differential equation related to $\mathcal{L}_b$ and the transition semigroup $(P_{t,b} : t \ge 0)$ which are needed for the proofs of Section 3. Although some results can be obtained using well-known, more general theory, our proofs are based on more or less elementary arguments, using the spectral analysis of $\mathcal{L}_b$ in Section 4.1.

4.1. Bounds on eigenvalues and eigenfunctions of $\mathcal{L}_b$

The following lemma summarizes some key properties of the eigenparis (u_j, λ_j) of $\mathcal{L}_b$. Note that the estimate (49) is an improvement on the bound in Lemma 6.6 of [12], and that (49) moreover coincides with the intuition from the eigenvalue equation $\mathcal{L}_b u_j = \lambda_j u_j$ that "two derivatives of u_j correspond to one order of growth in λ_j".

Lemma 12. *Let $s \ge 1$ be an integer and $B > 0$.*

1. *Suppose $b \in H^s \cap C_0^1$. Then for all $j \ge 0$, we have $u_j \in H^{s+2}$.*
2. *There exist $0 < C' < C < \infty$ such that for all $b \in C_0^1$ with $\|b\|_\infty \le B$,*

$$\forall j \ge 0, \qquad \lambda_j \in [-Cj^2, -C'j^2]. \tag{48}$$

   Moreover, we have $u_0 = 1, \lambda_0 = 0$.

3. *There exists $C < \infty$ such that for all $0 \le \alpha \le s+2$,*

$$\forall j \ge 1 : \sup_{b\in H^s\cap C_0^1 : \|b\|_{H^s}\le B} \|u_j\|_{H^\alpha} \le C|\lambda_j|^{\frac{\alpha}{2}}. \tag{49}$$

   In particular, we have $\|u_j\|_\infty \lesssim |\lambda_j|^{1/4+\epsilon}$ for all $\epsilon > 0$.

Proof. Using that $u_j \in \mathcal{D} \subseteq H^2$ and (18), we obtain that for all $j \geq 0$, $u_j'' = \lambda u_j - bu_j' \in H^1$. Differentiating this equation $s-1$ times and bootstrapping this argument yields that $u_j^{(s+1)} \in H^1$.

Next, we prove (48) by adapting arguments from Chapter 4 of [8]. The standard Laplacian $\mathcal{L}_0 = \Delta$ with domain $\mathcal{D}$ is a nonpositive operator, self-adjoint with respect to the L^2-inner product, with spectrum $\{-j^2\pi^2 : j = 0, 1, 2, ...\}$ and associated quadratic form

$$Q_0(f) = \langle f', f' \rangle_{L^2} \quad \text{for all } f \in Dom((-\mathcal{L}_0)^{1/2}) = H^1,$$

where the fact that $Dom((-\mathcal{L}_0)^{1/2}) = H^1$ is shown in Chapter 7 of [8]. Similarly, using (4) and integrating by parts using $f'(0) = f'(1) = 0$, we have that $\mathcal{L}_b$, with domain $\mathcal{D}$, is self-adjoint with respect to the $L^2(\mu_b)$-inner product, and that for any $f \in \mathcal{D}$, the associated quadratic form is given by

$$\begin{aligned} Q_b(f) = \langle -\mathcal{L}_b f, f \rangle_{L^2(\mu_b)} &= \int_0^1 f'^2 \mu_b dx + \int_0^1 f' f \mu_b' dx - \int_0^1 f' f b \mu_b dx \\ &= \langle f', f' \rangle_{L^2(\mu_b)}. \end{aligned} \tag{50}$$

For any finite-dimensional subspace $L \subseteq \mathcal{D}$, define

$$\lambda^{(0)}(L) := \inf_{f \in L, \|f\|_{L^2} \leq 1} -Q_0(f), \qquad \lambda^{(b)}(L) := \inf_{f \in L, \|f\|_{L^2(\mu_b)} \leq 1} -Q_b(f). \tag{51}$$

Then by Theorem 4.5.3 of [8], the eigenvalues of $\mathcal{L}_0$ and $\mathcal{L}_b$ are given by

$$\lambda_j^{(0)} = \sup_{L \subseteq \mathcal{D}, \dim L \leq j} \lambda^{(0)}(L) = -j^2\pi^2, \qquad \lambda_j^{(b)} = \sup_{L \subseteq \mathcal{D}, \dim L \leq j} \lambda^{(b)}(L) \tag{52}$$

respectively. This, combined with (50) and (24), yields (48).

