SUPPLEMENT TO: NONPARAMETRIC STATISTICAL INFERENCE FOR DRIFT VECTOR FIELDS OF MULTI-DIMENSIONAL DIFFUSIONS

By Richard Nickl and Kolyan Ray

University of Cambridge and King's College London

In this supplement, we present many of the technical details used in the proofs of the main article [10] along with a review of the relevant PDE material. To avoid confusion, we continue the numbering scheme (e.g. sections, equations, lemmas, etc.) from the main document in a linear way. Hence references to items that are not in this supplement concern the main manuscript, and vice versa.

4. Bayesian semi-parametric techniques for diffusions. We show here how techniques developed in [3–5] for Gaussian white noise and i.i.d. density models extend to the multi-dimensional diffusion case.

4.1. *Asymptotic expansion of the posterior Laplace transform.* We start with the following basic 'LAN expansion' for ℓ_T from (8).

Lemma 6. *Suppose $b_0 \in C^{(d/2+\kappa)\vee 1}(\mathbb{T}^d), h \in H^{d/2+\kappa}(\mathbb{T}^d), \kappa > 0$. Then*

$$\ell_T(b_0 + h/\sqrt{T}) - \ell_T(b_0) = W_T(h) - \frac{1}{2T}\int_0^T \|h(X_t)\|^2 dt,$$

where, as $T \to \infty$ and under P_{b_0},

$$W_T(h) \equiv \frac{1}{\sqrt{T}}\int_0^T h(X_t).dW_t \to^d N(0, \|h\|_{\mu_0}^2), \ \frac{1}{T}\int_0^T \|h(X_t)\|^2 dt \to^P \|h\|_{\mu_0}^2.$$

Proof. Using (1) with $b = b_0$ and (8),

$$\begin{aligned}\ell_T(b_0 + h/\sqrt{T}) - \ell_T(b_0) &= \frac{1}{\sqrt{T}}\int_0^T h(X_t).dX_t - \frac{1}{\sqrt{T}}\int_0^T b_0(X_t).h(X_t)dt \\ &\quad - \frac{1}{2T}\int_0^T \|h(X_t)\|^2 dt \\ &= \frac{1}{\sqrt{T}}\int_0^T h(X_t).dW_t - \frac{1}{2T}\int_0^T \|h(X_t)\|^2 dt.\end{aligned}$$

Since $x \mapsto \|x\|^2$ is a smooth map, the function $f_h(x) = \|h(x)\|^2 - \|h\|_{\mu_0}^2 \in L^2_{\mu_0}(\mathbb{T}^d) \cap H^{d/2+\kappa}(\mathbb{T}^d) \subset C(\mathbb{T}^d)$. In particular, $LL^{-1}[f_h] = f_h$ where $L^{-1} = L_{b_0}^{-1}$ is the inverse of the generator L constructed in Lemma 11 below. Moreover, by that Lemma and the Sobolev embedding theorem, $L^{-1}[f_h] \in H^{d/2+\kappa+2} \subset C^2$. By Itô's formula (Theorem 39.3 in [1]),

$$\int_0^T f_h(X_t)dt = \int_0^T LL^{-1}[f_h](X_t)dt$$
$$= L^{-1}[f_h](X_T) - L^{-1}[f_h](X_0) - \int_0^T \nabla L^{-1}[f_h](X_t).dW_t.$$

Since $L^{-1}[f_h] \in C^2$, the first term on the right-hand side is $O(1)$, while the second term satisfies

$$E_{b_0}\Big[\int_0^T \nabla L^{-1}[f_h](X_t).dW_t\Big]^2 = E_{b_0}\int_0^T \|\nabla L^{-1}[f_h](X_t)\|^2 dt \lesssim T\|L^{-1}[f_h]\|_{C^1}^2$$

so that $T^{-1}\int_0^T f_h(X_t)dt \to 0$ in $L^2(P_{b_0})$, and hence also in P_{b_0}-probability. Next, set $M_T^h = \int_0^T h(X_t).dW_t$, so that $(M_T^h)_{T\geq 0}$ is a continuous local L^2-martingale with quadratic variation $[M^h]_T = \int_0^T \|h(X_t)\|^2 dt$. By what precedes $T^{-1}[M^h]_T - \|h\|_{\mu_0}^2 = T^{-1}\int_0^T f_h(X_t)dt \to 0$ in $L^2(P_{b_0})$ as $T \to \infty$. Applying the martingale central limit theorem (p.338f. in [6]),

$$T^{-1/2}M_T^h \to^d N(0, \|h\|_{\mu_0}^2) \tag{46}$$

as $T \to \infty$, completing the proof. □

A key result for our proofs is the following expansion of the Laplace transform of the posterior distribution $\Pi^{\mathcal{D}_T}(\cdot|X^T)$ arising from a 'localised' prior $\Pi^{\mathcal{D}_T}$ for the choices of $\mathcal{D}_T$ from Lemma 5.

PROPOSITION 2. *Suppose $b_0 \in C^s(\mathbb{T}^d) \cap H^s(\mathbb{T}^d)$, $s > \max(d/2, 1)$, and consider the Gaussian prior Π from (11) with $2^J \approx T^{\frac{1}{2a+d}}$ and $\sigma_l = 2^{-l(\alpha+d/2)}$ for $a > \max(d-1, 1/2)$ and $0 \leq \alpha \leq a$. Let $\Gamma_T \subset V_J^{\otimes d}$ be a set of functions admitting envelopes as in (27) and let $D_T \subset V_J^{\otimes d}$ denote the set (28) for this choice of Γ_T and arbitrary $M > 0$. For $u \in \mathbb{R}$, $b \in V_J^{\otimes d}$ and fixed $\gamma \in \Gamma_T$, define the perturbations*

$$b_u = b_u(T, \gamma) = b - \tfrac{u}{\sqrt{T}}\gamma \in V_J^{\otimes d}. \tag{47}$$

For any measurable function $G : L^2(\mathbb{T}^d) \to \mathbb{R}$, write

$$E^{\Pi^{D_T}}[e^{u\sqrt{T}G(b)}|X^T] = e^{\Lambda_T(u)} \frac{\int_{D_T} e^{S_T(b)+\ell_T(b_u)} d\Pi(b)}{\int_{D_T} e^{\ell_T(b)} d\Pi(b)}, \tag{48}$$

for some Λ_T to be determined and where

$$S_T(b) = u\sqrt{T}\left(G(b) - \langle b - b_0, \gamma\rangle_{\mu_0}\right).$$

(i) If for some $\kappa > 0$ (or $\kappa = 0$ if $d = 1$),

$$R_T := 2^{J[d+(d/2+\kappa-1)_+]} M_T \varepsilon_T |\Gamma_T|_2 \left(1 + \sqrt{\log(1/(M_T\varepsilon_T))} + \sqrt{\log(1/|\Gamma_T|_2)}\right)$$

satisfies $R_T \to 0$ as $T \to \infty$, then we can take

$$\Lambda_T(u) = \frac{u}{\sqrt{T}} \int_0^T \gamma(X_t).dW_t + \frac{u^2}{2T}\int_0^T \|\gamma(X_t)\|^2 dt + ur_T, \ u \in \mathbb{R},$$

in (48) with $r_T = O_{P_{b_0}}(R_T) = o_{P_{b_0}}(1)$ uniformly over $\gamma \in \Gamma_T$.
(ii) Furthermore,

$$E_{b_0} \sup_{\gamma\in\Gamma_T} \left| \frac{1}{T}\int_0^T \|\gamma(X_t)\|^2 dt - \|\gamma\|_{\mu_0}^2 \right| \lesssim \tilde{R}_T,$$

where

$$\tilde{R}_T := T^{-1/2} 2^{J[d+(d/2+\kappa-1)_+]} |\Gamma_T|_2^2 \left(1 + \sqrt{\log(1/|\Gamma_T|_2)}\right)$$

for any $\kappa > 0$ ($\kappa = 0$ if $d = 1$). In particular, if both $R_T, \tilde{R}_T \to 0$, then we can take

$$\Lambda_T(u) = \frac{u}{\sqrt{T}} \int_0^T \gamma(X_t).dW_t + \frac{u^2}{2}\|\gamma\|_{\mu_0}^2 + ur_T + u^2\tilde{r}_T, \ u \in \mathbb{R},$$

in (48) with $r_T = O_{P_{b_0}}(R_T) = o_{P_{b_0}}(1)$ and $\tilde{r}_T = O_{P_{b_0}}(\tilde{R}_T) = o_{P_{b_0}}(1)$ uniformly over $\gamma \in \Gamma_T$.
(iii) Parts (i) and (ii) remain true if D_T is replaced by $\bar{D}_T$ from (29) and if M_T is replaced by $\bar{M}_T$ in the definition of R_T.

PROOF. *(i)* For $\gamma = (\gamma_1, \dots, \gamma_d) \in \Gamma_T$,

$$\begin{aligned} E^{\Pi^{D_T}}[e^{u\sqrt{T}G(b)}|X^T] &= E^{\Pi^{D_T}}[e^{S_T(b)+u\sqrt{T}\langle b-b_0,\gamma\rangle_{\mu_0}}|X^T] \\ &= Z_T^{-1}\int_{D_T} e^{S_T(b)+u\sqrt{T}\langle b-b_0,\gamma\rangle_{\mu_0}+\ell_T(b_u)+\ell_T(b)-\ell_T(b_u)} d\Pi(b), \end{aligned}$$

with $Z_T = \int_{D_T} e^{\ell_T(b)} d\Pi(b)$. Define the empirical process

$$\mathbb{G}_T[h] = \sqrt{T}\left(\frac{1}{T}\int_0^T h(X_t)dt - \int_{\mathbb{T}^d} h d\mu_0\right), \quad h \in L^2(\mathbb{T}^d).$$

Using the LAN expansion from Lemma 6,

$$\begin{aligned}
\ell_T(b) - \ell_T(b_u) &= \frac{u}{\sqrt{T}}\int_0^T \gamma(X_t).dW_t - \frac{u}{\sqrt{T}}\int_0^T [b(X_t) - b_0(X_t)].\gamma(X_t)dt \\
&\quad + \frac{u^2}{2T}\int_0^T \|\gamma(X_t)\|^2 dt \\
&= \frac{u}{\sqrt{T}}\int_0^T \gamma(X_t).dW_t - u\mathbb{G}_T[(b-b_0).\gamma] - u\sqrt{T}\langle b - b_0, \gamma\rangle_{\mu_0} \\
&\quad + \frac{u^2}{2T}\int_0^T \|\gamma(X_t)\|^2 dt.
\end{aligned}$$

The Laplace transform from the first equation of this proof therefore equals

$$e^{\frac{u}{\sqrt{T}}\int_0^T \gamma(X_t).dW_t + \frac{u^2}{2T}\int_0^T \|\gamma(X_t)\|^2 dt} Z_T^{-1}\int_{D_T} e^{-u\mathbb{G}_T[(b-b_0).\gamma]} e^{S_T(b)+\ell_T(b_u)} d\Pi(b).$$