We now prove (49). Iterating the equation $\mathcal{L}_b u_j = \lambda_j u_j$, we have

$$\begin{aligned} \lambda_j^2 u_j = \mathcal{L}_b^2 u_j &= (u_j'' + bu_j')'' + b(u_j'' + bu_j')' \\ &= u_j^{(4)} + b'' u_j' + 2b' u_j'' + b u_j''' + b u_j''' + b b' u_j' + b^2 u_j''. \end{aligned}$$

Note that in each summand above, except for the first one, the sum of the orders of all derivatives is at most 3. This generalizes to $n \geq 3$, in that there exist polynomials $P_{n,m}$ such that

$$\lambda_j^n u_j = \mathcal{L}_b^n u_j = u_j^{(2n)} + \sum_{m=1}^{2n-1} P_{n,m}(b, b', ..., b^{(2n-2)}) u_j^{(m)}, \tag{53}$$

for which one can check the following properties by induction:

1. For all $n \geq 1$ and $m \leq 2n-1$, $P_{n,m}$ has degree at most n.
2. The only summand in (53) with factor $b^{(2n-2)}$ is $u_j' b^{(2n-2)}$.

For the odd order derivatives of u_j, there similarly exist polynomials $\tilde{P}_{n,m}$ of degree at most n such that

$$\begin{aligned} u_j^{(2n+1)} &= \Big(\mathcal{L}_b^n u_j - \sum_{m=1}^{2n-1} P_{n,m}(b, b', ..., b^{(2n-2)}) u_j^{(m)}\Big)' \\ &= \lambda_j^n u_j' - \sum_{m=1}^{2n} \tilde{P}_{n,m}(b, b', ..., b^{(2n-1)}) u_j^{(m)}, \end{aligned} \tag{54}$$

where the only summand containing the factor $b^{(2n-1)}$ is $u_j' b^{(2n-1)}$.

We now use these facts to show (49) by an induction argument, consisting of the base case and two induction steps.

Base Case $\alpha \leq 2$*:* To show (49) for all $\alpha \leq 2$, it suffices to prove the case $\alpha = 2$, as the case $\alpha \in (0, 2)$ then follows from $\|u_j\|_{L^2(\mu)} = 1$ and (17). We also note that the estimate for $\|u_j\|_\infty$ then follows by the Sobolev embedding (19). The case $\alpha = 2$ follows immediately from (57) and (48):

$$\|u_j\|_{H^2}^2 \simeq \|\mathcal{L}_b u_j\|_{L^2(\mu)}^2 + \|u_j\|_{L^2(\mu)}^2 = (\lambda_j^2 + 1)\|u_j\|_{L^2(\mu)}^2 = \lambda_j^2 + 1 \lesssim \lambda_j^2, \quad j \geq 1.$$

Induction step $2n \to 2n+1$*:* Assume that for some integer n, (49) holds for all $\alpha \leq 2n < s+2$. Then, using (54), the Sobolev embedding $C^{2n-2} \subseteq H^s$ (note that $s \geq 2n-1$) and the induction hypothesis, we obtain

$$\|u_j^{(2n+1)}\|_{L^2} \lesssim |\lambda_j|^n \|u_j'\|_{L^2} + \|b^{(2n-1)}\|_{L^2} \|u_j'\|_\infty + \|b\|_{C^{2n-2}}^n \|u_j\|_{H^{2n}} \lesssim |\lambda_j|^{n+\frac{1}{2}}.$$

The non-integer case $\alpha \in (2n, 2n+1)$ follows by interpolation.

Induction step $2n-1 \to 2n$*:* Similarly, using (53), the embedding $C^{2n-3} \subseteq H^s$ (note that $s \geq 2n-2$) and the induction hypothesis, we have

$$\|u_j^{(2n)}\|_{L^2} \lesssim |\lambda_j|^n + \|b^{(2n-2)}\|_{L^2} \|u_j'\|_\infty + \|b\|_{C^{2n-3}}^n \|u_j\|_{H^{2n-1}} \lesssim |\lambda_j|^n,$$

and the non-integer case $\alpha \in (2n-1, 2n)$ again follows by interpolation. □

4.2. Characterisation of Sobolev norms in terms of (λ_j, u_j)

Using Lemma 12, we now prove that the graph norms of the non-negative self-adjoint operators $(-\mathcal{L}_b)^\theta$, $\theta \in \{0, \frac{1}{2}, 1\}$, on their respective domains, are equivalent to standard Sobolev norms. Let $\ell^2 = \ell^2(\mathbb{N} \cup \{0\})$ denote the usual space of square-summable sequences. For any Banach space $(X, \|\cdot\|_X)$ and linear operator $T : D \to X$ with domain $D \subseteq X$, we denote the graph norm of T by

$$\|x\|_T := (\|x\|_X^2 + \|Tx\|_X^2)^{1/2}, \qquad x \in D.$$

Lemma 13. *1. Let* $\theta \in [0, 1]$*. Then for any* $f \in L^2$*, we have*

$$f \in Dom\big((-\mathcal{L}_b)^\theta\big) \iff \sum_{j=0}^\infty (1 + |\lambda_j|^{2\theta}) |\langle f, u_j \rangle_{L^2(\mu)}|^2 < \infty \tag{55}$$

and for any $f \in Dom\big((-\mathcal{L}_b)^\theta\big)$*, we have*

$$(-\mathcal{L}_b)^\theta f = \sum_{j=1}^{\infty} (-\lambda_j)^\theta \langle f, u_j\rangle_{L^2(\mu)} u_j. \tag{56}$$

2. There exists $0 < C < \infty$ *such that for any* $\theta \in \{0, \frac{1}{2}, 1\}$ *, we have*

$$C^{-1}\|f\|_{H^{2\theta}} \le \|f\|_{(-\mathcal{L}_b)^\theta} \le C\|f\|_{H^{2\theta}}, \quad f \in Dom\big((-\mathcal{L}_b)^{\theta/2}\big). \tag{57}$$

3. There exists $0 < C < \infty$ *such that for all* $f \in L^2$*,*

$$C^{-1}\|f\|_{H^{-1}} \le \Big\|\Big(\frac{\langle f, u_j\rangle_{L^2(\mu)}}{\sqrt{1+|\lambda_j|}} : j \ge 0\Big)\Big\|_{\ell^2} \le C\|f\|_{H^{-1}} \tag{58}$$

Proof. 1. We first prove (55) for $\theta = 1$. Define the dense linear subspace

$$D := \bigcup_{n=0}^{\infty} \operatorname{span}\{u_j : j = 0, ..., n\} \subseteq L^2(\mu_b).$$

Then by Lemma 1.2.2 in [8], we know that the restriction of $\mathcal{L}_b$ to D, which we shall denote by $\mathcal{L}_b^D$, is an essentially self-adjoint operator on $L^2(\mu_b)$. Moreover, under the unitary operator

$$U : L^2(\mu_b) \to \ell^2, \quad f \mapsto \big(\langle f, u_j\rangle_{L^2(\mu_b)} : j \ge 0\big),$$

$\mathcal{L}_b^D$ is unitarily equivalent to the essentially self-adjoint multiplication operator $M^D : (a_j : j \ge 0) \mapsto (\lambda_j a_j : j \ge 0)$ on ℓ^2 with domain

$$U(D) = \{a \in \ell^2 : a_j = 0 \text{ for all } j \text{ large enough}\}.$$

Thus, the unique self-adjoint extentions of both operators (cf. [8], Theorem 1.2.7), which we denote by $\mathcal{L}_b$ and M, are also unitarily equivalent. Hence, for all $f \in L^2(\mu_b)$,

$$f \in \mathcal{D} \iff \sum_{j=0}^{\infty} (1+\lambda_j^2)|\langle f, u_j\rangle_{L^2(\mu_b)}|^2 < \infty$$