We use Lemma 1 to control the empirical process term uniformly over $b \in D_T$, $\gamma \in \Gamma_T$. Set

$$\mathcal{F}_T = \left\{ f_{b,\gamma} := (b-b_0).\gamma - \int_{\mathbb{T}^d}(b-b_0).\gamma d\mu_0 : b \in D_T, \gamma \in \Gamma_T \right\},$$

which is a subset of $L^2_{\mu_0}(\mathbb{T}^d) \cap H^{d/2+\kappa}(\mathbb{T}^d)$ for $0 < \kappa < s - d/2$ since $\Gamma_T \subset V_J^{\otimes d} \subset H^p$ for any $p \le S$. Suppose $d \ge 2$. Lemma 3 with $p = d/2 + \kappa - 1$

gives that for any $0<\kappa<s-d/2+1$, $b,\bar b\in D_T$ and $\gamma,\bar\gamma\in\Gamma_T$,

$$\begin{aligned}
&d_L(f_{b,\gamma},f_{\bar b,\bar\gamma})\lesssim \|f_{b,\gamma}-f_{\bar b,\bar\gamma}\|_{H^{d/2+\kappa-1}}\\
&\leq \sum_{j=1}^d \Big\|(b_j-\bar b_j)\gamma_j+(\bar b_j-b_{0,j})(\gamma_j-\bar\gamma_j)\\
&\qquad\qquad -\langle b_j-\bar b_j,\gamma_j\rangle_{\mu_0}-\langle \bar b_j-b_{0,j},\gamma_j-\bar\gamma_j\rangle_{\mu_0}\Big\|_{H^{d/2+\kappa-1}}\\
&\lesssim \sum_{j=1}^d 2^{J(d+\kappa-1)}\|b_j-\bar b_j\|_{L^2}\|\gamma_j\|_{L^2}\\
&\quad+\sum_{j=1}^d \|(\bar b_j-P_{V_J}b_{0,j})(\gamma_j-\bar\gamma_j)-\langle \bar b_j-P_{V_J}b_{0,j},\gamma_j-\bar\gamma_j\rangle_{\mu_0}\|_{H^{d/2+\kappa-1}}\\
&\quad+\sum_{j=1}^d \|(P_{V_J}b_{0,j}-b_{0,j})(\gamma_j-\bar\gamma_j)-\langle P_{V_J}b_{0,j}-b_{0,j},\gamma_j-\bar\gamma_j\rangle_{\mu_0}\|_{H^{d/2+\kappa-1}}.
\end{aligned}$$

The first sum above is bounded by $C2^{J(d+\kappa-1)}|\Gamma_T|_2\|b-\bar b\|_{L^2}$, while by Lemma 3 the second sum is bounded by $C\sum_{j=1}^d 2^{J(d+\kappa-1)}M_T\varepsilon_T\|\gamma-\bar\gamma\|_{L^2}$. Using Lemma 2, that $b_0\in C^s\cap H^s$ for $s>d/2$ and (16)-(17), the third sum is bounded by

$$\begin{aligned}
&\sum_{j=1}^d \|P_{V_J}b_{0,j}-b_{0,j}\|_{L^\infty}\|\gamma_j-\bar\gamma_j\|_{H^{d/2+\kappa-1}}+\|P_{V_J}b_{0,j}-b_{0,j}\|_{H^{d/2+\kappa-1}}\|\gamma_j-\bar\gamma_j\|_{L^\infty}\\
&\lesssim \sum_{j=1}^d \Big(2^{-Js}2^{J(d/2+\kappa-1)}\|\gamma_j-\bar\gamma_j\|_{L^2}+2^{J(d+\kappa-1-s)}\|\gamma_j-\bar\gamma_j\|_{L^2}\Big)\\
&\lesssim 2^{J(d+\kappa-1)}M_T\varepsilon_T\|\gamma-\bar\gamma\|_{L^2},
\end{aligned}$$

using again that $\Gamma_T\subset V_J^{\otimes d}$. Summarizing,

$$d_L(f_{b,\gamma},f_{\bar b,\bar\gamma})\lesssim 2^{J(d+\kappa-1)}(|\Gamma_T|_2\|b-\bar b\|_{L^2}+M_T\varepsilon_T\|\gamma-\bar\gamma\|_{L^2}).$$

In particular, $\mathcal{F}_T$ has d_L-diameter $D_{\mathcal{F}_T}\lesssim 2^{J(d+\kappa-1)}M_T\varepsilon_T|\Gamma_T|_2=o(R_T)=o(1)$. Since $D_T,\Gamma_T\subset(V_J^{\otimes d},\|\cdot\|_{L^2})$ where $v_J=dim(V_J)=O(2^{Jd})$, applying Proposition 4.3.34 of [9] yields

$$\begin{aligned}
&N(\mathcal{F}_{T,},d_L,\tau)\\
&\leq N(D_T,c2^{J(d+\kappa-1)}|\Gamma_T|_2\|\cdot\|_{L^2},\tau/2)N(\Gamma_T,c2^{J(d+\kappa-1)}M_T\varepsilon_T\|\cdot\|_{L^2},\tau/2)\\
&\leq (C2^{J(d+\kappa-1)}|\Gamma_T|_2/\tau)^{dv_J}(C2^{J(d+\kappa-1)}M_T\varepsilon_T/\tau)^{dv_J}
\end{aligned}$$

for some $c, C > 0$. Recall the inequality

$$\int_0^a \sqrt{\log(A/x)}dx \leq \frac{2\log A}{2\log A - 1} a\sqrt{\log(A/a)} \leq 4a\sqrt{\log(A/a)}$$

for any $A \geq 2$ and $0 < a \leq 1$ (p. 190 of [9]). Using the last two displays and that $D_{\mathcal{F}_T} \to 0$, $\int_0^{D_{\mathcal{F}_T}} \sqrt{\log 2N(\mathcal{F}_T, d_L, \tau)}d\tau$ is bounded by a multiple of

$$v_J^{1/2} \int_0^{D_{\mathcal{F}_T}} \Big[\sqrt{\frac{\log([C2^{J(d+\kappa-1)}|\Gamma_T|_2] \vee 2}{\tau}} + \sqrt{\frac{\log([C2^{J(d+\kappa-1)}M_T\varepsilon_T] \vee 2}{\tau}}\Big] d\tau$$
$$\lesssim 2^{Jd/2} D_{\mathcal{F}_T}\sqrt{\log([C2^{J(d+\kappa-1)}|\Gamma_T|_2] \vee 2/D_{\mathcal{F}_T})}$$
$$+ 2^{Jd/2} D_{\mathcal{F}_T}\sqrt{\log([C2^{J(d+\kappa-1)}M_T\varepsilon_T] \vee 2/D_{\mathcal{F}_T})}.$$

Taking $D_{\mathcal{F}_T} \approx 2^{J(d+\kappa-1)}M_T\varepsilon_T|\Gamma_T|_2$ for $\kappa > 0$ arbitrarily small, one can therefore bound the quantity $J(\mathcal{F}_T, d_L, D_{\mathcal{F}_T})$ in Lemma 1 by

$$2^{J(3d/2+\kappa-1)}M_T\varepsilon_T|\Gamma_T|_2(1 + \sqrt{\log(1/(M_T\varepsilon_T))} + \sqrt{\log(1/|\Gamma_T|_2)}) = R_T.$$

Using the Sobolev embedding theorem, Lemma 11, Lemma 3 and similar computations to the above, we see that

$$\sup_{f_{b,\gamma}\in\mathcal{F}_T} \|L^{-1}[f_{b,\gamma}]\|_\infty \lesssim \sup_{f_{b,\gamma}\in\mathcal{F}_T} \|f_{b,\gamma}\|_{H^{(d/2+\kappa-2)_+}} \lesssim 2^{J[d/2+(d/2+\kappa-2)_+]}M_T\varepsilon_T|\Gamma_T|_2$$

is also $o(R_T)$. Substituting these bounds into Lemma 1 yields

$$E_{b_0} \sup_{b\in D_T,\gamma\in\Gamma_T} |\mathbb{G}_T[(b-b_0).\gamma]| \lesssim R_T \to 0,$$

proving the first statement. The case $d = 1$ is proved similarly, using instead the simpler bound

$$d_L(f_{b,\gamma}, f_{\bar{b},\bar{\gamma}}) \lesssim 2^{J/2}|\Gamma_T|_2\|b-\bar{b}\|_{L^2} + 2^{J/2}M_T\varepsilon_T\|\gamma-\bar{\gamma}\|_{L^2}.$$

(*ii*) Since $x \mapsto \|x\|^2$ is a smooth map, the function $g_\gamma(x) = \|\gamma(x)\|^2 - \int_{\mathbb{T}^d}\|\gamma(y)\|^2 d\mu_0(y) \in L^2_{\mu_0}(\mathbb{T}^d) \cap H^{d/2+\kappa}$ for $\kappa > 0$. Since $\gamma \in V_J^{\otimes d}$, Lemma 3 with $p = (d/2+\kappa-1)_+$ gives that for any $\kappa > 0$ small enough and $\gamma, \bar{\gamma} \in \Gamma_T$,

$$\begin{aligned} d_L(g_\gamma, g_{\bar{\gamma}}) &\lesssim \|g_\gamma - g_{\bar{\gamma}}\|_{H^{(d/2+\kappa-1)_+}} \\ &\lesssim \sum_{j=1}^d \|\gamma_j^2 - \bar{\gamma}_j^2 - \int_{\mathbb{T}^d}(\gamma_j^2 - \bar{\gamma}_j^2)d\mu_0\|_{H^{(d/2+\kappa-1)_+}} \\ &\lesssim \sum_{j=1}^d 2^{J[d/2+(d/2+\kappa-1)_+]}\|\gamma_j - \bar{\gamma}_j\|_{L^2}\|\gamma_j + \bar{\gamma}_j\|_{L^2} \\ &\lesssim 2^{J[d/2+(d/2+\kappa-1)_+]}|\Gamma_T|_2\|\gamma-\bar{\gamma}\|_{L^2}. \end{aligned}$$

In particular, $\mathcal{G}_T = \{g_\gamma : \gamma \in \Gamma_T\} \cup \{0\}$ has d_L-diameter

$$D_{\mathcal{G}_T} \lesssim 2^{J[d/2+(d/2+\kappa-1)_+]}|\Gamma_T|_2^2.$$

Using the same arguments as above, one deduces

$$\begin{aligned} N(\mathcal{G}_T, d_L, \tau) &\leq N(\Gamma_T, 2^{J[d/2+(d/2+\kappa-1)_+]}|\Gamma_T|_2\|\cdot\|_2, \tau) \\ &\leq (C2^{J[d/2+(d/2+\kappa-1)_+]}|\Gamma_T|_2/\tau)^{dv_J} \end{aligned}$$

and hence

$$J(\mathcal{G}_T, d_L, D_{\mathcal{G}_T}) \lesssim 2^{J[d+(d/2+\kappa-1)_+]}|\Gamma_T|_2^2 \left(1 + \sqrt{\log(1/|\Gamma_T|_2)}\right) = \sqrt{T}\tilde{R}_T.$$