(The above condition defines the domain of the self-adjoint extension of M^D, see [8], Lemma 1.3.1), which proves (55) for $\theta = 1$. To see (55) for $\theta \in [0, 1)$, we note that the fractional power $(-\mathcal{L}_b)^\theta$ is unitarily equivalent to multiplication with $\big(|\lambda_j|^\theta : j \ge 0\big)$, and that $f \in \operatorname{Dom}\big((-\mathcal{L}_b)^\theta\big)$ iff

$$Uf \in \operatorname{Dom}(M^\theta) = \Big\{f \in L^2 : \sum_{j=0}^{\infty} (1+|\lambda_j|^{2\theta})|\langle f, u_j\rangle_{L^2(\mu)}|^2 < \infty\Big\}.$$

2. We now show (57). For $\theta = 0$, there is nothing to prove. For $\theta = 1/2$, note that by Theorem 7.2.1 in [8] and (50), we have $\operatorname{Dom}\big((-\mathcal{L}_b)^{1/2}\big) = H^1$ and

$$\forall f \in H^1 : \|f\|^2_{\mathcal{L}_b^{1/2}} = \|f\|^2_{L^2(\mu_b)} + \langle \mathcal{L}_b^{1/2} f, \mathcal{L}_b^{1/2} f\rangle_{L^2(\mu_b)} = \|f\|^2_{H^1(\mu_b)}.$$

The case $\theta = 1/2$ now follows from (24). Finally, let $\theta = 1$. It is clear that $\|f\|^2_{\mathcal{L}_b} \lesssim \|f\|^2_{H^2}$, so that it remains to show $\|f\|^2_{H^2} \lesssim \|f\|^2_{\mathcal{L}_b}$. For this, we use Cauchy's inequality with ϵ to obtain that for some c_1,

$$\|\mathcal{L}_b f\|^2_{L^2} = \|f''\|^2_{L^2} + 2\langle f'', bf'\rangle_{L^2} + \|bf'\|^2_{L^2} \geq \frac{1}{2}\|f''\|^2_{L^2} - c_1\|f'\|^2_{L^2}.$$

Hence, integrating by parts and using Cauchy's inequality with ϵ again yields that for some c_2,

$$\begin{aligned}\|f''\|^2_{L^2} &\leq 2\|\mathcal{L}_b f\|^2_{L^2} + 2c_1\|f'\|^2_{L^2} \leq 2\|\mathcal{L}_b f\|^2_{L^2} + 2c_1\|f\|_{L^2}\|f''\|_{L^2} \\ &\leq 2\|\mathcal{L}_b f\|^2_{L^2} + c_2\|f\|^2_{L^2} + \frac{1}{2}\|f''\|^2_{L^2},\end{aligned}$$

proving that $\|f\|^2_{H^2} \lesssim \|f\|^2_{\mathcal{L}_b}$.

3. For any $f \in L^2$ and any test function $\psi \in H^1$, let us write $f_j = \langle f, u_j\rangle_{L^2(\mu_b)}$ and $\psi_j = \langle \psi, u_j\rangle_{L^2(\mu_b)}$, $j \geq 0$ respectively. Then by (56)-(57), we have

$$\begin{aligned}\|f\|_{H^{-1}} &\simeq \sup_{\psi\in H^1, \|\psi\|_{H^1}\leq 1} \left|\langle f, \psi\rangle_{L^2(\mu_b)}\right| = \sup_{\psi\in H^1, \|\psi\|_{H^1}\leq 1} \Big|\sum_{j=0}^{\infty} f_j\psi_j\Big| \\ &\simeq \sup_{\psi\in L^2, \|\psi\|_{L^2}\leq 1} \Big|\sum_{j=0}^{\infty} f_j\,(1+|\lambda_j|)^{-1/2}\,\psi_j\Big| \qquad (59) \\ &\simeq \left\|\left(f_j\,(1+|\lambda_j|)^{-1/2} : j\geq 0\right)\right\|_{\ell^2}.\end{aligned}$$

□

4.3. Basic norm estimates for the one-dimensional Neumann problem

From the preceding Lemma, we can immediately derive some basic properties of the (elliptic) boundary value problem

$$\mathcal{L}_b u = f \quad \text{on } (0,1), \qquad u'(0) = u'(1) = 0 \tag{60}$$

needed in the proof of Lemma 9. Let us denote the orthogonal complement of the first eigenfunction $u_0 \equiv 1$ of $\mathcal{L}_b$ in $L^2(\mu_b)$ by

$$u_0^\perp = \Big\{f \in L^2 : \int f d\mu = 0\Big\}.$$

Lemma 14. *For every $f \in u_0^\perp$, there exists a unique function $u \in \mathcal{D}\cap u_0^\perp$ such that $\mathcal{L}_b u = f$, for which we use the notation $u = \mathcal{L}_b^{-1} f$. Moreover, for every $B > 0$ there exists $C < \infty$ such that for all $b \in C_0^1$ with $\|b\|_\infty \leq B$ and $f \in u_0^\perp$,*

$$\|u\|_{H^s} \leq C\|f\|_{H^{s-2}} \qquad \text{for } s \in \{0,1,2\}. \tag{61}$$

Proof. It follows immediately from the domain characterisation (55) and the spectral representation (22) that $\mathcal{L}_b$ is a one-to-one map from $\mathcal{D}\cap u_0^\perp$ to $L^2\cap u_0^\perp$, and that $\mathcal{L}_b^{-1}$ is unitarily equivalent to multiplication by $(\lambda_j^{-1}\mathbb{1}_{j\geq 1} : j \geq 0)$ in the spectral domain, so that the $L^2 \to L^2$ norm of $\mathcal{L}_b^{-1}$ is finite. Hence, for $s = 2$, the estimate (61) follows from (57):

$$\left\|\mathcal{L}_b^{-1} f\right\|_{H^2}^2 \simeq \left\|\mathcal{L}_b\mathcal{L}_b^{-1} f\right\|_{L_2}^2 + \|\mathcal{L}_b^{-1} f\|_{L^2}^2 \simeq \|f\|_{L_2}^2.$$

The case $s = 0$ is obtained by duality. Using that $\mathcal{L}_b^{-1}$ is self-adjoint on $u_0^\perp$ and the previous case $s = 2$, we have that

$$\begin{aligned}\|\mathcal{L}_b^{-1} f\|_{L^2(\mu_b)} &= \sup_{\phi\in u_0^\perp, \|\phi\|_{L^2}\leq 1} \Big| \int_0^1 \mathcal{L}_b^{-1} f\phi d\mu\Big| = \sup_{\phi\in u_0^\perp, \|\phi\|_{L^2}\leq 1} \Big| \int_0^1 f\mathcal{L}_b^{-1}\phi d\mu\Big| \\ &\lesssim \|f\|_{H^{-2}}.\end{aligned}$$

Finally, for $s = 1$, Lemma 13 implies that

$$\left\|\mathcal{L}_b^{-1} f\right\|_{H^1}^2 \simeq \sum_{j=1}^{\infty}(1+|\lambda_j|)\Big|\frac{\langle f, u_j\rangle_{L^2(\mu_b)}}{\lambda_j}\Big|^2 \lesssim \sum_{j=1}^{\infty}\frac{|\langle f, u_j\rangle_{L^2(\mu_b)}|^2}{1+|\lambda_j|} \lesssim \|f\|_{H^{-1}}^2. \quad \square$$

4.4. Estimates on $p_{t,b}(\cdot,\cdot)$ and $P_{t,b}$

Using Lemmata 12 and 13, we now collect some basic (partially well-known) results about the Lebesgue transition densities $p_{t,b}(\cdot,\cdot)$ (Lemma 15) and the semigroup $P_{t,b}$ (Lemma 16). Recall that they were defined in (5) and (6).

Lemma 15. *Let $s \geq 1$ be an integer, $t_0 > 0$ and $B > 0$. Then we have the following.*

1. *There exist constants $0 < C < C' < \infty$ such that for all $t \geq t_0$, $b \in C_0^1$ with $\|b\|_{C^1} \leq B$ and $x, y \in [0,1]$,*

$$C \leq p_{t,b}(x,y) \leq C'. \tag{62}$$

2. *There exists $C < \infty$ such that for all $t \in (0,1]$ and $b \in C_0^1$ with $\|b\|_\infty \leq B$,*

$$\|p_{t,b}(x,y)\|_\infty \leq Ct^{-\frac{1}{2}}, \quad x, y \in [0,1]. \tag{63}$$