In exactly the same way,

$$\sup_{g_\gamma \in \mathcal{G}_T} \|L^{-1}[g_\gamma]\|_\infty \lesssim 2^{J[d/2+(d/2+\kappa-2)_+]}\|\gamma\|_{L^2}^2 \leq 2^{J[d/2+(d/2+\kappa-2)_+]}|\Gamma_T|_2^2.$$

Applying Lemma 1 thus gives

$$\begin{aligned} &E_{b_0} \sup_{\gamma\in\Gamma_T} \left| \frac{1}{T}\int_0^T \|\gamma(X_t)\|^2 dt - \int_{\mathbb{T}^d} \|\gamma(x)\|^2 d\mu_0(x) \right| \\ &= \frac{1}{\sqrt{T}} E_{b_0} \sup_{g_\gamma \in \mathcal{G}_T} |\mathbb{G}_T(g_\gamma)| \lesssim \frac{1}{\sqrt{T}} 2^{J[d+(d/2+\kappa-1)_+]}|\Gamma_T|_2^2 \left(1 + \sqrt{\log(1/|\Gamma_T|_2)}\right) \end{aligned}$$

which equals $\tilde{R}_T$. Finally, the proof of Part (iii) follows in the same way, using that $\|\cdot\|_{L^2} \leq \|\cdot\|_\infty$ and replacing M_T by $\bar{M}_T$. □

4.2. *Change of measure.* In this section, let $\Pi = \Pi_T$ be the prior from (11). Using the lower bound for the small-ball probability (23) in the proof of Theorem 1, the proof of the following lemma is similar to the one of Theorem 8.20 in [7], and hence omitted.

LEMMA 7. *Suppose $b_0 \in C^s$ for $s > 0$. Then there exists a finite constant $C = C(b_0) > 0$ such that if B_T are measurable sets satisfying $\Pi_T(B_T) = o(e^{-CT\varepsilon_T^2})$ for $\varepsilon_T = T^{-\frac{a\wedge s}{2a+d}}(\log T)$, then $E_{b_0}\Pi_T(B_T|X^T) \to 0$.*

We now bound the ratio of Gaussian integrals from (48) in Proposition 2.

LEMMA 8. *(i) Suppose $b_0 \in C^s(\mathbb{T}^d) \cap H^s(\mathbb{T}^d)$ for some $s > \max(d/2, 1)$. Let $2^J \approx T^{\frac{1}{2a+d}}$ for $a > \max(d-1, 1/2)$ and $\varepsilon_T = T^{-\frac{a\wedge s}{2a+d}}(\log T)$. Let D_T be as in (28) for a choice of $\Gamma_T \subset V_J^{\otimes d}$ whose envelopes from (27) satisfy*

$|\Gamma_T|_2 = O(\sqrt{T}\varepsilon_T)$ *and* $\varepsilon_T \sigma_{\Gamma_T} \to 0$ *as* $T \to \infty$. *Then for all* $M > 0$ *large enough,* $\Pi(D_T|X^T) = 1 - o_{P_{b_0}}(1)$. *Moreover for* b_u *as in (47) and all* $u \in \mathbb{R}$,

$$\frac{\int_{D_T} e^{\ell_T(b_u)} d\Pi(b)}{\int_{D_T} e^{\ell_T(b)} d\Pi(b)} = 1 + \zeta_T(u) \leq C_T e^{r_T u^2}, \tag{49}$$

where $\zeta_T(u) = o_{P_{b_0}}(1)$ *for every fixed* u, *where both* $C_T = O_{P_{b_0}}(1)$ *and non-random* $r_T = o(1)$ *are independent of* u, *and all terms are uniform over* $\gamma \in \Gamma_T$.

(ii) The conclusion of Part (i) remains true for $d \leq 4$ *and under the conditions of Theorem 2 if* D_T *is replaced by the set* $\bar{D}_T$ *from (29) with* $\bar{M}_T = (\log T)^{\delta - 1}, \delta > 5/2$, *and if in addition* $|\Gamma_T|_2 = O(1)$ *as* $T \to \infty$.

PROOF. (i) The set of b's satisfying the L^2-constraint in the definition (28) of D_T has posterior probability tending to one by Theorem 1. Recall that by definition of the RKHS, $\langle b, \gamma \rangle_{\mathbb{H}} \sim N(0, \|\gamma\|_{\mathbb{H}}^2)$ for $b \sim \Pi$ and $\gamma \in \mathbb{H} = V_J^{\otimes d}$. By Dudley's metric entropy inequality (Section 2.3 in [9]) applied to the Gaussian process $(\langle b, \gamma \rangle_{\mathbb{H}} : \gamma \in \Gamma_T)$ indexed by bounded subsets of the finite-dimensional space $V_J^{\otimes d}$ (with covering numbers bounded in Proposition 4.3.34 in [9]), we have

$$E^{\Pi} \sup_{\gamma \in \Gamma_T} |\langle b, \gamma \rangle_{\mathbb{H}}| \lesssim 2^{Jd/2} \sigma_{\Gamma_T} \sqrt{\log(1/\sigma_{\Gamma_T})} \leq M_0 \sqrt{T} \varepsilon_T \sigma_{\Gamma_T} \tag{50}$$

for some $M_0 > 0$, since we may always take $\sigma_{\Gamma_T} \geq 1$. By the Borell-Sudakov-Tsirelson inequality (Theorem 2.5.8 of [9]), for $M > M_0$,

$$\begin{aligned} &\Pi\big(\sup_{\gamma \in \Gamma_T} |\langle b, \gamma \rangle_{\mathbb{H}}| > M\sqrt{T}\varepsilon_T \sigma_{\Gamma_T}\big) \\ &\quad \leq \Pi\big(\sup_{\gamma \in \Gamma_T} |\langle b, \gamma \rangle_{\mathbb{H}}| > E^{\Pi} \sup_{\gamma \in \Gamma_T} |\langle b, \gamma \rangle_{\mathbb{H}}| + (M - M_0)\sqrt{T}\varepsilon_T \sigma_{\Gamma_T}\big) \\ &\quad \leq e^{-\frac{1}{2}(M - M_0)^2 T \varepsilon_T^2}. \end{aligned}$$

Taking $M > 0$ large enough, the posterior probability of the set in the last display is then $o_{P_{b_0}}(1)$ by Lemma 7. In summary this establishes that $\Pi(D_T^c|X_T) = o_{P_{b_0}}(1)$.

We now establish (49). Letting Π_u denote the law of b_u under the prior and applying the Cameron-Martin theorem (Theorem 2.6.13 of [9]), the desired ratio equals

$$\frac{\int_{D_{T,u}} e^{\ell_T(g)} \frac{d\Pi_u}{d\Pi}(g) d\Pi(g)}{\int_{D_T} e^{\ell_T(g)} d\Pi(g)} = \frac{\int_{D_{T,u}} e^{\ell_T(g)} e^{-\frac{u}{\sqrt{T}}\langle \gamma, g \rangle_{\mathbb{H}} - \frac{u^2}{2T}\|\gamma\|_{\mathbb{H}}^2} d\Pi(g)}{\int_{D_T} e^{\ell_T(g)} d\Pi(g)}, \tag{51}$$

where $D_{T,u} = \{g = b_u : b \in D_T\}$. By the definition of D_T,

$$\sup_{g\in D_{T,u},\gamma\in\Gamma_T} \left| \frac{u}{\sqrt{T}}\langle \gamma, g\rangle_{\mathbb{H}} + \frac{u^2}{2T}\|\gamma\|_{\mathbb{H}}^2 \right| \leq \frac{|u|}{\sqrt{T}} \sup_{b\in D_T,\gamma\in\Gamma_T} |\langle \gamma, b - \frac{u\gamma}{\sqrt{T}}\rangle_{\mathbb{H}}| + \frac{u^2\sigma_{\Gamma_T}^2}{2T}$$
$$\leq |u| M\varepsilon_T\sigma_{\Gamma_T} + \frac{3u^2\sigma_{\Gamma_T}^2}{2T}.$$

We thus upper bound (51) by

$$e^{\tilde{r}_T u^2 + \tilde{r}'_T|u|} \frac{\int_{D_{T,u}} e^{\ell_T(g)} d\Pi(g)}{\int_{D_T} e^{\ell_T(g)} d\Pi(g)} = e^{\tilde{r}_T u^2 + \tilde{r}'_T|u|} \frac{\Pi(D_{T,u}|X^T)}{\Pi(D_T|X^T)}, \tag{52}$$

where $\tilde{r}_T, \tilde{r}'_T \to 0$ are non-random and uniform over $\gamma \in \Gamma_T$. Since $\alpha|u| \leq \alpha^2u^2 + 1$ for all $\alpha \geq 0$ and $u \in \mathbb{R}$, the exponential in the last display is bounded by $e^{r_T u^2 + 1}$ for all $u \in \mathbb{R}$, where $r_T = \tilde{r}_T + (\tilde{r}'_T)^2 = 3\sigma_{\Gamma_T}^2/(2T) + M^2\varepsilon_T^2\sigma_{\Gamma_T}^2 \to 0$. Since we have already shown that $\Pi(D_T|X^T) = 1 - o_{P_{b_0}}(1)$ and the posterior probability $\Pi(D_{T,u}|X^T)$ is bounded by one, the inequality in (49) follows.

Turning to the exact asymptotics for fixed $u \in \mathbb{R}$, the right hand side in (52) equals $\Pi(D_{T,u}|X^T)(1+o_{P_{b_0}}(1))$, and (51) can be lower bounded by (52) with $e^{\tilde{r}_T u^2 + \tilde{r}'_T|u|}$ replaced by $e^{-\tilde{r}_T u^2 - \tilde{r}'_T|u|}$. It consequently suffices to prove $\Pi(D_{T,u}|X^T) = 1 - o_{P_{b_0}}(1)$. Now

$$\Pi(D_{T,u}^c|X^T) \leq \Pi(g \in V_J^{\otimes d} : \|g + \tfrac{u}{\sqrt{T}}\gamma - b_0\|_{\mu_0} > M_T\varepsilon_T|X^T)$$
$$+ \Pi\Big(g \in V_J^{\otimes d} : \sup_{\gamma\in\Gamma_T} |\langle g + \tfrac{u}{\sqrt{T}}\gamma, \gamma\rangle_{\mathbb{H}}| > M\sqrt{T}\varepsilon_T\sigma_{\Gamma_T}\Big|X^T\Big).$$

By Proposition 1, $\|\frac{u}{\sqrt{T}}\gamma\|_{\mu_0} \lesssim \frac{|u|}{\sqrt{T}}|\Gamma_T|_2 = O(\varepsilon_T) = o(M_T\varepsilon_T)$, so that the first posterior probability tends to zero by Theorem 1. Using (50), that $\sigma_{\Gamma_T}^2/\sqrt{T} = o(\sqrt{T}\varepsilon_T\sigma_{\Gamma_T})$ and the Borell-Sudakov-Tsirelson inequality (Theorem 2.5.8 of [9]), the prior probability of the second event is bounded by

$$\Pi\Big(g : \sup_{\gamma\in\Gamma_T} |\langle g, \gamma\rangle_{\mathbb{H}}| + \frac{|u|\sigma_{\Gamma_T}^2}{\sqrt{T}} > E^{\Pi} \sup_{\gamma\in\Gamma_T} |\langle b, \gamma\rangle_{\mathbb{H}}| + (M - M_0)\sqrt{T}\varepsilon_T\sigma_{\Gamma_T}\Big)$$
$$\leq e^{-\frac{1}{4}(M-M_0)^2T\varepsilon_T^2}$$

for T large enough depending on u. For $M > 0$ large enough, Lemma 7 then yields that the posterior probability of this last set is $o_{P_{b_0}}(1)$, which shows $\Pi(D_{T,u}|X^T) = 1 - o_{P_{b_0}}(1)$ as desired.