3. *For each $n \leq s+2$, $m \leq s$ and $n', m' \leq s+1$,*

$$\begin{aligned}&\sup_{t\geq t_0}\sup_{y\in[0,1]}\sup_{b\in C_0^1\cap H^s:\|b\|_{H^s}\leq B}\|\partial_x^n\partial_y^m p_{t,b}(\cdot,y)\|_{L^2} < \infty \\ &\sup_{t\geq t_0}\sup_{x\in[0,1]}\sup_{b\in C_0^1\cap H^s:\|b\|_{H^s}\leq B}\|\partial_x^{n'}\partial_y^{m'} p_{t,b}(x,\cdot)\|_{L^2} < \infty.\end{aligned} \tag{64}$$

Proof. For a proof of (62), we refer to Proposition 9 in [24] and for a proof of (63), we refer to Theorem 2.12 in [7]. Let us now prove the first part of (64); the second is obtained analogously. Let $n \leq s+2$, $m \leq s$. Then (4) yields that

$$\sup_{\|b\|_{H^s} \leq B} \|\mu_b\|_{H^{s+1}} < \infty.$$

Using the multiplicative inequality (18), the spectral decomposition (21) and Lemma 12, we have

$$\begin{aligned}
\|\partial_x^n \partial_y^m p_{t,b}(\cdot, y)\|_{L^2} &\leq \sum_{j=0}^{\infty} e^{t\lambda_j} \|u_j^{(n)}\|_{L^2} |\left(u_j \mu_b\right)^{(m)}(y)| \\
&\leq \sum_{j=0}^{\infty} e^{t_0\lambda_j} \|u_j^{(n)}\|_{L^2} \|(u_j\mu_b)^{(m)}\|_\infty \lesssim \sum_{j=0}^{\infty} e^{t_0\lambda_j} \|u_j\|_{H^{s+2}} \|u_j\|_{H^{s+1}} \|\mu_b\|_{H^{s+1}} \\
&\lesssim \sum_{j=0}^{\infty} e^{-cj^2} |\lambda_j|^{\frac{s+2}{2}+\frac{s+1}{2}} \lesssim \sum_{j=0}^{\infty} e^{-cj^2} j^{2s+3} < \infty,
\end{aligned}$$

where Lemma 12 implies that the constants above are uniform in $\|b\|_{H^s} \leq B$. $\square$

Finally, we collect some properties of $(P_{t,b} : t \geq 0)$.

Lemma 16. *Let $B > 0$. The following holds.*

1. *For all $b \in C_0^1$, $p \in [1, \infty]$ and $f \in L^p$, we have $\|P_{t,b}f\|_{L^p(\mu)} \leq \|f\|_{L^p(\mu)}$.*
2. *For every $\epsilon > 0$, there exists $C < \infty$ such that for all $b \in C_0^1$ with $\|b\|_\infty \leq B$, $f \in H^1$ and $t > 0$,*

$$\|P_{t,b}f - f\|_{L^2} \leq Ct^{1/2}\|f\|_{H^1} \quad \text{and} \quad \|P_{t,b}f - f\|_\infty \leq Ct^{1/4-\epsilon}\|f\|_{H^1}. \tag{65}$$

   In particular, we have that $H^1 \subseteq \mathcal{D}(1/2)$, with $D(1/2)$ defined by (26).
3. *Let $s \geq 1$ be an integer. Then for all $t > 0$, $b \in H^s \cap C_0^1$ with $\|b\|_{H^s} \leq B$ and $f \in L^2$, we have $P_{t,b}f \in H^{s+2}$. Moreover, there exists $C < \infty$ such that for all such t, b, f and all $\alpha \leq s+2$,*

$$\|P_{t,b}f\|_{H^\alpha} \leq C(1 + t^{-\frac{\alpha}{2}-\frac{3}{4}})\|f\|_{H^{-1}}. \tag{66}$$

Proof. 1. For the case $p = 1$, we have by Fubini's theorem that

$$\begin{aligned}
\int_0^1 \Big| \int_0^1 p_{t,b}(x,z)f(z)dz \Big| d\mu(x) &\leq \int_0^1 \int_0^1 p_{t,b}(x,z)d\mu(x)|f(z)|dz \\
&= \int_0^1 |f(z)|d\mu(z).
\end{aligned}$$

For the case $p = \infty$, we observe that for all $x \in [0, 1]$

$$|P_{t,b}f(x)| \leq \|f\|_\infty \int p_{t,b}(x,z)dz = \|f\|_\infty.$$

The case $p \in (1, \infty)$ follows by the Riesz-Thorin interpolation theorem.

2. To prove the first part of (65), let $f \in H^1 = \mathrm{Dom}\big((-\mathcal{L}_b)^{1/2}\big)$. By the 1/2-Hölder continuity of $x \mapsto e^x$ on $(-\infty, 0]$ and Lemma 13, we have that for all $t \geq 0$,

$$\begin{aligned}\|P_{t,b}f - f\|^2_{L^2(\mu_b)} &= \sum_{j=1}^{\infty} \left(e^{\lambda_j t} - 1\right)^2 |\langle f, u_j\rangle_{L^2(\mu_b)}|^2 \\ &\lesssim t \sum_{j=1}^{\infty} |\lambda_j| |\langle f, u_j\rangle_{L^2(\mu_b)}|^2 \lesssim t\|f\|^2_{H^1}.\end{aligned}$$

The second estimate in (65) now follows from the $H^1 \to H^1$ boundedness of $P_{t,b}$, the embedding (19), the interpolation inequality (17) and the first part of (65). Indeed, we have for any $\varepsilon > 0$ that

$$\|P_{t,b}f - f\|_\infty \lesssim \|P_{t,b}f - f\|^{(1-4\varepsilon)/2}_{L^2} \|P_{t,b}f - f\|^{(1+4\varepsilon)/2}_{H^1} \lesssim t^{1/4-\varepsilon}\|f\|_{H^1}.$$

3. By Lemma 12, we have that $u_j \in H^{s+2}$ for all $j \geq 0$. Using the spectral representation (23), Lemma 12, Lemma 13 and Cauchy-Schwarz, we have

$$\begin{aligned}\|P_{t,b}f\|_{H^\alpha} &\lesssim \sum_{j=0}^{\infty} e^{\lambda_j t} \|u_j\|_{H^\alpha} (1 + |\lambda_j|)^{1/2} \frac{|\langle f, u_j\rangle_{L^2(\mu_b)}|}{(1 + |\lambda_j|)^{1/2}} \\ &\lesssim \Big(\sum_{j=0}^{\infty} e^{2\lambda_j t} (1 + |\lambda_j|)^{\alpha+1}\Big)^{1/2} \|f\|_{H^{-1}} \\ &\lesssim \Big(1 + \int_0^\infty e^{-2cx^2 t} x^{2(\alpha+1)} dx\Big)^{1/2} \|f\|_{H^{-1}} \\ &\lesssim \Big(1 + t^{-\alpha-1-\frac{1}{2}}\Big)^{1/2} \|f\|_{H^{-1}} \\ &\lesssim \big(1 + t^{-\frac{\alpha}{2}-\frac{3}{4}}\big)\|f\|_{H^{-1}}.\end{aligned}$$

□

Acknowledgments

I am very grateful to Richard Nickl for suggesting to pursue this project, for repeatedly proofreading the manuscript and for sharing with me some of the important ideas in this paper, including the one of applying a PDE approach. I would like to thank the Cantab Capital Institute for the Mathematics of Information, as well as Richard Nickl's ERC grant No. 647812 (UQMSI), for supporting my PhD.

References

[1] K. Abraham. Nonparametric Bayesian posterior contraction rates for scalar diffusions with high-frequency data. *Bernoulli, to appear, arXiv:1802.05635*, 2018.