Part (ii) is proved in the same way using Theorem 2 (whose proof only relies on Part (i) of the present lemma) to ensure that $\Pi(\bar{D}_T|X^T) \to^{P_{b_0}} 1$ as $T \to \infty$, and upon noting that

$$\|\tfrac{u}{\sqrt{T}}\gamma\|_\infty \lesssim \tfrac{2^{Jd/2}|u|}{\sqrt{T}}|\Gamma_T|_2 = O(\varepsilon_T) = o(\bar{M}_T\varepsilon_T)$$

since $|\Gamma_T|_2 = O(1)$ as $T \to \infty$. □

4.3. *An approximation lemma.*

LEMMA 9. *Suppose $b_0 \in C^s(\mathbb{T}^d)$ for some $s \geq 1$ and let $\lambda \leq J$, $1 \leq j \leq d$ and $b \in V_J^{\otimes d}$. If $a_\lambda = 2^{\lambda d/2}2^{-Jd/2}(\log T)^{-\eta}$ for some $\eta \geq 0$, then*

$$\begin{aligned} &a_\lambda|\langle \mu_0(b_j - b_{0,j}), \Phi_{\lambda,k}/\mu_0 - P_{V_J}[\Phi_{\lambda,k}/\mu_0]\rangle_{L^2}| \\ &\quad \leq C(2^{-2J}\|b_j - b_{0,j}\|_{L^2} + 2^{-J(s+d/2+1)})(\log T)^{-\eta}. \end{aligned}$$

If instead $a_\lambda = 2^{\lambda d/2}2^{-J(d-2)}(\log T)^{-\eta}$ for some $\eta \geq 0$, then

$$\begin{aligned} &a_\lambda|\langle \mu_0(b_j - b_{0,j}), \Phi_{\lambda,k}/\mu_0 - P_{V_J}[\Phi_{\lambda,k}/\mu_0]\rangle_{L^2}| \\ &\quad \leq C(2^{-Jd/2}\|b_j - b_{0,j}\|_{L^2} + 2^{-J(s+d-1)})(\log T)^{-\eta}. \end{aligned}$$

Finally,

$$\Big|\Big\langle \mu_0(b_j - b_{0,j}), \frac{\Phi_{\lambda,k}}{\mu_0} - P_{V_J}\Big[\frac{\Phi_{\lambda,k}}{\mu_0}\Big]\Big\rangle_{L^2}\Big| \leq C2^{-\lambda d/2}\big(2^{-2J}\|b_j - b_{0,j}\|_\infty + 2^{-J(s+1)}\big).$$

In all cases, the constant C depends only on b_0, Φ and d.

PROOF. By the triangle inequality, the desired quantity is bounded by

$$\begin{aligned} &a_\lambda|\langle \mu_0(b_j - P_{V_J}b_{0,j}), \Phi_{\lambda,k}/\mu_0 - P_{V_J}[\Phi_{\lambda,k}/\mu_0]\rangle_{L^2}| \\ &\quad + a_\lambda|\langle \mu_0(b_{0,j} - P_{V_J}b_{0,j}), \Phi_{\lambda,k}/\mu_0 - P_{V_J}[\Phi_{\lambda,k}/\mu_0]\rangle_{L^2}| =: (I) + (II). \end{aligned}$$

By Parseval's identity,

$$\begin{aligned} (I) &= a_\lambda\left|\sum_{l>J}\sum_r \langle \mu_0(b_j - P_{V_J}b_{0,j}), \Phi_{l,r}\rangle_{L^2}\langle \Phi_{\lambda,k}/\mu_0, \Phi_{l,r}\rangle_{L^2}\right| \\ &\leq a_\lambda \sum_{l>J}\max_r|\langle \mu_0(b_j - P_{V_J}b_{0,j}), \Phi_{l,r}\rangle_{L^2}|\sum_r|\langle \Phi_{\lambda,k}/\mu_0, \Phi_{l,r}\rangle_{L^2}|. \end{aligned} \tag{53}$$

By Proposition 1 we know that μ_0 has finite Lipschitz norm $\|\mu_0\|_{\mathrm{Lip}}$. Let $x_{l,r} \in I_{l,r} := \mathrm{supp}(\Phi_{l,r})$ and note that $\mathrm{diam}(I_{l,r}) = O(2^{-l})$ by construction of the wavelets. Using that $b_j - P_{V_J} b_{0,j} \in V_J$ is orthogonal to $\Phi_{l,r}$ for any $l > J$, $\|\Phi_{l,r}\|_{L^1} \lesssim 2^{-ld/2}$ and (17),

$$\begin{aligned}
&|\langle \mu_0(b_j - P_{V_J} b_{0,j}), \Phi_{l,r}\rangle_{L^2}| \\
&= \left| \int_{\mathbb{T}^d} (\mu_0(x) - \mu_0(x_{l,r}))(b_j(x) - P_{V_J} b_{0,j}(x))\Phi_{l,r}(x)dx \right| \\
&\leq \|\mu_0\|_{\mathrm{Lip}} \mathrm{diam}(I_{l,r}) \int_{\mathbb{T}^d} |b_j(x) - P_{V_J} b_{0,j}(x)||\Phi_{l,r}(x)|dx \\
&\leq C(b_0, \Phi) 2^{-l} \|b_j - P_{V_J} b_{0,j}\|_\infty \|\Phi_{l,r}\|_{L^1} \qquad (54)\\
&\leq C(b_0, \Phi) 2^{Jd/2} \|b_j - P_{V_J} b_{0,j}\|_{L^2} 2^{-l(d/2+1)}.
\end{aligned}$$

Moreover, using the following standard properties of wavelet bases,

$$\sup_x \sum_r |\Phi_{l,r}(x)| \lesssim 2^{ld/2}, l \geq 0, \quad \langle \Phi_{\lambda,k}, \Phi_{l,r}\rangle_{L^2} = 0, \quad \lambda \leq J < l,$$

we deduce

$$\begin{aligned}
\sum_r |\langle \Phi_{\lambda,k}/\mu_0, \Phi_{l,r}\rangle_{L^2}| &= \sum_r \left| \int_{\mathbb{T}^d} \left(\frac{1}{\mu_0(x)} - \frac{1}{\mu_0(x_{l,r})} \right) \Phi_{\lambda,k}(x)\Phi_{l,r}(x)dx \right| \\
&\leq \|1/\mu_0\|_{\mathrm{Lip}} \mathrm{diam}(I_{l,r}) \int_{\mathbb{T}^d} |\Phi_{\lambda,k}(x)| \sum_r |\Phi_{l,r}(x)|dx \\
&\leq C(b_0, \Phi) 2^{l(d/2-1)} 2^{-\lambda d/2}. \qquad (55)
\end{aligned}$$

Substituting (54) and (55) into (53) yields

$$(I) \lesssim a_\lambda 2^{Jd/2} 2^{-\lambda d/2} \|b_j - P_{V_J} b_{0,j}\|_{L^2} \sum_{l>J} 2^{-2l} \lesssim 2^{-2J} (\log T)^{-\eta} \|b_j - b_{0,j}\|_{L^2}$$

as desired. Next expanding (II) as in (53) and using (55),

$$\begin{aligned}
(II) &\leq C(b_0, \Phi) a_\lambda \sum_{l>J} \max_r |\langle \mu_0(b_{0,j} - P_{V_J}[b_{0,j}]), \Phi_{l,r}\rangle_{L^2}| 2^{l(d/2-1)} 2^{-\lambda d/2} \\
&\lesssim a_\lambda 2^{-Js} 2^{-\lambda d/2} \sum_{l>J} 2^{-l} \lesssim 2^{-J(s+1+d/2)} (\log T)^{-\eta},
\end{aligned}$$

where we have used Hölder's inequality and $\|b_{0,j} - P_{V_J}[b_0, j]\|_\infty \lesssim 2^{-Js}$. The last two displays imply the first inequality in the lemma. If instead $a_\lambda = 2^{\lambda d/2} 2^{-J(d-2)} (\log T)^{-\eta}$, then substituting this value into the final bounds for (I) and (II) gives the required result. The final inequality of the lemma is proved in the same way, but by using the penultimate $\|\cdot\|_\infty$ bound in (54) rather than the $\|\cdot\|_{L^2}$ bound in the last line. Taking $a_\lambda = 1$ in the rest of the argument gives the result. □

4.4. *Proof of Lemma 5.*

LEMMA 10. *Suppose $b_0 \in C^s(\mathbb{T}^d)$ for some $s > d/2$. Then for Γ_T as in (30) and the RKHS norm $\|\cdot\|_{\mathbb{H}}$ defined in (12) with $\sigma_l = 2^{-l(\alpha+d/2)}$ for $\alpha \geq 0$, we can take the envelopes from (27) as*

$$|\Gamma_T|_2 = C(d,\mu_0)\max_{\lambda\leq J} a_\lambda, \quad \sigma_{\Gamma_T} = C(d,\mu_0,\Phi)2^{J(\alpha+d/2)}\max_{\lambda\leq J} a_\lambda.$$

PROOF. Using Proposition 1,

$$\|\tilde{\Phi}_{\lambda,k,j}\|_{L^2}^2 = \|P_{V_J}[a_\lambda \Phi_{\lambda,k}/\mu_0]\|_{L^2}^2 \leq a_\lambda^2 \|1/\mu_0\|_\infty^2 \|\Phi_{\lambda,k}\|_{L^2}^2 \lesssim a_\lambda^2.$$

To prove the second bound, note that Proposition 1 implies that $1/\mu_0$ has finite Lipschitz norm $\|1/\mu_0\|_{\mathrm{Lip}}$ on $\mathbb{T}^d$. Let $x_{l,r} \in \mathrm{supp}(\Phi_{l,r})$ and note that $\mathrm{diam}(\mathrm{supp}(\Phi_{l,r})) = O(2^{-l})$ by construction of the wavelets. Using the orthogonality of the wavelets and Hölder's inequality, for $(l,r) \neq (\lambda,k)$ with $\lambda \leq J$,

$$\begin{aligned}|\langle P_{V_J}[\Phi_{\lambda,k}/\mu_0], \Phi_{l,r}\rangle_{L^2}| &= \left|\int_{\mathbb{T}^d} \Phi_{\lambda,k}(x)\Phi_{l,r}(x)\left(\frac{1}{\mu_0(x)} - \frac{1}{\mu_0(x_{l,r})}\right)dx\right| \\ &\lesssim \|1/\mu_0\|_{\mathrm{Lip}} 2^{-\max(l,\lambda)} \int_{\mathbb{T}^d} |\Phi_{\lambda,k}(x)||\Phi_{l,r}(x)|dx \\ &\lesssim 2^{-\max(l,\lambda)} 2^{-|l-\lambda|d/2},\end{aligned}$$

while $|\langle P_{V_J}[\Phi_{\lambda,k}/\mu_0], \Phi_{\lambda,k}\rangle_{L^2}| \leq \|1/\mu_0\|_\infty \|\Phi_{\lambda,k}\|_{L^2}^2 \lesssim 1$. Note that for $l \leq \lambda$, there are a constant number of wavelets $\Phi_{l,r}$ intersecting $\mathrm{supp}(P_{V_J}[\Phi_{\lambda,k}/\mu_0])$, while for $l \geq \lambda$, there are $O(2^{(l-\lambda)d})$ such wavelets. Splitting the following sum into these two cases, while separately keeping track of the term $(l,r) = (\lambda,k)$, and using the above bounds gives that for $\lambda \leq J$,

$$\begin{aligned}\|\tilde{\Phi}_{\lambda,k,j}\|_{\mathbb{H}}^2 &= \sum_{l\leq J}\sum_r \sigma_l^{-2} |\langle P_{V_J}[a_\lambda \Phi_{\lambda,k}/\mu_0], \Phi_{l,r}\rangle_{L^2}|^2 \\ &\lesssim a_\lambda^2 \left(\sum_{l=0}^{\lambda} \sigma_l^{-2} 2^{-2\lambda-(\lambda-l)d} + \sum_{l=\lambda+1}^{J} \sigma_l^{-2} 2^{(l-\lambda)d} 2^{-2l-(l-\lambda)d} + \sigma_\lambda^{-2}\right) \\ &\lesssim a_\lambda^2 \left(2^{-(d+2)\lambda} \sum_{l=0}^{\lambda} 2^{2l(\alpha+d)} + \sum_{l=\lambda+1}^{J} 2^{(2\alpha+d-2)l} + 2^{(2\alpha+d)\lambda}\right) \\ &\lesssim a_\lambda^2 2^{(2\alpha+d)J}.\end{aligned}$$

This yields $\sigma_{\Gamma_T}^2 \lesssim a_\lambda^2 2^{(2\alpha+d)J}$. □

We now prove Lemma 5 from above. The first assertion in Part (i) follows from Lemmas 8 and 10 since the envelopes satisfy $|\Gamma_T|_2 = O(a_J) = O((\log T)^{-\eta})$, $\sqrt{T}\varepsilon_T \to \infty$ and $\varepsilon_T \sigma_{\Gamma_T} = o(1)$ for $0 \le \alpha < a \wedge s - d/2$ and the specified choice $2^J \approx T^{1/(2a+d)}$. To prove the first two maximal inequalities, we start by verifying the conditions of Proposition 2 for the case $d \le 4$. Using the envelopes from Lemma 10, we can bound R_T and $\tilde{R}_T$ in Proposition 2 by

$$R_T \lesssim M_T T^{\frac{d - a\wedge s + (d/2+\kappa-1)_+}{2a+d}} (\log T)^{3/2-\eta} \to 0$$

for $a \wedge s > \max(3d/2 - 1, 1)$, $0 < \kappa < a \wedge s - 3d/2 + 1$ ($\kappa = 0$ if $d = 1$) and $M_T \to \infty$ slowly enough, while

$$\tilde{R}_T \lesssim T^{-\frac{1}{2} + \frac{d + (d/2+\kappa-1)_+}{2a+d}} (\log T)^{-2\eta} \sqrt{\log\log T} \to 0 \tag{56}$$

for $a > \max(d-1, 1/2)$ and $0 < \kappa < s - d + 1$ ($\kappa = 0$ if $d = 1$). We may thus apply Proposition 2 to $G(b) = \langle b_j - b_{0,j}, a_\lambda \Phi_{\lambda,k}\rangle_{L^2}$ with $\gamma = \tilde{\Phi}_{\lambda,k,j} \in V_J^{\otimes d}$, so that for $b_u = b - \frac{u}{\sqrt{T}}\tilde{\Phi}_{\lambda,k,j}$, $u \in \mathbb{R}$,

$$E^{\Pi^{D_T}}[e^{u\sqrt{T}\langle b_j - b_{0,j}, a_\lambda \Phi_{\lambda,k}\rangle_{L^2}} | X^T] = e^{\frac{u}{\sqrt{T}}\int_0^T \tilde{\Phi}_{\lambda,k,j}(X_t).dW_t + \frac{u^2}{2}\|\tilde{\Phi}_{\lambda,k,j}\|_{\mu_0}^2 + u r_T + u^2 \tilde{r}_T} \times \frac{\int_{D_T} e^{S_T(b) + \ell_T(b_u)} d\Pi(b)}{\int_{D_T} e^{\ell_T(b)} d\Pi(b)},$$

where $r_T = O_{P_{b_0}}(R_T)$, $\tilde{r}_T = O_{P_{b_0}}(\tilde{R}_T)$ uniformly over $\lambda \le J$ and k, j, and

$$\begin{aligned} S_T(b) &= u\sqrt{T}\langle b_j - b_{0,j}, a_\lambda \Phi_{\lambda,k}\rangle_{L^2} - u\sqrt{T}\langle b_j - b_{0,j}, a_\lambda P_{V_J}[\Phi_{\lambda,k}/\mu_0]\rangle_{\mu_0} \\ &= u\sqrt{T} a_\lambda \langle \mu_0(b_j - b_{0,j}), \Phi_{\lambda,k}/\mu_0 - P_{V_J}[\Phi_{\lambda,k}/\mu_0]\rangle_{L^2}. \end{aligned}$$

Applying the first bound from Lemma 9 with $d \le 4$, $\eta > 1$,

$$\sup_{b \in D_T} |S_T(b)| \lesssim |u|\sqrt{T}(\log T)^{-\eta}(2^{-2J} M_T \varepsilon_T + 2^{-J(s+1+d/2)}) = |u| \times o(1), \tag{57}$$

for $M_T \to \infty$ slowly enough and $s \ge a - 1$. Applying $\alpha|u| \le \alpha^2 u^2 + 1$ for all $\alpha \ge 0$ to (57), and using Lemmas 8 and 10 gives for all $u \in \mathbb{R}$,

$$\frac{\int_{D_T} e^{S_T(b) + \ell_T(b_u)} d\Pi(b)}{\int_{D_T} e^{\ell_T(b)} d\Pi(b)} \le C_T e^{c_T u^2},$$

where $C_T = O_{P_{b_0}}(1)$ and non-random $c_T = o(1)$ are independent of u and uniform over $\lambda \le J$ and k, j.

Setting $Z_{\lambda,k,j} = \langle b_j - b_{0,j}, a_\lambda \Phi_{\lambda,k}\rangle_{L^2} - \frac{1}{T}\int_0^T \tilde{\Phi}_{\lambda,k,\cdot}(X_t).dW_t$ and using again that $\alpha|u| \le \alpha^2 u^2 + 1$ for all $\alpha \ge 0$, the Laplace transform satisfies the conditional subgaussian bound

$$E^{\Pi^{D_T}}[e^{u\sqrt{T}Z_{\lambda,k,j}}|X^T] \le C'_T e^{\frac{u^2}{2}(\|\tilde{\Phi}_{\lambda,k,j}\|^2_{\mu_0} + c'_T)}$$

for sequences $C'_T = O_{P_{b_0}}(1)$ and $c'_T = 2(\tilde{r}_T + r_T^2 + c_T) = o_{P_{b_0}}(1)$, which are independent of u and uniform over $\lambda \le J$ and k, j. Since $\|\tilde{\Phi}_{\lambda,k,j}\|^2_{\mu_0} = a_\lambda^2 \|P_{V_J}[\Phi_{\lambda,k}/\mu_0]\|^2_{\mu_0} \le a_\lambda^2 \|\mu_0\|_\infty \|1/\mu_0\|^2_\infty \le C(b_0)$, standard subgaussian inequalities (Lemmas 2.3.2 and 2.3.4 of [9]) give

$$\begin{aligned}(58)\quad &\sqrt{T}E^{\Pi^{D_T}}[\max_{\lambda \le J,k,j}|Z_{\lambda,k,j}||X^T] \\ &\lesssim \sqrt{2\log 2\dim(V_J^{\otimes d})}(C'_T+1)\max_{\lambda\le J,k,j}(\|\tilde{\Phi}_{\lambda,k,j}\|^2_{\mu_0} + c'_T) = O_{P_{b_0}}(\sqrt{J})\end{aligned}$$

since $\dim(V_J^{\otimes d}) = O(d2^{Jd})$. Similarly, for $\lambda \le J$,

$$\begin{aligned}(59)\quad &\sqrt{T}E^{\Pi^{D_T}}[\max_{k,j}|Z_{\lambda,k,j}||X^T] \\ &\lesssim \sqrt{2\log 2\dim(V_\lambda^{\otimes d})}(C'_T+1)\max_{k,j}(\|\tilde{\Phi}_{\lambda,k,j}\|^2_{\mu_0} + c'_T) = O_{P_{b_0}}(\sqrt{\lambda}).\end{aligned}$$

We now deduce (33) from (58), the same arguments then also show that (32) follows from (59). Decompose

$$\begin{aligned}&E^{\Pi^{D_T}}\left[\max_{\lambda\le J,k,j}\sqrt{T}|\langle b_j - b_{0,j}, a_\lambda\Phi_{\lambda,k}\rangle_{L^2}|\,\middle|\, X^T\right] \\ &\le E^{\Pi^{D_T}}\left[\max_{\lambda \le J,k,j}\sqrt{T}|Z_{\lambda,k,j}|\,\middle|\, X^T\right] + \max_{\lambda\le J,k,j}\left|\frac{1}{\sqrt{T}}\int_0^T \tilde{\Phi}_{\lambda,k,j}(X_t).dW_t\right|\end{aligned}$$

and we have shown the first term is $O_{P_{b_0}}(\sqrt{J})$. We now control the P_{b_0}-expectation of the second term by showing that $M_T^{\lambda,k,j} = \int_0^T \tilde{\Phi}_{\lambda,k,j}(X_t).dW_t$ are subgaussian with uniform constants on a suitable event A_T. For $\epsilon > 0$ fixed, set

$$(60)\qquad A_T = \left\{\max_{\lambda,k,j}\left|\frac{1}{T}\int_0^T \|\tilde{\Phi}_{\lambda,k,j}(X_t)\|^2 dt - \|\tilde{\Phi}_{\lambda,k,j}\|^2_{\mu_0}\right| \le \epsilon\right\}.$$

Applying Markov's inequality, Proposition 2(ii) and (56), $P_{b_0}(A_T^c) \lesssim \epsilon^{-1}\tilde{R}_T \to 0$. On A_T we have

$$T^{-1}[M^{\lambda,k,j}]_T = T^{-1}\int_0^T \|\tilde{\Phi}_{\lambda,k,j}(X_T)\|^2 dt \le \|\tilde{\Phi}_{\lambda,k,j}\|^2_{\mu_0} + \epsilon \le C_0(b_0) + \epsilon,$$

so that applying Bernstein's inequality (15), for any $x > 0$,

$$P_{b_0}(T^{-1/2}|M_T^{\lambda,k,j}|1_{A_T} \ge x) \le P_{b_0}(|M_T^{\lambda,k,j}| \ge x\sqrt{T}, [M^{\lambda,k,j}]_T \le (C_0+\epsilon)T) \le 2e^{-\frac{x^2}{2(C_0+\epsilon)}}.$$

Consequently, $(T^{-1/2}M^{\lambda,k,j}1_{A_T} : \lambda, k)$ are subgaussian random variables with uniformly bounded constants, so that $E_{b_0}\max_{\lambda\le J,k,j} T^{-1/2}|M_T^{\lambda,k,j}|1_{A_T} = O(\sqrt{J})$ by Lemma 2.3.4 of [9]. When $d \ge 5$, one proceeds exactly as above with the only difference to the case $d \le 4$ being that we use the second bound in Lemma 9 with $a \le s + d/2 - 1$ rather than the first bound, which is needed to ensure $\sup_{b\in D_T}|S_T(b)| = o(|u|)$.