[2] D. Bakry, I. Gentil, and M. Ledoux. *Analysis and Geometry of Markov Diffusion Operators*. Springer, 2013. MR3155209
[3] R. Bass. *Diffusions and Elliptic Operators*. Springer, 1998. MR1483890
[4] B. Brown. Martingale Central Limit Theorems. *Ann. Math. Statist.*, 42:59–66, 1971. MR0290428
[5] I. Castillo and R. Nickl. Nonparametric Bernstein-von Mises theorems in Gaussian White Noise. *Annals of Statistics*, 41:1999–2028, 2013. MR3127856
[6] I. Castillo and R. Nickl. On the Bernstein - von Mises phenomenon for nonparametric Bayes procedures. *Annals of Statistics*, 42:1941–1969, 2014. MR3262473
[7] J. Chorowski. *Statistics for diffusion processes with low and high-frequency observations*. PhD Thesis, 2016.
[8] E. B. Davies. *Spectral Theory and Differential Operators*. Cambridge University Press, 1995. MR1349825
[9] I. Gihman and A. Skorohod. *Stochastic differential equations*. Springer, 1972. MR0346904
[10] E. Giné and R. Nickl. *Mathematical Foundations of Infinite-Dimensional statistical models*. Cambridge University Press, 2016. MR3588285
[11] E. Gobet. LAN property for ergodic diffusions with discrete observations. *Annales de l'Institut Henri Poincaré B*, 38:711–737, 2002. MR1931584
[12] E. Gobet, M. Hoffmann, and M. Reiß. Nonparametric estimation of scalar diffusions based on low frequency data. *Annals of Statistics*, 32, 2004.
[13] S. Gugushvili and P. Spreij. Nonparametric Bayesian drift estimation for multidimensional stochastic differential equations. *Lith. Math. J.*, 54:127–141, 2014. MR3212631
[14] L. P. Hansen, J. A. Scheinkman, and N. Touzi. Spectral methods for identifying scalar diffusions. *Journal of Econometrics*, 86:1–32, 1998. MR1650012
[15] M. Hoffmann, A. Munk, and J. Schmidt-Hieber. Adaptive wavelet estimation of the diffusion coefficient under additive error measurements. *Annals of the Institute H. Poincaré*, 48:1186–1216, 2012. MR3052408
[16] J. Koskela, D. Spano, and P. A. Jenkins. Consistency of Bayesian nonparametric inference for discretely observed jump diffusions. *arXiv:1506.04709*, 2015.
[17] J. Lions and E. Magenes. *Non-Homogeneous Boundary Value Problems and Applications, Vol. 1*. Springer, 1972. MR0350178
[18] Y. Lu. On the bernstein-von mises theorem for high dimensional nonlinear bayesian inverse problems. *arXiv:1706.00289*, 2017.
[19] A. Lunardi. *Analytic Semigroups and Optimal Regularity for Parabolic Problems*. Birkhäuser, 1995. MR1944167
[20] F. Monard, R. Nickl, and G. Paternain. Efficient nonparametric Bayesian inference for X-ray transforms. *Annals of Statistics*, to appear. MR3909962
[21] R. Nickl. Bernstein - von Mises theorems for statistical inverse problems I: Schrödinger equation. *Journal of the European Mathematical Society*, to appear. MR2514176
[22] R. Nickl and K. Ray. Nonparametric statistical inference for drift vector

fields of multi-dimensional diffusions. *arXiv:1810.01702*, 2018.

[23] R. Nickl and J. Söhl. Bernstein - von Mises theorems for statistical inverse problems II: compound Poisson processes. *arXiv preprint*, 2017.

[24] R. Nickl and J. Söhl. Nonparametric Bayesian posterior contraction rates for discretely observed scalar diffusions. *Annals of Statistics*, 45:1664–1693, 2017. MR3670192

[25] O. Papaspiliopoulos, Y. Pokern, G. O. Roberts, and A. M. Stuart. Nonparametric estimation of diffusions: a differential equations approach. *Biometrika*, 99:511–531, 2012. MR2966767

[26] F. van der Meulen, M. Schauer, and H. van Zanten. Reversible jump MCMC for nonparametric drift estimation for diffusion processes. *Comput. Statist. Data Anal.*, 71:615–632, 2014. MR3131993

[27] A. W. van der Vaart. *Asymptotic Statistics*. Cambridge University Press, 1998. MR1652247

[28] J. van Waaij and H. van Zanten. Gaussian process methods for one-dimensional diffusions: Optimal rates and adaptation. *Electron. J. Stat.*, 10:628–645, 2013. MR3471991