For Part (ii), we can invoke Lemma 8(ii) to obtain $\Pi(\bar{D}_T|X^T) \to 1$ in P_{b_0}-probability. The maximal inequality then follows from the same proof as in (i), using Proposition 2(iii), that the $\|\cdot\|_\infty$-contraction rate implies the same rate in L^2-norm, and replacing the bias bound (57) by the third inequality of Lemma 9 so that scaling by a_λ is not necessary. The conditions $s > a-1+d/2$ and $d < 4$ ensure that the terms $\sqrt{T}2^{-J(s+1)}$ and $\sqrt{T}2^{-2J}\|b-b_0\|_\infty$ are both $o(1)$ and hence asymptotically negligible.

5. Proofs for Section 2.5.

5.1. *Proof of Theorem 5.* It suffices to prove the theorem for

$$\sqrt{T}(\mu_b - \mu_{\hat{b}_T})|X^T, \quad b \sim \Pi^{\bar{D}_T}(\cdot|X^T),$$

where $\Pi^{\bar{D}_T}(\cdot|X^T)$ was introduced at the beginning of the proof of Theorem 3 and $\bar{D}_T$ is given by (29). On the set $\bar{D}_T$ we have the estimate

$$\|b\|_{B^1_{\infty\infty}} \lesssim 2^J\|b-b_0\|_\infty + \max_j \|b_{0,j} - P_{V_J}(b_{0,j})\|_{B^1_{\infty\infty}} = O(1)$$

as $T \to \infty$, and the same argument shows $\|\hat{b}_T\|_{B^1_{\infty\infty}} = O_{P_{b_0}}(1)$ by virtue of Corollary 1. Proposition 1 then further implies that $\|\mu_b\|_{Lip}, \|\mu_{\hat{b}_T}\|_{Lip}$ are also $O(1)$ and $O_{P_{b_0}}(1)$, respectively – these bounds will be used repeatedly in the proof without further mention. Recall that $v_h = (L_b^*)^{-1}[f_h]$ and $w_{b,h} = (L^*_{b+h})^{-1}[\bar{f}_h]$ for $f_h = -\sum_{j=1}^d \frac{\partial}{\partial x_j}(h_j\mu_b)$ and $\bar{f}_h = \sum_{j=1}^d \frac{\partial}{\partial x_j}[h_j v_h]$. We will use the decomposition (45) with $h = \hat{b}_T - b$. First, for the 'remainder' term, we can use (73) below and (3) to deduce that, uniformly in $\|g\|_{\mathbb{B}_r} \le 1$,

$$
\begin{aligned}
\Big| \int_{\mathbb{T}^d} w_{b,h} g \Big| &\le \|g\|_{L^2} \|(L^*_{b+h})^{-1}[\bar{f}_h]\|_{L^2} \\
&\lesssim \|\bar{f}_h\|_{H^{-2}} = \sup_{\|\phi\|_{H^2}\le 1} \Big| \sum_{j=1}^d \int_{\mathbb{T}^d} \phi \frac{\partial}{\partial x_j}[h_j v_h] \Big| \\
&\le \sup_{\|\phi\|_{H^2}\le 1} \Big| \sum_{j=1}^d \int_{\mathbb{T}^d} h_j v_h \frac{\partial}{\partial x_j}\phi \Big| \lesssim \|h\|_\infty \|(L^*_b)^{-1}[f_h]\|_{L^2} \qquad (61) \\
&\lesssim \|h\|_\infty \|f_h\|_{H^{-2}} \lesssim \|h\|_\infty \|h\mu_b\|_{L^2} \lesssim \|h\|_\infty^2
\end{aligned}
$$

is $O_{P_{b_0}}((\|\hat{b}_T - b_0\|_\infty + \|b - b_0\|_\infty)^2) = o_{P_{b_0}}(1/\sqrt{T})$ on $\bar{D}_T$ and by Corollary 1. For the 'linear' term we may write, noting the dependence $f_h = f_{h,b}$ on b,

$$\int v_{b,h} g = \int (L^*_{b_0})^{-1}[f_{h,b_0}]g + \int (L^*_{b_0})^{-1}[f_{h,b} - f_{h,b_0}]g + \int [(L^*_b)^{-1} - (L^*_{b_0})^{-1}][f_{h,b}]g$$

and we denote the right hand side as $A_0 + A_1 + A_2$. The last term A_2 is $o_{P_{b_0}}(1/\sqrt{T})$ in $\mathbb{B}_r^*$ since $[(L^*_b)^{-1} - (L^*_{b_0})^{-1}][f_{h,b}]$ can be written as $-(L^*_b)^{-1}[(b-b_0).\nabla\omega + div(b-b_0)\omega]$ for $\omega = (L^*_{b_0})^{-1}[f_{h,b}]$ (arguing just as in (43)), so that using (73) gives (as in (61)) the inequality

(62)

$$\|[(L^*_b)^{-1} - (L^*_{b_0})^{-1}][f_h]\|_{L^2} \lesssim \|b - b_0\|_\infty (\|\hat{b}_T - b_0\|_\infty + \|b - b_0\|_\infty) = o_{P_{b_0}}(1/\sqrt{T}).$$

Similarly the term A_1 can be bounded in $\mathbb{B}_r^*$ by

$$\|(L^*_{b_0})^{-1}[f_{h,b} - f_{h,b_0}]\|_{L^2} \lesssim \|f_{h,b} - f_{h,b_0}\|_{H^{-2}} \lesssim \|h\|_\infty \|b - b_0\|_\infty = o_{P_{b_0}}(1/\sqrt{T}).$$

Finally, for the term A_0, we show that the linear operator

$$h \mapsto v_{b_0,h} = (L^*_{b_0})^{-1}[f_{h,b_0}]$$

is Lipschitz on $C^1(\mathbb{T}^d)$ for the norms $(B^{r+1}_{1\infty})^*$ and $\mathbb{B}_r^*$ for any $d/2-1<r<1$. Using that $v_{b_0,h}\in L^2_0(\mathbb{T}^d)$ and writing $\bar g = g-\int g d\mu_0$,

$$\begin{aligned}
&\|v_{b_0,h}\|_{\mathbb{B}_r^*}\\
&= \sup_{\|g\|_{L^2}+\|g\|_{B^r_{1\infty}}\le 1}\left|\int_{\mathbb{T}^d} g v_{b_0,h}\right| = \sup_{\|g\|_{L^2}+\|g\|_{B^r_{1\infty}}\le 1}\left|\int_{\mathbb{T}^d} L_{b_0}L_{b_0}^{-1}[\bar g] v_{b_0,h}\right|\\
&= \sup_{\|g\|_{L^2}+\|g\|_{B^r_{1\infty}}\le 1}\left|\int_{\mathbb{T}^d} L^*_{b_0} v_{b_0,h} L_{b_0}^{-1}[\bar g]\right|\\
&= \sup_{\|g\|_{L^2}+\|g\|_{B^r_{1\infty}}\le 1}\left|\sum_{j=1}^d\int_{\mathbb{T}^d} h_j\big(\frac{\partial}{\partial x_j}L_{b_0}^{-1}[\bar g]\big)\mu_0\right|\\
(63)\quad &= \sup_{\|g\|_{L^2}+\|g\|_{B^r_{1\infty}}\le 1}\left|\langle h, \mu_0\nabla L_{b_0}^{-1}[\bar g]\rangle_{L^2}\right|\\
&\lesssim \|\mu_0\|_{C^{r+1}} \sup_{\|g\|_{L^2}+\|g\|_{B^r_{1\infty}}\le 1}\|\nabla L_{b_0}^{-1}[\bar g]\|_{B^{r+1}_{1\infty}} \sup_{\tilde g:\|\tilde g\|_{B^{r+1}_{1\infty}}\le 1}|\langle h,\tilde g\rangle_{L^2}| \lesssim \|h\|_{(B^{r+1}_{1\infty})^*},
\end{aligned}$$

where we have used (3), (74) below and that ∇ maps $B^{r+2}_{1\infty}$ continuously into $B^{r+1,\otimes d}_{1\infty}$. We also used that $\mu_0\in C^t(\mathbb{T}^d), t\ge 1$, whenever $b_0\in C^s\cap H^s, s> t+d/2-1$. Indeed, $\|\mu_0\|_{H^1}\lesssim\|\mu_0\|_{Lip}\le C(d,\|b_0\|_\infty)$ (by Proposition 1 and Rademacher's theorem) allows an iterated application of the inequality (72) below with $u=\mu_0$ to bound $\|\mu_0\|_{H^{s+1}}$ by a constant $C(d,s,\|b_0\|_{C^s})$, which in turn bounds $\|\mu_0\|_{C^t}$ by the Sobolev embedding theorem.

Summarizing, with $h=b-\hat b_T$ we have proved uniformly in $\|g\|_{\mathbb{B}_r}\le 1$,

$$\sqrt{T}\int_{\mathbb{T}^d}(\mu_b-\mu_{\hat b_T})g = \int_{\mathbb{T}^d} v_{b_0,\sqrt{T}(b-\hat b_T)}g + o_{P_{b_0}}(1)$$

and that the linear operator $h\mapsto v_{b_0,h}$ is continuous from $(C^1(\mathbb{T}^d),\|\cdot\|_{(B^{r+1}_{1\infty})^*})$ to $\mathbb{B}_r^*$. Theorem 5 now follows from Theorem 3 with $r+1=\rho$ and the continuous mapping theorem for weak convergence applied to $\sqrt{T}(b-\hat b_T)$. We note that the calculation leading to (63) shows that the covariance of the limiting Gaussian process is the one of the Gaussian process $g\mapsto \mathbb{W}_0(\mu_0\nabla L_{b_0}^{-1}[\bar g])$, $\mathbb{W}_0\sim\mathcal{N}_{b_0}$, of the required form. In particular, $\mathcal{N}_{\mu_0}$ exists as a tight Gaussian probability measure in $\mathbb{B}_r^*$ as the image of $\mathcal{N}_{b_0}$ under the continuous map $v_{b_0,\cdot}$. The limit of the MAP-estimate follows from similar (in fact simpler) arguments and Theorem 4, and is left to the reader.

5.2. *Proof of Theorem 6.* We finally prove Theorem 6 and explain the necessary modifications to the arguments from the proof of Theorem 5. As in the proof of Theorem 5, one shows that $\hat{b}_T, b$ are (in the former case, stochastically) bounded in $B^1_{\infty\infty}$ on the set $\bar{D}_T$, and so are then $\mu_{\hat{b}_T}, \mu_b$ by Proposition 1. Using the Sobolev-embedding $H^1(\mathbb{T}) \subset C(\mathbb{T})$ and then repeatedly Lemma 12, (73) and the basic interpolation inequality $\|g\|_{H^1} \lesssim \|g\|^{1/2}_{H^2}\|g\|^{1/2}_{L^2}$, the second term in the decomposition (45) can be bounded by

$$\|(L^*_{b+h})^{-1}[\bar{f}_h]\|_{H^1} \lesssim \|\bar{f}_h\|^{1/2}_{L^2}\|\bar{f}_h\|^{1/2}_{H^{-2}} \lesssim \|h\|_{H^1}\|h\|_{L^2},$$

which for $a > 3/2$ and $h = b - \hat{b}_T = b - b_0 - (\hat{b}_T - b_0), b \sim \Pi^{\bar{D}_T}(\cdot|X^T)$, is of order $\|h\|_{H^1}\|h\|_{L^2} = o_P(1/\sqrt{T})$ since $\|h\|_{H^1} \lesssim 2^J\|h\|_{L^2}$ for $h \in V_J$. The linear term in (45) can be decomposed as

$$(L^*_{b_0})^{-1}[f_h](x) - [(L^*_{b_0})^{-1} - (L^*_b)^{-1}][f_h](x).$$

Then arguing as before (62) and using the Sobolev embedding $H^1 \subset C(\mathbb{T})$ as well as Lemma 12, the second term is bounded, for $a > 3/2$, by

$$\|[(L^*_b)^{-1}-(L^*_{b_0})^{-1}][f_h]\|_\infty \lesssim \|b-b_0\|_\infty(\|\hat{b}_T-b_0\|_{H^1}+\|b-b_0\|_{H^1}) = o_{P_{b_0}}(1/\sqrt{T}).$$

Similarly, noting the dependence $f_h = f_{h,b}$ on b, the term $(L^*_{b_0})^{-1}[f_{h,b}] - (L^*_{b_0})^{-1}[f_{h,b_0}]$ can be shown to be $o_{P_{b_0}}(1/\sqrt{T})$ in $C(\mathbb{T})$.

We next establish continuity of the linear operator

$$h \mapsto v_{b_0,h} = (L^*_{b_0})^{-1}[f_{h,b_0}]$$

on $C^1(\mathbb{T})$ for the norms of $(B^1_{1\infty}(\mathbb{T}))^*$ and $C(\mathbb{T})$, so that the theorem follows from Theorem 3 and the continuous mapping theorem for weak convergence, just as in the proof of Theorem 5. We use a dual representation for the weighted wavelet sequence norms characterising Besov spaces – more precisely, that the classical identities $(c_0)^* = \ell_1, (\ell_1)^* = \ell_\infty$, where $c_0 = \{(a_k) : \lim_{k\to\infty} a_k = 0\}$ is equipped with the supremum-norm on sequences, imply

$$\|g\|_{B^0_{\infty 1}} \lesssim \sup_{\phi\in C(\mathbb{T}):\|\phi\|_{B^0_{1\infty}}\le 1} |\langle g, \phi\rangle_{L^2}|, \quad g \in C(\mathbb{T}).$$

Moreover (3) and (74) imply

$$\sup_{\|\phi\|_{B^0_{1\infty}}\le 1} \|\mu_0 \frac{d}{dy} L^{-1}_{b_0}[\bar{\phi}]\|_{B^1_{1\infty}} \lesssim \|\mu_0\|_{Lip} \sup_{\|\phi\|_{B^0_{1\infty}}\le 1} \|L^{-1}_{b_0}[\bar{\phi}]\|_{B^2_{1\infty}} < \infty,$$

which will be used in the following estimate. For $\bar\phi = \phi - \int \phi d\mu_{b_0}$, and since $v_{b_0,h} \in L^2_0(\mathbb{T}) \cap H^2 \subset C(\mathbb{T})$ in view of Lemma 12 below,

$$\begin{aligned}\|v_{b_0,h}\|_\infty \lesssim \|v_{b_0,h}\|_{B^0_{\infty 1}} &\lesssim \sup_{\phi\in C(\mathbb{T}):\|\phi\|_{B^0_{1\infty}}\le 1} \left|\int v_{b_0,h}\phi\right| \\ &= \sup_{\phi\in C(\mathbb{T}):\|\phi\|_{B^0_{1\infty}}\le 1} \left|\int (L^*_{b_0})^{-1}[f_{h,b_0}] L_{b_0} L^{-1}_{b_0}[\bar\phi]\right| \\ &= \sup_{\phi\in C(\mathbb{T}):\|\phi\|_{B^0_{1\infty}}\le 1} \left|\int f_{h,b_0} L^{-1}_{b_0}[\bar\phi]\right| \\ &\le \sup_{\|\phi\|_{B^0_{1\infty}}\le 1} \|\mu_0 \frac{d}{dy} L^{-1}_{b_0}[\bar\phi]\|_{B^1_{1\infty}} \|h\|_{(B^1_{1\infty})^*} \lesssim \|h\|_{(B^1_{1\infty})^*}.\end{aligned}$$

The covariance of the limiting Gaussian process is obtained as follows: since $G_{b_0}(x,y)$ is the periodic Green kernel of $L^{-1}_{b_0}$, the Green kernel of $(L^*_{b_0})^{-1}$ is $G_{b_0}(y,x)$, and thus by the definitions and integration by parts,

$$(L^*_{b_0})^{-1}[f_{h,b_0}] = -\int_{\mathbb{T}} G_{b_0}(y,\cdot)\frac{d}{dy}[h\mu_0](y)dy = \int_{\mathbb{T}} \frac{d}{dy}G_{b_0}(y,\cdot)h(y)\mu_0(y)dy.$$

Inserting for h the limit $\mathbb{W}_0 \sim \mathcal{N}_{b_0}$ of $\sqrt{T}(b-\hat b_T)$ gives the desired form of the limiting covariance. Finally, the limit distribution of the MAP estimate follows from the same (in fact simpler) arguments and Theorem 4.

6. Appendix: Some basic facts on the elliptic PDEs involved.

We record here some basic facts about elliptic PDEs and refer to, e.g., Chapter II.3 in [2] as a reference for standard background material in the periodic setting considered here. The generator $L = L_b$ of the diffusion process given in (5) is a strongly elliptic second order partial differential operator. We will suppress the dependence on b in most of what follows; all that is required is that b is 'smooth enough', and $b \in V_J^{\otimes d}$ for a S-regular wavelet basis with S large enough will be sufficient throughout. The maximum principle for elliptic operators (see [2, 8]) implies that any (strong and then also weak) periodic solution of the Laplace equation

$$Lu = 0 \text{ on } \mathbb{T}^d \tag{64}$$

equals a constant. The adjoint operator $L^* = L^*_b$ was defined in (6), and in the periodic setting considered here the operators (L, L^*) form a Fredholm pair on $L^2(\mathbb{T}^d)$, see p.175f. in [2]. As a consequence, the inhomogeneous equation

$$Lu = f, \quad f \in L^2(\mathbb{T}^d), \tag{65}$$

has a solution u if and only if $\langle f, m\rangle_{L^2} = 0$ for every solution $m \in L^2(\mathbb{T}^d)$ of

$$L^* m = 0 \text{ on } \mathbb{T}^d. \tag{66}$$

By the Fredholm property the kernel of L^* has the same dimension as the kernel of L and inspection of the form of L^* shows that the solutions $m \in L^2(\mathbb{T}^d)$ to (66) are determined up to a normalising constant. It follows that

$$L^* m = 0 \iff m \in \mathcal{K} = \{c\mu : c \in \mathbb{R}\}, \tag{67}$$

where $\mu > 0$ is the unique solution m ('invariant measure') satisfying $\int_{\mathbb{T}^d} m = 1$. Positivity of μ can be deduced from appropriate heat kernel estimates: in fact (arguing, e.g., as on p.167f. in [11]) the solution μ can be seen to be Lipschitz continuous and bounded away from zero on $\mathbb{T}^d$, and $\|\mu\|_{Lip}$ is bounded by a fixed constant that only depends on d and on an upper bound for $\|b\|_\infty$, proving in particular Proposition 1.

We can now state the following basic result for the PDE (65).

LEMMA 11. *Let $t \geq 2$ and assume $b \in C^{t-2}(\mathbb{T}^d)$. For any $f \in L^2_\mu(\mathbb{T}^d)$, there exists a unique solution $L_b^{-1}[f] \in L^2_0(\mathbb{T}^d)$ of equation (65) satisfying $L_b L_b^{-1}[f] = f$ almost everywhere. Moreover,*

$$\|L_b^{-1}[f]\|_{H^t} \lesssim \|f\|_{H^{t-2}},$$

with constants depending on t, d and on an upper bound B for $\|b\|_{B^{t-2}_{\infty\infty}}$.

PROOF. By standard Sobolev space theory and definition of the Laplacian we have for any $u \in \mathcal{H} \equiv H^t \cap \{u : \langle u, 1\rangle_{L^2} = 0\}$ the inequality

$$\|u\|_{H^t} \lesssim \|\Delta u\|_{H^{t-2}}. \tag{68}$$

Indeed, for $\{e_k : k = (k_1, \dots, k_d) \in \mathbb{Z}^d\}$ the usual trigonometric basis of $L^2(\mathbb{T}^d)$ we have $\langle u, e_0\rangle_{L^2} = \langle u, 1\rangle_{L^2} = 0$, $\langle \Delta u, e_k\rangle = -(2\pi)^2 \sum_j k_j^2 \langle u, e_k\rangle$, and $\sup_{k\neq 0}(1 + \|k\|^2)/\|k\|^2 < \infty$, which gives the result using the characterisation of Sobolev norms in the basis $\{e_k\}$. We then also have, by the triangle inequality and (3),

$$\|u\|_{H^t} \lesssim \|Lu\|_{H^{t-2}} + \|b.\nabla u\|_{H^{t-2}} \lesssim \|Lu\|_{H^{t-2}} + \|b\|_{B^{t-2}_{\infty\infty}} \|u\|_{H^{t-1}} \tag{69}$$

for all $u \in \mathcal{H}$, with constants depending on t, d. We now deduce from this the inequality

$$\|u\|_{H^t} \lesssim \|Lu\|_{H^{t-2}} \quad \forall u \in \mathcal{H}. \tag{70}$$

Indeed, if the latter inequality does not hold true, then there exists a sequence $u_m \in \mathcal{H}$ such that $\|u_m\|_{H^t} = 1$ for all m but $\|Lu_m\|_{H^{t-2}} \to 0$ as $m \to \infty$. At the same time, by compactness, u_m converges in $\|\cdot\|_{H^{t-1}}$-norm (if necessary along a subsequence) to some $u \in \mathcal{H}$ satisfying $Lu = 0$. Using (69) with fixed constant depending only on B, t, d, we see that u_m is also Cauchy in H^t, and its limit must necessarily satisfy $\|u\|_{H^t} = 1$. However, as remarked after (64), the only solution $u \in \mathcal{H}$ to $Lu = 0$ on $\mathbb{T}^d$ equals $u = const = 0$, a contradiction to $\|u\|_{H^t} = 1$, proving (70).

By the Fredholm property and (67), a solution u_f to (65) exists whenever $\int f d\mu = 0$, and for $f \in H^{t-2}(\mathbb{T}^d)$ any such solution belongs to $H^t(\mathbb{T}^d)$ (see Theorem 3.5.3 in [2], which is proved for smooth b, but the proof remains valid for $b \in C^{t-2}(\mathbb{T}^d)$). The weak maximum principle (p.179 in [8]) now implies that u_f is unique up to an additive constant, and applying (70) to the unique selection $u_f = L^{-1}[f] \in \mathcal{H}$ completes the proof. □

We next obtain corresponding results for the adjoint PDE. It follows from (67) that the unique element $m \in \mathcal{K}$ satisfying $\int_{\mathbb{T}^d} m = 0$ must necessarily vanish identically, and we can study the solution operator $(L^*)^{-1}$ of the inhomogeneous adjoint PDE

$$L^* u = f \text{ on } \mathbb{T}^d, \tag{71}$$

which assigns to any $f \in L^2_0(\mathbb{T}^d)$ the unique solution $u = (L^*)^{-1}[f] \in L^2_0(\mathbb{T}^d)$. Indeed, using the Fredholm property from Section 3.6 in [2] in a reverse way (with L equal to our L^* so that the new L^* is our $(L^*)^* = L$), we see that solutions $u = u_f$ to (71) exist for any periodic f for which $\int f = 0$ (since solutions to $Lu = 0$ equal constants), and if u_1, u_2 are two such solutions, so that $L^*(u_1 - u_2) = 0$ and $\int u_1 = \int u_2$, then necessarily $u_1 = u_2$ by what precedes.

LEMMA 12. *Let $t \geq 2$ and assume $b \in C^{t-1}(\mathbb{T}^d)$. Then for any $f \in H^{t-2}(\mathbb{T}^d) \cap L^2_0(\mathbb{T}^d)$, we have $(L_b^*)^{-1}[f] \in H^t(\mathbb{T}^d)$ and*

$$\|(L_b^*)^{-1}[f]\|_{H^t} \lesssim \|f\|_{H^{t-2}},$$

with constants depending on t, d and on an upper bound B for $\|b\|_{B^{t-1}_{\infty\infty}}$.

PROOF. The proof is similar to the one of Lemma 11 after deriving the basic inequality

$$\begin{aligned} \|u\|_{H^t} &\lesssim \|L^* u\|_{H^{t-2}} + \|b.\nabla u + div(b)u\|_{H^{t-2}} \\ &\lesssim \|L^* u\|_{H^{t-2}} + \|b\|_{B^{t-2}_{\infty\infty}} \|u\|_{H^{t-1}} + \|b\|_{B^{t-1}_{\infty\infty}} \|u\|_{H^{t-2}} \end{aligned} \tag{72}$$

in analogy to (69). □

We can also give a version of Lemma 12 with $t=0$. Since $(L_b^*)^{-1}[f] \in L^2_0(\mathbb{T}^d)$ we have for all $f \in L^2_0(\mathbb{T}^d)$ and $\bar\phi = \phi - \int \phi d\mu_b$ the estimate

$$\|(L_b^*)^{-1}[f]\|_{L^2} = \sup_{\|\phi\|_{L^2}\le 1} \left| \int (L_b^*)^{-1}[f] L_b L_b^{-1}[\bar\phi] \right| = \sup_{\|\phi\|_{L^2}\le 1} \left| \int f L_b^{-1}[\bar\phi] \right|$$
$$\le \|f\|_{H^{-2}} \sup_{\|\phi\|_{L^2}\le 1} \|L_b^{-1}[\bar\phi]\|_{H^2} \lesssim \|f\|_{H^{-2}}, \tag{73}$$

where we have used Lemma 11 with $t=2$ in the last inequality and where the constants in the last inequality depend only on d and on bounds for $\|b\|_{B^1_{\infty\infty}}$ and $\|\mu_b\|_{L^2}$.

6.0.1. *Refinements on the Besov scale.* For the the proofs of Theorems 5 and 6 we need more refined regularity estimates for the solutions of the PDE involved, replacing the Sobolev norms in Lemma 11 by appropriate Besov norms. The inequality

$$\|L_b^{-1}[f]\|_{B^t_{1\infty}} \lesssim \|f\|_{B^{t-2}_{1\infty}},\ t-2\ge 0, \forall f \in L^2_{\mu_b}(\mathbb{T}^d), \tag{74}$$

with constants depending on b only via a bound B for $\|b\|_{B^{t-1}_{\infty\infty}}$, is proved in the same way as Lemma 11, replacing the basic inequality (68) by its analogue for Besov norms

$$\|u\|_{B^t_{1\infty}} \lesssim \|\Delta u\|_{B^{t-2}_{1\infty}} \quad \forall u \in B^t_{1\infty} \cap \{u : \langle u, e_0\rangle = 0\}, \tag{75}$$

which is proved as follows: for all u such that $\langle u, e_0\rangle = 0$ and $\mathbb{N}_0 = \mathbb{N}\cup\{0\}$, an equivalent Littlewood-Paley norm on any Besov space $B^r_{1\infty}$ is given by

$$\|u\|_{B^r_{1\infty}} = \sup_{j\in\mathbb{N}_0} 2^{jr} \Big\| \sum_{k\in\mathbb{Z}, k\neq 0} \psi_j(k)\langle u, e_k\rangle e_k \Big\|_{L^1(\mathbb{T}^d)},$$

where the $\psi_j = \psi(\cdot/2^j), \mathrm{supp}(\psi) \in (1/2,2)^d$ form a Littlewood-Paley resolution of unity, see p.162f. in [12]. Then as after (68),

$$\|u\|_{B^t_{1\infty}} = \sup_{j\in\mathbb{N}_0} 2^{jt} \Big\| \sum_{k\in\mathbb{Z}, k\neq 0} \frac{1}{4\pi^2\|k\|^2} \psi_j(k)\langle \Delta u, e_k\rangle e_k \Big\|_{L^1(\mathbb{T}^d)}$$
$$= \sup_{j\in\mathbb{N}_0} 2^{j(t-2)} \Big\| \sum_{k\in\mathbb{Z}, k\neq 0} M_j(k)\psi_j(k)\langle \Delta u, e_k\rangle e_k \Big\|_{L^1(\mathbb{T}^d)}$$

where $M_j = M(\cdot/2^j)$ and $M = \Phi/(4\pi^2\|\cdot\|^2)$ with Φ a smooth function supported in $(1/4, 9/4)^d$ such that $\Phi = 1$ on $(1/2,2)^d$. By a standard Fourier multiplier inequality (e.g., Lemma 4.3.27 in [9], which easily generalises to $d > 1$) the last norm can be estimated by

$$\sup_{j\in\mathbb{N}_0} 2^{j(t-2)} \Big\| \sum_{k\in\mathbb{Z},k\neq 0} \psi_j(k)\langle \Delta u, e_k\rangle e_k \Big\|_{L^1(\mathbb{T}^d)} \times \|F^{-1}M_j\|_{L^1(\mathbb{R}^d)},$$

where F^{-1} is the inverse Fourier transform. Since Φ is smooth and supported in $(-1/4, 3/4)^d$, both M and $F^{-1}M$ belong to the Schwartz-class $\mathcal{S}$, so that (75) follows from

$$\sup_j \|F^{-1}[M_j]\|_{L^1(\mathbb{R}^d)} = \|F^{-1}[M]\|_{L^1(\mathbb{R}^d)} < \infty.$$

References.

[1] Bass, R. F. *Stochastic processes.* Cambridge Univ. Press, Cambridge, 2011.

[2] Bers, L., John, F., and Schechter, M. *Partial differential equations.* Lectures in Applied Mathematics, Vol. III. Wiley & Sons, Inc. New York-London-Sydney, 1964.

[3] Castillo, I. On Bayesian supremum norm contraction rates. *Ann. Statist. 42*, 5 (2014), 2058–2091.

[4] Castillo, I., and Nickl, R. On the Bernstein–von Mises phenomenon for nonparametric Bayes procedures. *Ann. Statist. 42*, 5 (2014), 1941–1969.

[5] Castillo, I., and Rousseau, J. A Bernstein–von Mises theorem for smooth functionals in semiparametric models. *Ann. Statist. 43*, 6 (2015), 2353–2383.

[6] Ethier, S. N., and Kurtz, T. G. *Markov processes, Characterization and convergence.* John Wiley & Sons, Inc., New York, 1986.

[7] Ghosal, S., and van der Vaart, A. W. *Fundamentals of Nonparametric Bayesian Inference.* Cambridge University Press, New York, 2017.

[8] Gilbarg, D., and Trudinger, N. S. *Elliptic partial differential equations of second order.* Classics in Mathematics. Springer-Verlag, Berlin, 2001. Reprint of the 1998 edition.

[9] Giné, E., and Nickl, R. *Mathematical foundations of infinite-dimensional statistical models.* Cambridge University Press, New York, 2016.

[10] Nickl, R., and Ray, K. Nonparametric statistical inference for drift vector fields of multi-dimensional diffusions. 2019.

[11] Norris, J. R. Long-time behaviour of heat flow: global estimates and exact asymptotics. *Arch. Rational Mech. Anal. 140*, 2 (1997), 161–195.

[12] Schmeisser, H.-J., and Triebel, H. *Topics in Fourier analysis and function spaces.* A Wiley-Interscience Publication. John Wiley & Sons, Ltd., Chichester, 1987.

Statistical Laboratory
Department of Pure Mathematics
and Mathematical Statistics
University of Cambridge
CB3 0WB Cambridge
United Kingdom
E-mail: r.nickl@statslab.cam.ac.uk

Department of Mathematics
King's College London
Strand
London WC2R 2LS
United Kingdom
E-mail: kolyan.ray@kcl.ac.uk